"INTRODUCCION A LA TEORIA DE PROBABILIDADES Y DEDUCCION DE LA DISTRIBUCION BINOMIAL"

Ruiz Gale, María Fernanda.

Dasso, Sergio Ricardo.

Buenos Aires, Argentina.

Febrero, 1995.

Facultad de Ciencias Exactas y Naturales, U.B.A.

" Todo el mundo cree en ella, me decía un día Lippmann, pues los experimentadores se imaginan que es un teorema de matemáticas y los matemáticos que es un hecho experimental".

Sobre la LEY de GAUSS.

Henri Poincaré.

RESUMEN

En la primera parte de este trabajo se introducen algunos conceptos básicos afines a la teoría de probabilidad tales como espacio muestral, suceso, álgebra booleana, medida de probabilidad, variable aleatoria (discreta, contínua y mixta) y función de distribución acumulativa.

Luego, haciendo uso de estos conceptos, se realiza un desarrollo que nos permite mostrar los fundamentos de la Distribución Binomial.

CONTENIDOS

1. INTRODUCCION

Muchas veces nos asombramos al descubrir cómo han nacido ciertas teorías que en principio fueron utilizadas para resolver problemas que nada tienen que ver con la "ciencia". Una de estas teorías, que comenzó a formarse a partir de una disputa entre jugadores, es el Cálculo de Probabilidades, a partir de la cual se han desarrollado la estadística, la teoría de errores, la teoría de juegos de azar, y muchas otras.

Históricamente el Cálculo de Probabilidades estuvo lleno de controversias y ambigüedades, incluso hubo dos grandes interpretaciones filosóficas opuestas acerca del carácter objetivo o subjetivo del concepto de probabilidad. Recién a principios del siglo XX se llegaron a establecer claramente las bases axiomáticas de esta rama de la matemática.

Es importante dejar en claro que la teoría de probabilidades es un modelo matemático *no determinista* o *estocástico*. Estos están en contraposición con los modelos determinísticos que estipulan que las condiciones bajo las cuales se realiza un experimento determinan el resultado del mismo.

2.1. ESPACIOS MUESTRALES

Si para representar algún fenómeno utilizamos un modelo no determinista, y al realizar cierto experimento relacionado con dicho fenómeno conocemos la totalidad de los resultados que pueden ocurrir, se define como espacio muestral $\mathbb{E}$ al conjunto de todos los resultados posibles. Así, dado un experimento, cada resultado posible será un elemento del conjunto $\mathbb{E}$. Entonces, para poder realizar un modelo estocástico de la realidad, debemos poder definir con certeza *todos* los resultados posibles del experimento en cuestión.

En la teoría de probabilidad es común referirnos a una colección determinada de resultados posibles, es decir a un subconjunto del espacio muestral, llamaremos *suceso* a este subconjunto. Definido un suceso $\mathbb{B}$, al realizar el experimento obtendremos un resultado x, diremos que ocurrió el suceso $\mathbb{B}$ si $x \in \mathbb{B}$. Si en cambio tenemos que $x \notin \mathbb{B}$ diremos que el suceso $\mathbb{B}$ no ha ocurrido.

2.2. ALGEBRA BOOLEANA

Definido el espacio muestral, es frecuente realizar operaciones con subconjuntos (sucesos) de éste: unión, intersección, complemento.

Si tomamos una clase $\mathbb{A}$ de subconjuntos[1] de un Universal $\mathbb{S}$, sería interesante garantizar que $\mathbb{A}$ sea cerrado con respecto a las operaciones mencionadas. Para esto definimos un *álgebra booleana* de conjuntos.

Una clase no vacía $\mathbb{A}$ de subconjuntos de un Universal $\mathbb{S}$ se llama álgebra booleana si para todo par de subconjuntos $\mathbb{B}$ y $\mathbb{C}$ tal que $\mathbb{B} \in \mathbb{A}$ y $\mathbb{C} \in \mathbb{A}$ se cumple :

$$\mathbb{B} \cup \mathbb{C} \in \mathbb{A} \qquad \text{y} \qquad \mathbb{S} - \mathbb{B} \in \mathbb{A}$$ [2]

Estas dos propiedades implican que tanto el conjunto vacío como el Universal $\mathbb{S}$ pertenecen a $\mathbb{A}$ [3].

2.3. MEDIDAS CON ADITIVIDAD FINITA

Existen funciones que poseen ciertas características en común:

1 - Están definidas sobre una colección $\mathbb{A}$ de conjuntos, es decir que cada elemento del dominio de la función es un conjunto que pertenece a la clase $\mathbb{A}$.

2 - Su imagen pertenece a los números reales.

[1]Una clase de subconjuntos es una colección de subconjuntos.
[2]El complemento de B con respecto a S pertenece a la clase A.
[3]Ver Apéndice.

3 - Cumplen con la propiedad aditiva :

$f: \mathbb{A} \text{---} > \mathbb{R}$ es de aditividad finita si

$f(\mathbb{B} \cup \mathbb{C}) = f(\mathbb{B}) + f(\mathbb{C})$

para todos los conjuntos *disjuntos* $\mathbb{B}$ y $\mathbb{C}$ que pertenezcan a la clase $\mathbb{A}$.

4 - Cumplen con la condición

$f(\mathbb{B}) \geq 0 \quad \forall \quad \mathbb{B} \in \mathbb{A}$

Si f cumple con estas cuatro propiedades se dice que $f: \mathbb{A} \text{---} > \mathbb{R}$ es una medida con aditividad finita o simplemente una medida.

2.4. DEFINICION DE PROBABILIDAD PARA ESPACIOS MUESTRALES FINITOS

En la teoría de probabilidad, al Universal $\mathbb{S}$ se lo asocia con el espacio muestral $\mathbb{E}$, es decir que $\mathbb{S} = \mathbb{E}$.

Eligiendo una clase $\mathbb{A}$ de subconjuntos de un espacio muestral finito $\mathbb{E}$ tales que formen un álgebra de Boole, se define a la medida de probabilidad, o simplemente probabilidad, a la función de conjunto P definida sobre $\mathbb{A}$ que cumple con las siguientes propiedades:

1 - P es una medida con aditividad finita.

2 - $P(\mathbb{E}) = 1$

Tenemos entonces que para definir una medida de probabilidad es necesario definir tres conceptos :

- El espacio muestral $\mathbb{E}$.

- El álgebra de Boole $\mathbb{A}$, esto es, la colección de subconjuntos de $\mathbb{E}$ que serán los sucesos considerados.

- La propia función de conjunto P.

A la terna ($\mathbb{E}$, $\mathbb{A}$, P) se la denomina frecuentemente *Espacio de Probabilidad*.

2.5. CONJUNTOS NUMERABLES Y NO NUMERABLES

Dos conjuntos $\mathbb{B}$ y $\mathbb{C}$ están en correspondencia uno a uno si existe cierta función f con las siguientes propiedades:

- El Dominio de f es $\mathbb{B}$. El recorrido de f es $\mathbb{C}$.

- Si x e y son elementos distintos de $\mathbb{B}$, se implica que $f(x)$ y $f(y)$ son elementos distintos de $\mathbb{C}$.

Dos conjuntos $\mathbb{B}$ y $\mathbb{C}$ que están en correspondencia uno a uno se llaman también conjuntos *equivalentes* y se denota con $\mathbb{B} \sim \mathbb{C}$.

Un conjunto $\mathbb{B}$ se llama finito y se dice que contiene n elementos si $\mathbb{B} \sim \{1,2, \ldots ,n\}$.

El conjunto vacío también se considera finito.

Un conjunto $\mathbb{B}$ se llama infinito numerable si $\mathbb{B} \sim \{1,2,3, \ldots \}$, es decir que $\mathbb{B}$ es equivalente al conjunto de los números naturales.

Los números naturales nos ayudan a *marcar* los elementos del conjunto $\mathbb{B}$ y todos los elementos reciben marcas.

Un conjunto se dice numerable en sentido amplio si es finito o infinito numerable.

Existen conjuntos que son *no* numerables.

2.6. PROPIEDAD DE ADITIVIDAD NUMERABLE

Para toda colección infinita numerable de elementos de $\mathbb{A}$, $\{\mathbb{B}_1,\mathbb{B}_2,\ldots\}$ tales que sean disjuntos entre ellos, es decir que $\mathbb{B}_i \cap \mathbb{B}_j = \phi \quad \forall \ i \neq j$, se debe cumplir :

$$P\left[\bigcup_{k=1}^{\infty} \mathbb{B}_k\right] = \sum_{k=1}^{\infty} P(\mathbb{B}_k)$$

Entonces, si P es una medida de aditividad finita, decimos que P es completamente aditiva o de aditividad numerable.

Para pedir esta propiedad debemos suponer que el álgebra de Boole también cumple que :

$$\bigcup_{k=1}^{\infty} \mathbb{B}_k \in \mathbb{A} \qquad \text{si cada uno de los } \mathbb{B}_k \in \mathbb{A}.$$

Las álgebras de Boole que cumplen con esta última propiedad se denominan σ-álgebras de Boole. El ejemplo más típico es el caso en que $\mathbb{A}$ incluye a todos los subconjuntos de $\mathbb{E}$.

2.7. DEFINICION DE PROBABILIDAD PARA ESPACIOS MUESTRALES INFINITOS NUMERABLES

Sea $\mathbb{E}$ un conjunto infinito numerable y $\mathbb{A}$ una σ-álgebra de Boole de subconjuntos de $\mathbb{E}$, se dice que una función de conjunto P definida sobre $\mathbb{A}$ es una medida de probabilidad si:

1 - P es una medida con aditividad finita.

2 - P es una medida con aditividad numerable.

3 - $P(\mathbb{E}) = 1$.

Si $\mathbb{A}$ es la clase de *todos* los subconjuntos que pertenecen a $\mathbb{E}$, entonces la función P queda completamente determinada con sólo asignarle valores a todos y a cada uno de los subconjuntos de un elemento (probabilidades puntuales).

Si $\mathbb{B}$ es un subconjunto cualquiera de $\mathbb{E}$, tal que $\mathbb{B} \in \mathbb{A}$, entonces:

$$P(\mathbb{B}) = \sum_{x \in \mathbb{B}} P(\{x\})$$

2.8. DEFINICION DE PROBABILIDAD PARA ESPACIOS MUESTRALES INFINITOS NO NUMERABLES

Intentar definir una medida de probabilidad teniendo un espacio muestral no numerable con el mismo proceso que para espacios muestrales numerables ocasiona ciertas dificultades que exceden la esencia de este trabajo. Para salvarlas sin ahondar demasiado, vamos a realizar algunas restricciones:

- El espacio muestral $\mathbb{E}$, será un subconjunto no numerable del eje real $\mathbb{R}$.

- Para el álgebra Booleana utilizaremos subconjuntos de $\mathbb{E}$ que sean *medibles*, intentaremos mencionar sólo algunas propiedades esenciales que poseen los conjuntos medibles.

a) Si $\mathbb{B}$ es medible, entonces $\mathbb{R} - \mathbb{B}$ es medible.

b) La unión de una colección infinita numerable de conjuntos medibles, es medible. Es decir que si $\{\mathbb{B}_1, \mathbb{B}_2, \ldots\}$ es una colección numerable de conjuntos medibles $\mathbb{B}_i$, entonces $\mathbb{C} = \bigcup_{i=1}^{\infty} \mathbb{B}_i$ es medible.

c) Todo intervalo es medible, tanto si es abierto, cerrado, semiabierto, finito o infinito.

Así, las clases restringidas $\mathbb{A}$ serán σ-álgebras de Boole que contienen sólo subconjuntos medibles de $\mathbb{E}$, a su vez $\mathbb{E}$ está incluido en $\mathbb{R}$.

Existe una σ-álgebra de Boole mínima que contiene a todos los intervalos de $\mathbb{E}$, cuyos elementos son llamados conjuntos de Borel en $\mathbb{R}$.

Definimos entonces a las medidas de probabilidad sobre una σ-álgebra de Boole $\mathbb{A}$, formada por subconjuntos de Borel de $\mathbb{E}$ (con $\mathbb{E} \subseteq \mathbb{R}$) a una función de conjunto P que cumple con las siguientes propiedades:

- P es función de conjunto no negativa.

- P es función de conjunto completamente aditiva, definida en $\mathbb{A}$.

- $P(\mathbb{E}) = 1$.

Se puede demostrar que la mayor parte (podrían ser todas) de las probabilidades puntuales de un espacio de probabilidad con espacio muestral infinito no numerable serán cero. Entonces *no* se pueden calcular las probabilidades de sucesos cualesquiera sumando las correspondientes probabilidades puntuales. La solución vendrá dada por el cálculo de integrales que reemplazarán a las sumatorias.

2.9. VARIABLE ALEATORIA

A veces se hace necesario asociar un número a cada uno de los resultados de un experimento, es decir, a cada miembro del espacio muestral $\mathbb{E}$. Para esto definamos a la variable aleatoria[4] como una función que tiene como dominio a $\mathbb{E}$ y sus valores pertenecen a los reales $\mathbb{R}$.

Si notamos con X a esta variable aleatoria, resulta $X:\mathbb{E} \dashrightarrow \mathbb{R}$.

En general, dado un conjunto $\mathbb{E}$, pueden existir más de una variable aleatoria; conviene elegir una sencilla que simplifique la forma de escribir los sucesos que nos interesarán estudiar.

Ejemplo: Si tomamos como experimento tirar una moneda dos veces, y notamos con C si sale cara y con S si sale ceca, el espacio muestral será : {CC, CS, SC, SS}.

Si estamos interesados en cada uno de los cuatro resultados, podríamos definir a la variable aleatoria de la siguiente manera :

$$X(CC) = 0 \quad ; \quad X(CS) = 1 \quad ; \quad X(SC) = 2 \quad ; \quad X(SS) = 3$$

Para describir la probabilidad del suceso "que salga solamente una cara en las dos tiradas", podemos escribir $P(X=1 \text{ ó } X=2)$.

Una manera alternativa podría ser definir una variable aleatoria distinta Y:

$$Y(CC) = 0 \quad ; \quad Y(CS) = Y(SC) = 1 \quad ; \quad Y(SS) = 2$$ [5]

[4] A veces llamada variable aleatoria unidimensional.

[5] Notar que la función para definir esta variable aleatoria no es uno a uno.

Entonces para describir la probabilidad del mismo suceso anterior, "solamente una cara", escribiremos $P(Y=1)$.

Si nos interesa diferenciar cada uno de los resultados del experimento, debemos restringir la función que elegimos como variable aleatoria a ser uno a uno. Es decir que si ε_1 y ε_2 son dos resultados cualesquiera diferentes del espacio muestral $\mathbb{E}$, entonces X debe ser tal que $X(\varepsilon_1) \neq X(\varepsilon_2)$.

Si nos quedamos con este tipo de funciones, la variable aleatoria heredará necesariamente la numerabilidad o no numerabilidad del conjunto $\mathbb{E}$[6].

2.10. SUCESOS EQUIVALENTES

Dado un espacio muestral $\mathbb{E}$ y una variable aleatoria X con recorrido[7] $\mathbb{R}_x$, diremos que un subconjunto $\mathbb{B}$ de $\mathbb{E}$ y un subconjunto $\mathbb{C}$ de $\mathbb{R}_x$ son equivalentes si :

$$\mathbb{B} = \{ s \in \mathbb{E} \text{ tal que } X(s) \in \mathbb{C} \}$$

Es decir que $\mathbb{B}$ contiene a todos los resultados en $\mathbb{E}$ que se corresponden con $\mathbb{C}$ en $\mathbb{R}_x$ mediante X.

[6]Vale la pena recordar que, tanto los números Enteros como los Racionales son infinitos numerables, mientras que los Irracionales y en concecuencia los Reales son infinitos **no** numerables.

[7]Al ser X una función, su recorrido es el conjunto de valores que toma la función.

Definimos la probabilidad del suceso $\mathbb{C}$ como:

$$P(\mathbb{C}) = P(\mathbb{B}) \tag{2.1}$$

donde $\mathbb{B}$ es equivalente con $\mathbb{C}$.

Esto significa que las probabilidades en el espacio muestral $\mathbb{E}$ se definen según el modelo que se utilice para representar a la realidad, mientras que las probabilidades en $\mathbb{R}_x$ quedan determinadas a partir de la definición dada a través de la ecuación (2.1).

2.11. VARIABLES ALEATORIAS DISCRETAS

Decimos que una variable aleatoria X es discreta si la cantidad de elementos que contiene el conjunto Imagen de la función que define a X[8] es finito o infinito numerable. Entonces podremos enumerar a cada uno de los elementos como un miembro de la lista $\{x_1, x_2, ...\}$.

Si a cada x_i le asignamos una probabilidad, según queda definido por la probabilidad de los correspondientes conjuntos equivalentes en $\mathbb{E}$, se debe cumplir:

$$p(x_i) \geq 0 \qquad \text{y} \qquad \sum_{i=1}^{\infty} p(x_i) = 1$$

La función p se denomina *función de probabilidad puntual*

[8]La cantidad de valores posibles de X.

de la variable aleatoria X. La colección de pares $(x_i, p(x_i))$ se denomina frecuentemente distribución de probabilidad de la variable aleatoria X.

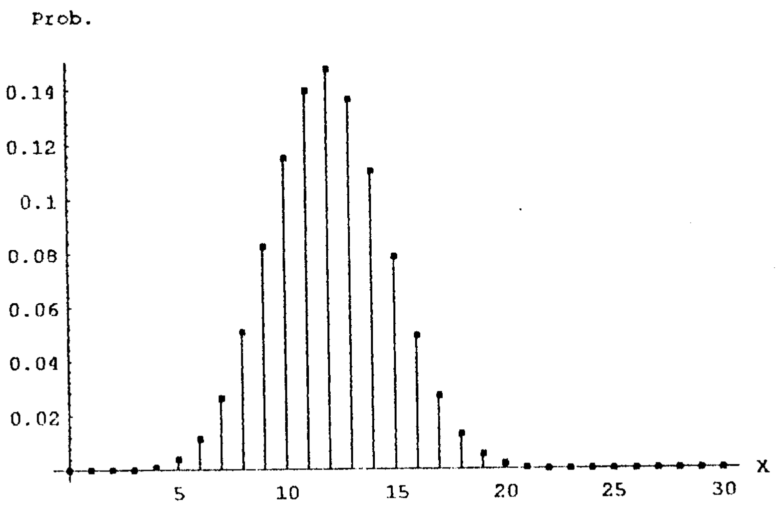

Figura 1. *Gráfico de probabilidad puntual.* Las variables aleatorias discretas pueden tomar una cantidad numerable de valores.

Si tenemos un suceso asociado con X, esto es $\mathbb{C} \subseteq \mathbb{R}_x$ y tenemos que $\mathbb{C} = \{x_{i1}, x_{i2}, x_{i3}, \ldots\}$.

Entonces $$P(\mathbb{C}) = \sum_{j=1}^{\infty} p(x_{ij})$$

Es interesante notar que cuando el espacio muestral es infinito numerable no puede existir equiprobabilidad en todos sus elementos puntuales, si lo suponemos tendríamos que $p(x_i) = k \quad \forall i$, pero entonces no podríamos obtener la propiedad $\sum_{i=1}^{\infty} p(x_i) = 1$ pues debería cumplirse $\sum_{i=1}^{\infty} k = 1$.

2.12. VARIABLES ALEATORIAS CONTINUAS

Decimos que una variable aleatoria X es contínua si la cantidad de elementos que contiene el conjunto Imagen de la función que define a X es infinito no numerable.

En este caso deja de tener sentido $p(x_i)$, ya que al no ser numerables los valores que toma X, no podemos contabilizarlos con $\{x_1, x_2, \ldots\}$ y entonces sustituiremos esta probabilidad puntual por una función $f(x)$ que tendrá como dominio a $\mathbb{R}_x$, el conjunto no numerable de valores que puede tomar la variable aleatoria X.

Así $f(x)$ cumplirá las propiedades :

- $f(x) \geq 0 \qquad \forall\, x \in \mathbb{R}_x$

- $\displaystyle\int_{\mathbb{R}_x} f(x)\, dx = 1$

Podemos entonces extender al dominio de los reales $\mathbb{R}$ a la función $f(x)$ definiendo $f^*(x)$ tal que $f^*(x)=0 \quad \forall\, x \notin \mathbb{R}_x$.

Así tenemos entonces que

$$P(\mathbb{B}) = \int_{\mathbb{B}} f^*(x)\, dx$$

para todo conjunto $\mathbb{B}$ que pertenece a la σ-álgebra de Boole $\mathbb{A}$ formada por los subconjuntos de Borel que pertenecen a $\mathbb{R}$.

Esta función $f(x)$ se llama *función de densidad de probabilidad* de la variable aleatoria X.

Vale la pena acotar que :

- Si f(x) es acotada, $P(X = x_o) = \lim_{\varepsilon \to 0} \int_{x_o}^{x_o+\varepsilon} f(x)\, dx = 0$, sin embargo esto no significa que sea **imposible** (en el sentido estricto) que X tome el valor x_o.

- $f(x)$ no es ninguna probabilidad, lo que sí es una probabilidad es $\int_{\mathbb{B}} f(x)\, dx$

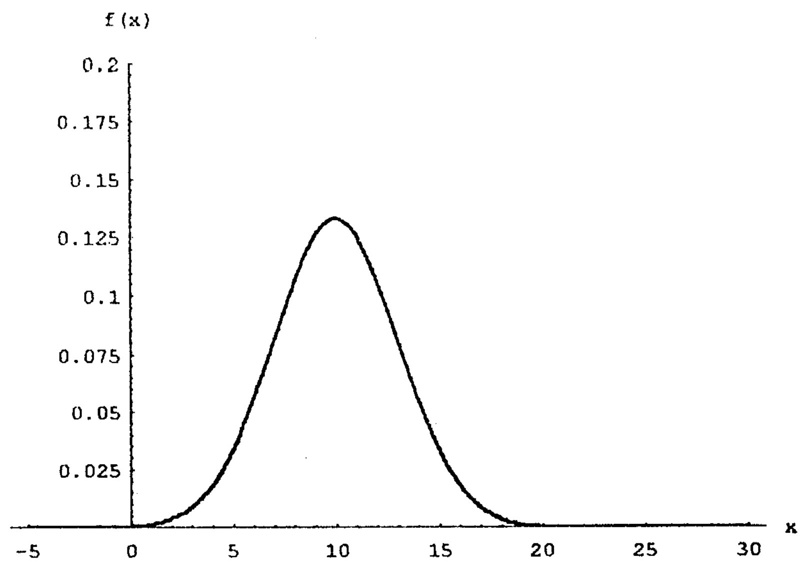

Figura 2. *Gráfico de una función de densidad de probabilidad.* Las variables aleatorias contínuas tienen asociada una función de probabilidad $f(x)$.

2.13. VARIABLES ALEATORIAS MIXTAS

Existen ciertos experimentos, cuyos resultados se describen mejor a través de una variable aleatoria X con valores pertenecientes a $\mathbb{R}_x \leq \mathbb{R}$, y tales que tienen una medida de probabilidad P que es discreta y contínua a la vez. Esto es, X por un lado toma valores $x_1, x_2, \ldots, x_n$ donde cada uno de éstos tiene asociado una probabilidad puntual $p(x_i) \geq 0$, y tal que $\sum_{i=1}^{n} p(x_i) = P$ [0]. A su vez X también toma todos los valores $x \in I_j$, siendo $I_j=(a_j, b_j)$ algún intervalo abierto de los reales, con $j=1,2,\ldots,m$ [1]. Cada uno de los intervalos I_j es disjunto con respecto a todos los otros, es decir $I_j \cap I_k = \phi$ para $j \neq k$. Entonces definimos una función $f(x)$ tal que

$$P_j = \int_{I_j} f(x)\, dx = \int_{a_j}^{b_j} f(x)\, dx \quad , \quad \text{con } P_j \geq 0$$

y tal que

$$\sum_{j=1}^{m} \left[\int_{a_j}^{b_j} f(x)\, dx \right] = \sum_{j=1}^{m} P_j = \tilde{P} = 1 - P$$

Así tenemos que, si $\mathbb{E}$ es el espacio muestral completo, la probabilidad de que ocurra un resultado cualquiera es :

$$P(\mathbb{E}) = P(-\infty < X < \infty) = \left(\sum_{i=1}^{n} p(x_i) \right) + \sum_{j=1}^{m} P_j = P + \tilde{P} = P + (1-P) = 1$$

[0] Podrían ser infinitos valores de X, pero el conjunto debe ser numerable. En ese caso se reemplaza en la sumatoria n por ∞.

[1] Este conjunto particular de intervalos $\{ I_j \}$ será un subconjunto del conjunto de Borel de $\mathbb{R}$.

Si ahora extendemos la función $f(x)$ a los reales, definiendo $f^*(x) = f(x)$ para los x que pertenezcan a algún intervalo I_j, y $f^*(x) = 0$ para los x que no pertenezcan a ningún intervalo I_j, tenemos que :

$$P(a \leq X \leq b) = \left[\sum_{\{i/a \leq x_i \leq b\}} p(x_i) \right] + \int_a^b f^*(x)\, dx$$

A este tipo de variable aleatoria se le llama *variable aleatoria de tipo mixta.*

2.14. FUNCION DE DISTRIBUCION ACUMULATIVA

Si tenemos una variable aleatoria X, definimos la *función de distribución acumulativa* $F(x)$ de la siguiente manera :

$$F(x) = P(X \leq x)$$

i) Si X es una variable aleatoria discreta

$$F(x) = \sum_{\{i/\ x_i \leq x\}} p(x_i)$$

ii) Si X es una variable aleatoria contínua

$$F(x) = \int_{-\infty}^{x} f^*(x')\, dx'$$

iii) Si X es variable aleatoria mixta

$$F(x) = \int_{-\infty}^{x} f^{*}(x')\, dx' + \sum_{\{i/\ x_i \le x\}} p(x_i)$$

La función de distribución acumulativa, a veces notada con fda, tiene varias propiedades interesantes, algunas de las cuales enunciaremos:

a) $F(x)$ es una función que tiene como dominio a *todo* el eje real $\mathbb{R}$.

b) $F(x)$ es no decreciente, es decir que si $x<x'$, entonces será $F(x) \le F(x')$ $\forall\ x, x' \in \mathbb{R}$.

c) $\lim\limits_{x \to -\infty} F(x) = F(-\infty) = 0$

d) $\lim\limits_{x \to +\infty} F(x) = F(+\infty) = 1$

e) $P(a<X\le b) = P(X\le b) - P(X\le a) = F(b) - F(a)$

f) Si X es una variable aleatoria contínua con función de probabilidad $f(x)$, tenemos que $f(x) = \frac{d}{dx}F(x)$ para todos los x en que $F(x)$ sea diferenciable.

g) Si X es variable aleatoria discreta y rotulamos los $x_1, x_2, x_3, \ldots$ tales que $x_1<x_2<x_3<\ldots$, tenemos que

$$p(x_i) = P(X=x_i) = F(x_i) - F(x_{i-1})$$

h) Si X es una variable aleatoria mixta con una función de distribución $f(x)$, tendremos que la fda $F(x)$ será discontínua con saltos solamente en los valores x_i tales que $p(x_i) > 0$, así tendremos que

$$\lim_{x \to x_i^+} F(x) = F(x_i)$$

y

$$\lim_{x \to x_i^-} F(x) = F(x_i) - P(X=x_i)$$

Es decir que en cada punto x_i tal que $p(x_i) > 0$, $F(x)$ es contínua solamente a derecha. El salto de $F(x)$ en la discontinuidad es igual a la probabilidad en el punto x_i correspondiente.

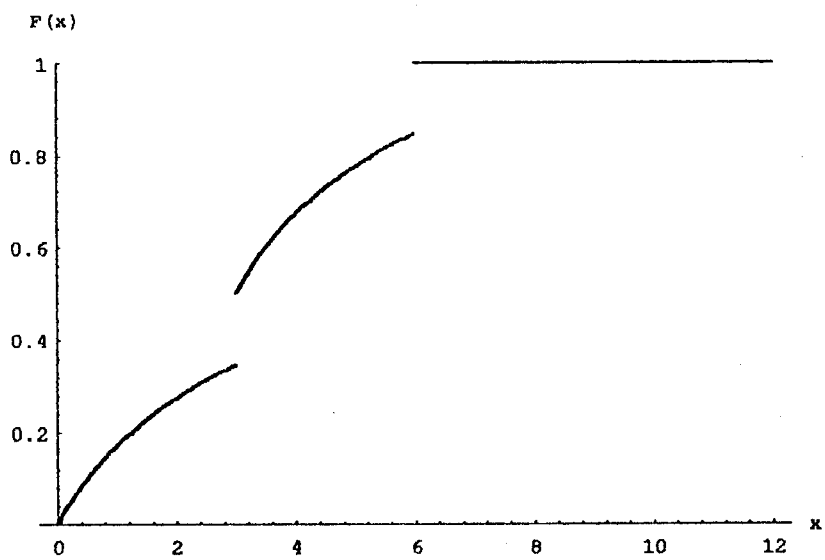

Figura 3. *Gráfico de una función de distribución acumulativa.* Las funciones de distribución acumulativas de variables aleatorias mixtas tienen saltos en los puntos que poseen probabilidad no nula.

3.1. PROBABILIDAD CONDICIONADA

Veamos cómo podemos utilizar en algún tipo de experimento cierta información adicional. Para ello utilizaremos un ejemplo.

Se lanzan dos dados y se sabe que la suma de los mismos es seis. Cuál es la probabilidad de que los dos números sean impares? Formularemos este problema de otra manera. Sea A el suceso: "la suma de los números es seis" y B el suceso: "los dos números son impares". Cuál es la probabilidad de que ocurra el suceso B sabiendo que A ha ocurrido?

Hay que tener cuidado de no confundir esta probabilidad con la probabilidad del suceso $A \cap B$, que no son necesariamente lo mismo.

Calculemos estas dos probabilidades para ver sus diferencias. Comenzaremos calculando $P(B/A)$[11]. El espacio muestral correspondiente al lanzamiento de dos dados es: $S = \{(1,1);(1,2);(1,3);\ldots;(6,6)\}$ que tiene 36 elementos[12]. Ahora queremos calcular cuál es la probabilidad de que los dos números sean impares sabiendo que este suceso ya está en A, es decir que los dos números suman seis; por lo tanto hay que restringirse al subconjunto $S_A = \{(1,5);(5,1);(3,3);(2,4);(4,2)\}$. De esta manera debemos tomar de este subconjunto los elementos tal que los dos números sean

[11] Léase "la probabilidad de B dado A".

[12] Considerando al par ordenado (a,b) diferente de (b,a) si $b \neq a$.

impares, esto es $S_{B/A} = \{(1,5);(5,1);(3,3)\}$. Entonces, si asignamos a cada uno de los resultados de A la misma probabilidad 1/5, $P(B/A) = 3.(1/5) = 3/5$.

Veamos qué sucede con $P(A \cap B)$. Esto se puede enunciar de la siguiente manera: se lanzan dos dados. Cuál es la probabilidad de que los dos números sean impares y sumen seis? En este caso el espacio muestral es S y cada uno de los elementos tienen la misma probabilidad 1/36. Como el suceso $A \cap B$ es el subconjunto $S_{A \cap B} = \{(1,5);(5,1);(3,3)\}$, tenemos que $P(A \cap B) = 3.(1/36) = 1/12$. La diferencia fundamental entre $P(B/A)$ y $P(A \cap B)$ es el espacio muestral con el que estamos trabajando. Sin embargo estas dos probabilidades requieren casi lo mismo, es decir que sucedan A y B. Esto nos da la pauta que deben estar relacionadas de alguna manera.

Lo que debemos hacer es definir una nueva función de probabilidad P' que esté restringida al nuevo espacio muestral y un álgebra booleana $\mathbb{B}'$ de subconjuntos de A de modo que (A, $\mathbb{B}'$, P') sea un espacio de probabilidad. En nuestro caso el nuevo espacio de probabilidad es $S_A = A$ (conjunto de elementos relacionado con el suceso A), estando P' definido de tal manera que la probabilidad asignada a dicho espacio sea 1, esto es $P'(A) = 1$. Una forma de definir la función de probabilidad P' es dividiendo cada una de las antiguas probabilidades por P(A). Esto es:

si $C \subseteq S_A$ y $P(A) \neq 0$

$$P'(C) = \frac{P(C)}{P(A)}$$

Se ve fácilmente que $P'(C)$ es no negativa y que $P'(A) = 1$. C es un suceso que está incluído en S_A y que también puede estar incluído en otro suceso; en nuestro caso $C = A \cap B$. De esta manera podemos definir P' de la siguiente forma:

$$P'(A \cap B) = \frac{P(A \cap B)}{P(A)}$$

Esto nos sugiere una definición adecuada a nuestro problema de interés $P(B/A)$.

Definición de probabilidad condicionada. Sea $(S, \mathbb{B}, P)$ un espacio de probabilidad y A un elemento tal que $P(A) \neq 0$. La probabilidad condicionada de que un suceso B ocurra, en el supuesto de que A ha ocurrido, se representa mediante el símbolo $P(B/A)$ y se define por la igualdad

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

La probabilidad $P(B/A)$ no está definida si $P(A) = 0$.

Sigamos con nuestro ejemplo tomando en cuenta que $P(A \cap B) = 1/12$ y que $P(A) = 5/36$, entonces

$$P(B/A) = \frac{1/12}{5/36} = 3/5$$

Este resultado concuerda con nuestro cálculo anterior.

Habiendo definido la probabilidad condicional se pueden hacer dos afirmaciones generales acerca de $P(B/A)$. El primer caso es cuando $A \cap B = \phi$, es decir que los sucesos A y B no pueden ocurrir simultáneamente. Entonces $P(B/A) = 0$ pues si ocurre uno no puede ocurrir el otro.

El segundo caso a considerar es cuando $A \subset B$, entonces

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

Si ocurre A (un conjunto incluído en B) con seguridad ocurre B.

3.2. INDEPENDENCIA

Comenzaremos con un ejemplo para ilustrar este concepto. Se lanza dos veces una moneda, sean los sucesos A y B como sigue:

$A = \{$ la primera moneda es cara $\}$

$B = \{$ la segunda moneda es cara $\}$

Intuímos que saber que A ha ocurrido no nos proporciona ninguna información adicional acerca de la ocurrencia de B. Hagamos un simple cálculo para comparar $P(B)$ y $P(B/A)$. Nuestro espacio muestral es $S = \{(c,c);(c,s);(s,c);(s,s)\}$, donde c indica "cara" y s indica "ceca". La probabilidad de que ocurra B, considerando todos los resultados igualmente probables, es $P(B) = 2/4 = 1/2$; para calcular $P(B/A)$ necesitamos saber $P(A) = 2/4 = 1/2$ y $P(A \cap B) = 1/4$ entonces

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/4}{1/2} = 1/2 = P(B)$$

Vemos entonces que $P(B/A) = P(B)$, es decir que la ocurrencia de A no ha afectado la ocurrencia de B. Del cálculo anterior vemos que se tiene que cumplir

$$\frac{P(A \cap B)}{P(A)} = P(B)$$

Entonces debe ser $P(A \cap B) = P(A).P(B)$.

Hemos llegado así al concepto de independencia, que se define formalmente como sigue:

Definición de independencia. Dos sucesos A y B se llaman independientes (o estocásticamente independientes) si, y sólo si,

$$P(A \cap B) = P(A).P(B)$$

Muchas veces es difícil determinar de antemano si dos sucesos son independientes, sin embargo la definición de independencia nos proporciona un método útil para llegar a esta determinación.

3.3. EXPERIMENTOS O PRUEBAS COMPUESTAS

Se considera como prueba compuesta al resultado de ejecutar sucesivamente dos o más pruebas que pueden ser distintas o no. Estas pruebas pueden estar relacionadas entre sí o pueden ser estocásticamente independientes.

Comenzaremos nuestra discusión citando un interesante ejemplo histórico en el cual se realizan pruebas iguales sucesivas, estocásticamente independientes. Antoine Gombaud, caballero de Méré, en 1654 llevó a Blaise Pascal y a Pierre de Fermat a interesarse, a través de cuestiones de juego y apuestas, en el cálculo de probabilidades.

Uno de los problemas que planteó de Méré a Pascal, aparentemente contradictorio según el primero, fue el siguiente: se lanzan 24 veces un par de dados. Es conveniente apostar a favor o en contra de la aparición de por lo menos un doble seis?

Haciendo un análisis muy superficial se puede razonar de la siguiente manera: existen 36 resultados diferentes en la tirada de un par de dados, si uno hace 24 tiradas debe ser bastante probable que salga un doble seis.

Hay que tener mucho cuidado con este tipo de razonamientos apresurados. Como hemos dicho al principio de esta sección, este problema es una prueba compuesta cada una de las cuales es estocásticamente independiente una de las otras. Entonces haciendo uso de un razonamiento un poco más profundo pensemos de la siguiente manera: en la primera tirada tengo una probabilidad de 1/36 de que salga un doble seis. Cuando tiro por segunda vez los dados el resultado de la primera prueba no tiene influencia, entonces, la probabilidad de obtener un doble seis es también 1/36. Puesto que las pruebas son independientes la probabilidad debería ser el producto de las probabilidades; pero si multiplicamos 1/36 tantas veces como pruebas hagamos, la probabilidad de obtener por lo menos un doble seis disminuiría, en contraposición con lo que experimentalmente se observa.

El error que se ha cometido en estos razonamientos es no definir adecuadamente el espacio muestral y la correspondiente medida de probabilidad.

Definamos el mencionado espacio como sigue: sea x_i el resultado de la i-ésima tirada, es decir, un par ordenado (a,b). El resultado del juego se puede representar con la 24-pla ordenada $(x_1, x_2, \ldots, x_{24})$. Por ejemplo, un resultado puede ser que en las primeras 23 tiradas haya salido el par ordenado (1,1) y en la última el (1,2), y otro resultado diferente es que en la primera tirada haya salido el par ordenado (1,2) y en las restantes el (1,1). Entonces el espacio muestral es el producto cartesiano de los espacios muestrales de cada una de las pruebas, esto es $S^{24} = S \times S \times \ldots \times S$, donde S^{24} es el espacio que queremos definir. De esta manera el espacio S^{24} tiene 36^{24} elementos diferentes, y, asignando a cada una de las 24-plas igual probabilidad, se tiene $P(x) = 1/36^{24}$, donde x representa un elemento del nuevo espacio muestral.

Volvamos a nuestro ejemplo pero calculemos el suceso complementario por simplicidad. Sea este suceso $\bar{A}$: "que no salga ningún doble seis en 24 tiradas de un par de dados". Calculemos ahora la cantidad de elementos que tiene este suceso. Los resultados posibles de $\bar{A}$ son las 24-plas que no contengan el par (6,6), por lo tanto podemos tener 35 valores posibles para cada x_i (con $1 \leq i \leq 24$), entonces existen 35^{24} posibilidades. Como cada uno de estos resultados tiene la misma probabilidad $(1/36^{24})$, la probabilidad del suceso $\bar{A}$ es

$$P(\bar{A}) = \left(\frac{35}{36} \right)^{24}$$

De esta manera podemos encontrar el resultado tan esperado por de Méré

$$P(A) = 1 - \left(\frac{35}{36} \right)^{24} = 0{,}49.$$

Es decir, es conveniente apostar en contra de la aparición de un doble seis en 24 tiradas, lo que contradice al primer razonamiento hecho.

Esta discusión nos lleva a encontrar un método más general para tratar los experimentos sucesivos. Por simplicidad consideremos dos pruebas independientes combinadas en una prueba compuesta.

Comenzaremos estableciendo en una prueba compuesta el nuevo espacio de probabilidades $(S, \mathbb{B}, P)$ donde S es el nuevo espacio muestral, $\mathbb{B}$ el álgebra booleana de subconjuntos de S y P la medida de probabilidad.

Si tenemos dos experimentos E_1 y E_2, asociamos a cada uno de ellos un espacio de probabilidad $(S_1, \mathbb{B}_1, P_1)$ y $(S_2, \mathbb{B}_2, P_2)$ respectivamente. Sea E el experimento compuesto que resulta de hacer E_1 y E_2 sucesivamente. El espacio muestral de E será, como hicimos en el ejemplo anterior, el producto cartesiano de S_1 y S_2, esto es, $S = S_1 \times S_2$. Entonces, si el espacio de probabilidad S_1 tiene n elementos y el espacio S_2, m elementos, S tendrá nm elementos. Un resultado de S se podrá expresar a través del par ordenado (x,y) donde x es un resultado de hacer E_1 e y de hacer E_2.

Como nueva álgebra booleana $\mathbb{B}$ tomamos la colección de todos los subconjuntos de S.

Ahora debemos definir la nueva medida de probabilidad P. Para llegar a una definición razonable consideramos dos sucesos A y B del nuevo espacio muestral S definidos como sigue:

$$A = \left\{(x_1,y_1);(x_1,y_2)\ldots(x_1,y_m)\right\}$$

$$B = \left\{(x_1,y_1);(x_2,y_1)\ldots(x_n,y_1)\right\}$$

El conjunto unitario $\left\{(x_1,y_1)\right\}$ es la intersección entre A y B. Considerando a estos dos sucesos independientes entre sí, P debe ser tal que

$$P(A\cap B) = P(A).P(B) \qquad (3.1)$$

Nótese que $P(A\cap B)$ es en definitiva $P(x_1,y_1)$. Debemos encontrar ahora la forma de asignar las probabilidades P(A) y P(B). El suceso A ocurre solamente si el resultado de la primera prueba es x_1 sin importar cuál es el resultado de la segunda. Como el suceso A está determinado a partir del valor que adquiere x (si $x = x_1$ o no), es razonable considerar que $P(A) = P_1(x_1)$. De manera similar podemos pensar que $P(B) = P_2(y_1)$. Entonces la ecuación (3.1) nos queda

$$P(x_1,y_1) = P_1(x_1).P_2(y_1)$$

Este razonamiento se puede generalizar a cada elemento (x,y) de S. Entonces, si dos pruebas E_1 y E_2 son estocásticamente independientes, definimos P mediante la ecuación

$$P(x,y) = P_1(x).P_2(y) \qquad (3.2)$$

Comprobemos que P es efectivamente una medida de probabilidad. P es un producto de dos factores no negativos, entonces P debe ser no negativo. La suma de todas las probabilidades puntuales es

$$\sum_{(x,y)\in S} P(x,y) = \left[\sum_{x\in S_1} P_1(x) \right]\left[\sum_{y\in S_2} P_2(y) \right] = 1.1 = 1$$

La asignación de probabilidades (3.2) implica que para todo par de subconjuntos U de $\mathbb{B}_1$ y V de $\mathbb{B}_2$ vale

$$P(U\times V) = P_1(U).P_2(V) \qquad (3.3)$$

Demostremos esto:

$$P(U\times V) = \sum_{(x,y)\in U\times V} P(x,y) = \sum_{\substack{x\in U\\ y\in V}} P(x,y) = \sum_{x\in U}\sum_{y\in V} P_1(x).P_2(y) =$$

$$= \left[\sum_{x\in U} P_1(x) \right].\left[\sum_{y\in V} P_2(y) \right] = P_1(U).P_2(V)$$

Podemos deducir entonces algunas consecuencias importantes.

Consideremos un suceso A tal que cada elemento es un par ordenado (x,y), donde x es un resultado de la primera prueba E_1 que pertenece a un conjunto C_1, e y puede ser cualquier resultado de S_2. Es decir, A es de la forma $A = C_1 \times S_2$ donde $C_1 \in \mathbb{B}_1$.

Utilizando la ecuación (3.3) llegamos a

$$P(A) = P(C_1 \times S_2) = P_1(C_1).P_2(S_2) = P_1(C_1)$$

Un suceso como A (tal que $P(A) = P_1(C_1)$) se dice que está *determinado mediante la primera prueba* E_1. Análogamente definamos un suceso B que esté *determinado por la segunda prueba* E_2. Esto es, si $C_2 \in \mathbb{B}_2$

$$B = S_1 \times C_2$$

Tenemos entonces

$$P(B) = P(S_1 \times C_2) = P_1(S_1).P_2(C_2) = P_2(C_2)$$

Considerando todo esto, analicemos lo que sucede con $P(A \cap B)$. Primeramente encontremos qué conjunto es $A \cap B$. Por definición de intersección son los pares ordenados tal que (x,y) pertenecen a $C_1 \times S_2$ y a $S_1 \times C_2$. La primera de estas condiciones ($(x,y) \in C_1 \times S_2$) dice que se debe considerar los x que pertenecen a C_1 sin importar el resultado de la segunda prueba, pues y siempre pertenece a S_2; la otra condición impone que $y \in C_2$ sin considerar el resultado de la primera prueba. En resumen (x,y) debe ser tal que $x \in C_1$ e $y \in C_2$. Esto es $A \cap B = C_1 \times C_2$.

Utilizando la ecuación (3.3) obtenemos

$$P(A \cap B) = P(C_1 \times C_2) = P_1(C_1).P_2(C_2)$$

Como $P_1(C_1) = P(A)$ y $P_2(C_2) = P(B)$ resulta

$$P(A \cap B) = P(A).P(B)$$

Demostramos entonces que los sucesos A y B son independientes.

Podemos generalizar estos resultados haciendo una deducción análoga a experimentos compuestos de n pruebas independientes $E_1, E_2, ..., E_n$. Cada resultado del nuevo espacio muestral $S = S_1 \times S_2 \times S_3 \times ... \times S_n$ será una n-pla de la forma $(x_1, x_2, ..., x_n)$ y la medida de probabilidad será

$$P(x_1, x_2, ..., x_n) = P_1(x_1).P_2(x_2)...P_n(x_n) \qquad (3.4)$$

Cuando utilizamos esta definición de P se dice que E está *determinado por n pruebas independientes* $E_1, E_2, ..., E_n$.

Existe un caso especial de pruebas compuestas en el que todas las pruebas están asociadas al mismo espacio de probabilidad. Se dice que la prueba compuesta E es un caso de *pruebas independientes repetidas bajo idénticas condiciones*. Un ejemplo de este tipo de pruebas es el de las llamadas pruebas de Bernoulli.

3.4. DISTRIBUCION BINOMIAL

A las pruebas compuestas que tienen una distribución binomial se las conoce también por el nombre de *suceción de pruebas de Bernoulli*, tal nombre fue puesto en honor a Jacobo Bernoulli. Se trata de una suceción de pruebas repetidas ejecutadas en las mismas condiciones, siendo cada resultado estocásticamente independiente de los demás. En las pruebas de Bernoulli solamente existen dos resultados posibles en cada una de las pruebas, llamadas comunmente "éxito" y "fallo". Se asocia al éxito una probabilidad p y al fallo una probabilidad q, donde $q = 1 - p$.

Consideremos un caso sencillo para ejemplificar estas ideas.

Se lanza una moneda tres veces. Se asume que esto cumple con la condición de que cada una de las pruebas ("se lanza una moneda") es estocásticamente independiente una de las otras. Además podemos tener solamente dos resultados posibles. Asociemos el éxito al resultado "cara" y el fallo a "ceca". Supongamos que la moneda no está bien balanceada y que la probabilidad de que salga cara es $p = 2/3$, entonces la probabilidad de fallo es $q = 1/3$. Designemos al éxito con E y al fallo con F, el espacio muestral es

$$S = \{(E,E,E);(F,E,E);(E,F,E);(E,E,F);(F,F,E);(F,E,F);(E,F,F);(F,F,F)\}$$

Utilizando la ecuación (3.4) obtenemos las respectivas probabilidades

$$(2/3)^3;\ (1/3).(2/3)^2;\ (1/3).(2/3)^2;\ (1/3).(2/3)^2;$$

$$(1/3)^2.(2/3);\ (1/3)^2.(2/3);\ (1/3)^2.(2/3);\ (1/3)^3.$$

Pero comúnmente no nos interesa saber, por ejemplo, si en la segunda tirada salió cara o ceca, sino que queremos saber cuántas caras salieron o cuántas cecas. En este caso necesitamos definir una variable aleatoria X que indique el número de veces que se obtuvo el éxito en las tres tiradas (también se podría haber definido el fallo). Esto es, X puede tomar los valores 0, 1, 2 y 3. De esta manera $x = 0$ si ocurre (F,F,F), $x = 1$ si ocurre (E,F,F), (F,E,F) o (F,F,E), etcétera.

Entonces las probabilidades serán como sigue:

$$P(x = 0) = P(0) = (1/3)^3 \qquad P(1) = 3.(1/3)^2.(2/3)$$

$$P(2) = 3.(1/3).(2/3)^2 \qquad P(3) = (2/3)^3$$

Notemos que la suma de las pobabilidades es

$$(1/3)^3 + 3.(1/3)^2.(2/3) + 3.(1/3).(2/3)^2 + (2/3)^3 =$$
$$= (1/3 + 2/3)^3 = 1$$

Ahora enunciemos y demostremos el teorema principal relacionado con las sucesiones de Bernoulli.

Fórmula de Bernoulli. La probabilidad de k éxitos en n pruebas de Bernoulli es

$$\binom{n}{k} p^k q^{n-k}$$

donde $\binom{n}{k}$ representa el coeficiente binomial $\dfrac{n!}{k!\,(n-k)!}$,
p es la probabilidad de éxito en una de las pruebas y $q = 1 - p$.

Demostración. Consideremos un resultado particular del espacio muestral, esto es la n-pla

$$(x_1, x_2, \ldots, x_n)$$

Si representamos el éxito con E y el fallo con F cada x_i $(1 \leq i \leq n)$ es F o E. En particular nosotros estamos interesados en los resultados que contienen k veces E y n − k veces F; designemos con A este suceso. Calculemos la probabilidad de uno de los resultados de A, por ejemplo

$$\bar{x}_0 = (E, E, \ldots, E, F, F, \ldots, F)$$

Utilizando la ecuación (3.4), la probabilidad de que ocurra $\bar{x}_0$ es

$$P(\bar{x}_0) = P_1(E).P_2(E)\ldots P_k(E).P_{k+1}(F)\ldots P_n(F) = p^k q^{n-k}$$

Como cada uno de los elementos de A tiene la misma probabilidad que $\bar{x}_0$, solamente necesitamos contar el número de elementos de A y multiplicar dicho número por $p^k q^{n-k}$.

Para contar los elementos de A hay que ver de cuántas maneras es posible colocar k veces la E en las n posiciones de la n-pla. Este número es $\binom{n}{k}$. Entonces sumando las probabilidades de todos los elementos de A obtenemos

$$P(A) = \binom{n}{k} p^k q^{n-k}$$

Observación: llamemos $P_k(n,p)$ a la probabilidad de obtener k éxitos en un sucesión de pruebas de Bernoulli con n ensayos y con probabilidad de éxito p en cada una de las pruebas. Veamos que $P_k(n,p)$ es una medida de probabilidad. Este número es un producto de factores no negativos, por lo tanto es no negativo. Sumemos las probabilidades sobre todos los posibles valores de k

$$\sum_{k=0}^{n} P_k(n,p) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1^n = 1$$

Con esto probamos que P es una medida de probabilidad.

3.5. APROXIMACION PARA EL CALCULO DE LA DISTRIBUCION BINOMIAL

Muchas veces es dificil calcular $P_k(n,p)$ cuando n es un número grande debido a que esta probabilidad involucra factoriales de n. Haciendo uso de la *fórmula de Stirling*, la cual aproxima factoriales de enteros que son mucho mayores que 1, puede demostrarse la siguiente aproximación:

$$P_k(n,p) \simeq \frac{1}{\sqrt{2\pi\ n\ p\ q}} \exp\left[-\frac{1}{2} \left(\frac{k - n\ p}{\sqrt{n\ p\ q}} \right)^2 \right]$$

Si se define una variable aleatoria X como la cantidad de éxitos obtenidos en las n pruebas se tiene que

$$P(X \leq K) = \sum_{i=0}^{K} P_i(n,p) \quad = P\left[\frac{X - n\ p}{\sqrt{n\ p\ q}} \leq \frac{K - n\ p}{\sqrt{n\ p\ q}} \right] \simeq$$

$$\simeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(K-np)/(npq)^{1/2}} e^{-t^2/2}\ dt$$

Este último término representa a la función de distribución acumulativa de la conocida densidad de probabilidad normal.

4. CONCLUSIONES

A lo largo de este trabajo vemos que cuando describimos un experimento a través de un modelo no determinista es de suma importancia definir en forma adecuada el espacio de probabilidad ($\mathbb{E}$, $\mathbb{A}$, P) con el que estamos trabajando. Este espacio debe ser tal que $\mathbb{E}$ refleje todos los resultados posibles de nuestra experiencia y P sea una medida de probabilidad bien definida.

También mencionamos que la diferencia entre el cálculo de probabilidades utilizando variables aleatorias discretas o contínuas se fundamenta en la cantidad de elementos del espacio muestral, ya sea éste numerable en sentido amplio o no numerable.

Al ser frecuente encontrarnos con experimentos compuestos, es de mucha importancia y utilidad el estudio de un caso particular que es la sucesión de pruebas de Bernoulli. Dicha sucesión está formada por n pruebas estocásticamente independientes cada una de las cuales está asociada al mismo espacio de probabilidad, teniendo el espacio muestral de cada prueba solamente dos resultados (comunmente llamadas éxito y fallo respectivamente).

5. APENDICE

En este apéndice mostraremos que, dado un Universal $\mathbb{S}$, la mínima colección de subconjuntos $\mathbb{A}_o$ posible para lograr un álgebra de Boole está formada por dos conjuntos : $\{\phi, \mathbb{S}\}$.

Recordemos las dos propiedades básicas del álgebra de Boole, si $\mathbb{B} \in \mathbb{A}_o$ y $\mathbb{C} \in \mathbb{A}_o$ entonces tenemos que :

1 - $\mathbb{B} \cup \mathbb{C} \in \mathbb{A}_o$,

2 - $\mathbb{B}' = \mathbb{S} - \mathbb{B} \in \mathbb{A}_o$.

i) Demostremos que si $\mathbb{A}_o$ es un álgebra de Boole, entonces el conjunto vacío pertenece a $\mathbb{A}_o$.

De la teoría de conjuntos tenemos que:

3 - $\mathbb{B} \cap \mathbb{C} = (\mathbb{B}' \cup \mathbb{C}')'$,

4 - $\mathbb{B} - \mathbb{C} = \mathbb{B} \cap \mathbb{C}'$.

Como la clase $\mathbb{A}_o$ es no vacía, elijiendo un conjunto cualquiera $\mathbb{B}$ que pertenezca a $\mathbb{A}_o$, se tiene que :

5 - $\mathbb{B}' \in \mathbb{A}_o$,

6 - $\phi = \mathbb{B} - \mathbb{B} = \mathbb{B} \cap \mathbb{B}' = (\mathbb{B}' \cup \mathbb{B}'')' = (\mathbb{B}' \cup \mathbb{B})'$.

Entonces por la propiedad 1, se tiene que $(\mathbb{B}' \cup \mathbb{B})' = \phi \in \mathbb{A}_o$.

ii) Por la propiedad 2 se tiene que si $\phi \in \mathbb{A}_o$, entonces su complemento $\phi' = \mathbb{S} - \phi = \mathbb{S}$ también pertenece a la colección $\mathbb{A}_o$.

De esto podemos concluir que, dado un conjunto Universal $\mathbb{S}$, la menor de las álgebras de Boole que se pueden construir es $\mathbb{A}_o = \{\phi, \mathbb{S}\}$.

Desde ya, la mayor será la clase $\mathbb{A}_1$, que contiene a *todos* los subconjuntos de $\mathbb{S}$.

Todas las álgebras Booleanas $\mathbb{A}$ que se pueden construir con subconjuntos de $\mathbb{S}$ satisfacen : $\mathbb{A}_o \subseteq \mathbb{A} \subseteq \mathbb{A}_1$.

6. BIBLIOGRAFIA

1 - Apóstol, T. M., *Calculus*, 2da edición, Reverté, S.A., Buenos Aires, 1984.

2 - Feller, W., *Introducción a la teoría de probabilidades y sus aplicaciones*, 2da edición, LIMUSA, México, D.F., 1978.

3 - Le Lionnais, F., *Las grandes corrientes del pensamiento matemático*, 2da edición, EUDEBA, Buenos Aires, 1965.

4 - Meyer, P. L., *Probabilidad y Aplicaciones Estadísticas*, Sistemas Técnicos de Edición, S.A. de C.V., México, D.F., 1986.