

Capítulo I

Síndrome de Bernoulli

1. Colectivos

Se tiene un experimento u observación, el cual se denotará por \mathcal{E} y se denominará *experimento*, que cada vez que se *ejecuta*, *efectúa* o *repite* se obtiene u observa uno y sólo uno de dos resultados posibles pero excluyentes (blanco o negro, águila o sol, prendido o apagado, sí o no, está o no está, pertenece a o no pertenece a, 0 ó 1, etc.) y que por simplicidad y conveniencia uno se denotará por el número 0 y el otro por el número 1; el conjunto constituido por estos dos resultados se denotará por $X_1 = \{0, 1\}$ y se le llamará *espacio de resultados de \mathcal{E}* .

Una de las características adicionales que se necesitan para \mathcal{E} , es que pueda *repetirse* un gran número de veces y esto tantas veces como se necesiten; *repetir*, *efectuar* o *ejecutar \mathcal{E}* en las mismas condiciones experimentales físicamente posibles; esto es, que una vez efectuado \mathcal{E} , se reproducen las condiciones físicas, se hacen los arreglos necesarios (sin que intervengan, entre otras cosas el resultado obtenido, 0 ó 1, anteriormente) para conseguir que sean las mismas condiciones ("volver las cosas a como se tenían antes"), en lo posible, para repetir \mathcal{E} . A este tipo de ejecuciones de \mathcal{E} algunos autores las llaman *repeticiones independientes*, tal como se les nominará aquí.

Un ejemplo prototípico de este tipo de parejas *experimento* \mathcal{E} - *espacio de resultados* $\{0, 1\}$ es el siguiente: en un recipiente se tienen bolas negras, 0, y blancas, 1, de igual tamaño y textura; el experimento \mathcal{E} consiste en revolver bien las bolas dentro del recipiente y sin ver dentro del recipiente, *a ciegas*, extraer una bola, si es posible “por medio de una mano inocente”; este experimento es una típica *extracción al azar*. Es razonable pensar que \mathcal{E} se efectúa con algún objetivo y no sólo por ocio; el objetivo que interesa es qué tipo de bola *sale*, negra (= 0) o blanca (=1); esto es, qué resultado en $X_1 = \{0, 1\}$ se obtiene. Una vez registrado el resultado obtenido, se devuelve la bola al recipiente, como una parte del proceso de conseguir que el experimento pueda *repetirse*, claro está, en las *mismas condiciones* e *independientemente* de la ejecución anterior.

Como puede colegirse, en las actividades humanas y las terrestres, y aún en la celeste, abundan este tipo de parejas *experimento* \mathcal{E} - *espacio de resultados* $\{0, 1\}$, así que en esta monografía se tratará, en principio, con parejas \mathcal{E} - $\{0, 1\}$ con las características

C1. \mathcal{E} puede repetirse un número de veces tan grande como sea necesario y esto siempre que se requiera.

pero en cuanto a los resultados posibles de efectuar \mathcal{E} , 0 ó 1, interesarán las parejas con la característica adicional acerca de la *indeterminación* del resultado cada vez que \mathcal{E} se ejecute:

C2. Las condiciones en que se efectúa \mathcal{E} , no determinan su resultado; es decir, los resultados de ejecutar \mathcal{E} , 0 ó 1, son aleatorios; por lo tanto,

el resultado de efectuar \mathcal{E} es imprevisible: a iguales causas, diferentes efectos.

En lo que sigue, estas parejas *experimento* \mathcal{E} - *espacio de resultados* $\{0, 1\}$, con las características $C1$ y $C2$, serán denotadas por el arreglo de símbolos (\mathcal{E}, X_1) .

El producto que interesará de estas parejas (\mathcal{E}, X_1) , será, en primer lugar, sus *colectivos de resultados*, esto es, las diferentes sucesiones de 0s y 1s que se obtienen en cada *corrida* o *tanda* de sucesiones de ejecuciones sucesivas y numerosas de \mathcal{E} ; se registrará cada colectivo en una tabla como la siguiente:

$$\begin{array}{l} \text{Ejecución de } \mathcal{E} \quad 1, \quad 2, \quad \dots, \quad n, \quad \dots \\ \text{Resultado de } X \quad \delta_1, \quad \delta_2, \quad \dots, \quad \delta_n, \quad \dots \end{array}$$

siendo δ_i 0 ó 1, el resultado de la ejecución i de \mathcal{E} en esa corrida de repeticiones independientes de \mathcal{E} . Y debido a la importancia de respetar el orden en que *salen* o se obtienen los resultados δ_i , importará, entonces, como *sucesiones* de 0s y 1s, los resultados obtenidos,

$$\delta_1, \delta_2, \dots, \delta_n, \dots$$

y a la cual se le nombrará *colectivo de resultados*, o *serie de tiempo*, de la pareja (\mathcal{E}, X_1) ; preservando ese orden, las sucesiones así obtenidas, para cada tanda de ejecuciones de \mathcal{E} , son *sucesiones aleatorias* de 0s y 1s.

En segundo lugar, lo que aquí se expondrá en relación a *los* colectivos de *algunas* parejas (\mathcal{E}, X_1) es cierto comportamiento común, comportamiento parecido de un colectivo a otro aunque no idéntico y esto debido al pecado

(o virtud) de origen: que los colectivos de una (\mathcal{E}, X_1) son siempre diferentes; es un comportamiento *parecido*, pero de *carácter estadístico*. Tal comportamiento de los colectivos de ciertos (\mathcal{E}, X_1) que aquí se tratará, será el llamado *síndrome de Bernoulli*.

Hay que advertir enfáticamente que no toda pareja (\mathcal{E}, X_1) produce colectivos con tal síndrome; es más, no se tiene un criterio general (aún no se ha descubierto) el cual diga qué tipo de experimentos u observaciones \mathcal{E} tiene, junto con X_1 , la propiedad de que sus colectivos de resultados adolezcan del síndrome de Bernoulli. Así que para cada pareja (\mathcal{E}, X_1) hay que analizar si sus colectivos tienen o no tal síndrome. Sin embargo hay innumerables parejas (\mathcal{E}, X_1) cuyos colectivos tienen tal comportamiento (en la mecánica cuántica, en la física estadística, en la demografía, en la transmisión de información, en los juegos de azar, y en un largo etcétera; ah, y por cierto, en el ejemplo dado en un principio); hay ciencias que se basan totalmente en este síndrome, lo cual avala justificadamente su estudio; en el aspecto teórico se tienen las teorías matemáticas de la probabilidad y de los procesos estocásticos. Así pues, su importancia, tanto teórica como práctica, es indiscutible.

Para la formulación de tal comportamiento se necesitan algunos conceptos analíticos y geométricos que a continuación se expondrán.

Para comenzar, se tienen un par de conceptos analíticos (equivalentes) útiles para el análisis cuantitativo de los colectivos de una pareja (\mathcal{E}, X_1) , asociados al número n de ejecuciones de \mathcal{E} y a los correspondientes n resul-

tados, en X_1 , que de éstas se obtengan.

Se efectua n ejecuciones de \mathcal{E} y se obtienen los correspondientes resultados

$$\delta_1, \delta_2, \dots, \delta_n \ ;$$

a) *frecuencia acumulada* (del resultado 1): es la suma

$$0 \leq \delta_1 + \dots + \delta_n \leq n,$$

la cual es, ni más ni menos que el número de veces que se obtuvo el resultado 1, de $X_1 = \{0, 1\}$, en esa corrida de ejecuciones de \mathcal{E} (o, equivalentemente, la frecuencia acumulada del resultado 0, de $X_1 = \{0, 1\}$, es la suma

$$0 \leq (1 - \delta_1) + \dots + (1 - \delta_n) \leq n \ ;$$

el número de veces, en esas n ejecuciones de \mathcal{E} , que se obtuvo el resultado 0);

b) *frecuencia relativa* (del resultado 1): es el cociente

$$0 \leq \frac{\delta_1 + \dots + \delta_n}{n} \leq 1,$$

que es la proporción del número de veces que se obtuvo el resultado 1 en esas n ejecuciones de \mathcal{E} , respecto a n ; o lo que es lo mismo, pero en otras palabras: es la proporción del número de 1s que *salieron* en esas n ejecuciones de \mathcal{E} respecto a n (o, equivalentemente, el cociente

$$0 \leq \frac{(1 - \delta_1) + \dots + (1 - \delta_n)}{n} \leq 1,$$

es la frecuencia relativa de 0s obtenidos en esas n ejecuciones de \mathcal{E} respecto a n).

Nótese que tanto la frecuencia relativa como la acumulada son números aleatorios *respecto* a n repeticiones de \mathcal{E} ; una vez más, *a iguales causas* (n repeticiones de \mathcal{E}), *diferentes efectos* (frecuencia acumulada o frecuencia relativa del resultado 1 ó 0).

2. Graficación de colectivos de (\mathcal{E}, X_1)

Otro recurso (camino, método o herramienta) que es un gran auxiliar para desvelar propiedades ocultas de los colectivos o series de tiempo de una pareja (\mathcal{E}, X_1) es su *graficación* (y esto no sólo para estos colectivos sino para cualquier sucesión de datos numericos), aunque aquí se utilizará una graficación muy peculiar:

Considérese una sucesión de 0s y 1s,

$$\delta_1, \delta_2, \dots, \delta_n, \dots;$$

tómense las parejas ordenadas

$$(k, \delta_1 + \dots + \delta_k), \text{ para } k = 1, 2, \dots, n, \dots,$$

y gráfiquense en el plano \mathbb{R}^2 ; lo que se obtendrá será una sucesión de puntos

$$(1, \delta_1), (2, \delta_1 + \delta_2), \dots, (n, \delta_1 + \dots + \delta_n), \dots$$

en el primer cuadrante de \mathbb{R}^2 , con, entre otras, las características siguientes

- i) El punto $(k + 1, \delta_1 + \dots + \delta_k + \delta_{k+1})$ está a la misma altura (tiene la misma ordenada) o una unidad más arriba (la ordenada +1) que el punto $(k, \delta_1 + \dots + \delta_k)$, ya que, por el tipo de sucesión se tienen que

$$\delta_1 + \dots + \delta_k \leq \delta_1 + \dots + \delta_k + \delta_{k+1},$$

y todo esto para $k = 1, 2, \dots, n, \dots$. La gráfica formada por esta sucesión de puntos es no decreciente.

ii) La grafica obtenida parte del punto $(1, 0)$ o del $(1, 1)$, según el valor de δ_1 , 0 ó 1. Para la abscisa k , la altura, la ordenada, del punto $(k, \delta_1 + \dots + \delta_k)$ es igual a la suma de 1s o número de 1s, o *frecuencia acumulada* de 1s, desde el término 1 hasta el k de la sucesión considerada.

iii) Como

$$0 \leq \delta_1 + \dots + \delta_k \leq k,$$

la gráfica de esta sucesión de puntos se encuentra en la región que resulta de intersecar las regiones $x \geq 1$, $y \geq 0$ y $x \geq y$ de \mathbb{R}^2 .

Por ejemplo, para la sucesión (o vector de $n = 9$ componentes) de 0s y 1s, de $n = 9$ términos

$$(1, 0, 0, 1, 0, 0, 0, 1, 1),$$

siguiendo lo dicho, se obtiene la sucesión de puntos de \mathbb{R}^2

$$(1, 1), (2, 1), (3, 1), (4, 2), (5, 2), (6, 2), (7, 2), (8, 3), (9, 4)$$

y cuya gráfica es la que se ve en la Fig.1.

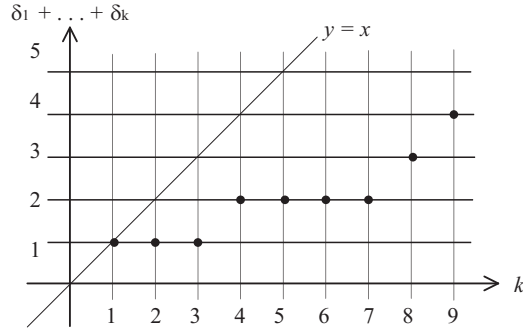


Figura 1: Gráfica de la sucesión $(1, 0, 0, 1, 0, 0, 0, 1, 1)$

Y viceversa: si se tiene una gráfica constituida por una sucesión de puntos del plano \mathbb{R}^2 , $(k, n(k))$ con abscisa natural $k \geq 1$ y ordenada un entero $n(k) \geq 0$ y esas propiedades geométricas, expresadas analíticamente,

$$i') 0 \leq n(k) \leq k$$

$$ii') n(k) \leq n(k+1) \leq n(k) + 1, \forall k = 1, 2, \dots$$

entonces, esta gráfica determina una y sólo una sucesión (o vector) de 0s y 1s, a saber

$$(n(1), n(2) - n(1), \dots, n(k+1) - n(k), \dots).$$

Por ejemplo, véase la gráfica compuesta por la sucesión de puntos de la figura anterior: sólo de verla se deduce la sucesión de 0s y 1s que la originó según el método descrito: el término 1 es 1; el término 2 es 0; el término 3 es 0; el término 4 es 1; el término 5 es 0; el término 6 es 0; el término 7 es 0; el término 8 es 1 y el término 9 es 1.

Con tal procedimiento pueden graficarse los colectivos, que se obtengan de cualquier pareja (\mathcal{E}, X_1) , con objeto de desvelar, de descubrir en ellos ten-

dencias, regularidades, periodicidades, etc., ya sea individualmente o colectivamente, es decir, tipos de comportamiento o características comunes.

3. Síndrome de Bernoulli

Ahora bien, al graficar los colectivos de *ciertas* parejas (\mathcal{E}, X_1) se ha descubierto que sus gráficas, se comportan *similarmente* (y las nuevas que se obtengan también lo harán) aunque no *identicamente*, de cierta manera muy especial; abusando del lenguaje: es como si sus colectivos padecieran la *misma enfermedad*, el *mismo complejo*, el *mismo síndrome*, el *mismo pecado* o, si se quiere, la *misma virtud*, aunque claro está, con sus peculiaridades individuales, como toda enfermedad, complejo, síndrome, pecado o virtud.

Este comportamiento muy especial, que se llamará *síndrome de Bernoulli*, y que se da o manifiesta en los colectivos de sólo ciertas parejas *experimento-espacio de resultados* $\{0, 1\}$ con las características $C1$ y $C2$, consiste, se reconoce, por el hecho de que sus gráficas aparecen *correteando* o *persiguiendo*, a veces por arriba, a veces por abajo, a una recta que parte del origen y cuya pendiente es mayor que cero pero menor que uno (pendiente que depende del experimento de aquella pareja), pero además que la distancia vertical de cada gráfica a esa recta no es grande sino más bien pequeña y esto respecto al número de repeticiones del experimento, particularidad que se acentúa siempre que este número sea cada vez más grande.

Con objeto de aclarar más esta descripción, se formulará tanto en términos geométricos como analíticos: tómesese una de esas parejas *experimento-espacio de resultados* $\{0, 1\}$ con las características $C1$ y $C2$ y convéngase que

ésta es (\mathcal{E}, X_1) , y obténgase un colectivo de resultados

$$\begin{array}{l} \text{Ejecución de } \mathcal{E} \quad 1, \quad 2, \quad \dots, \quad n, \quad \dots \\ \text{Resultado} \quad \delta_1, \quad \delta_2, \quad \dots, \quad \delta_n, \quad \dots \quad ; \end{array}$$

al ir graficando los puntos

$$(k, \delta_1 + \dots + \delta_k), \quad k = 1, 2, \dots, n, \dots$$

en \mathbb{R}^2 , se observa que estos puntos van *correteando* o *persiguiendo* una recta, cuya ecuación es de la forma $y = px$, con $0 < p < 1$, pero que, además, las distancias entre los puntos $(k, \delta_1 + \dots + \delta_k)$ y (k, pk) , las distancias verticales de los puntos de la gráfica del colectivo a la recta $y = px$ para los valores $x = 1, 2, \dots$,

$$|\delta_1 + \dots + \delta_k - pk|$$

son muy pequeñas en relación al tamaño de k , acentuándose este hecho siempre y cuando k es grande.

La Figura 2 da una idea del comportamiento geométrico de la gráfica de un colectivo con el síndrome de Bernoulli.

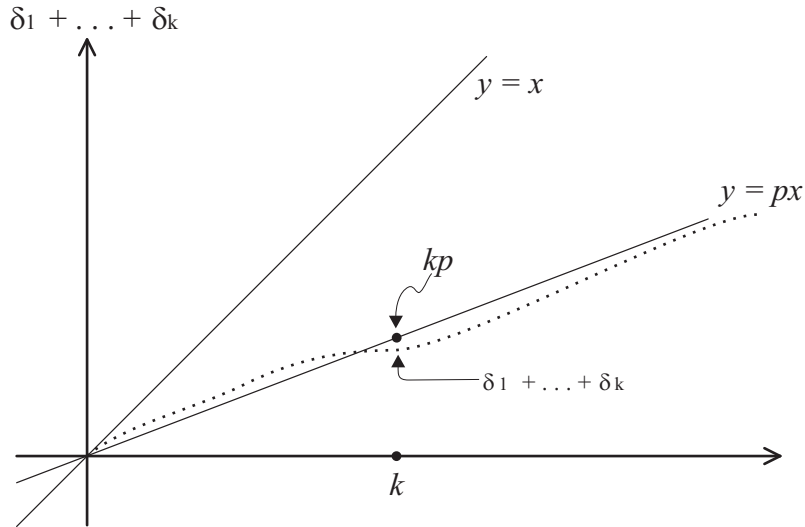


Figura 2: Gráfica de un colectivo con el síndrome de Bernoulli.

En esta *persecución*, el colectivo a veces *corre por arriba* de la recta, lo que analíticamente se expresa en la forma

$$\delta_1 + \dots + \delta_n - np > 0,$$

para varios valores de n consecutivos; pero otras corre por *abajo* de la recta, en cuyo caso, analíticamente, se tendrá que

$$\delta_1 + \dots + \delta_n - np < 0,$$

para valores de n consecutivos.

Los *tramos*, número de ejecuciones consecutivas de \mathcal{E} , durante los cuales el colectivo corretea a la recta ya sea por arriba, o por abajo, por lo general son *largos*, sucediendo que estos tramos se van alargando más y más uno después del otro. Pero, sin embargo, manteniéndose siempre esa cercanía del

colectivo a la recta, esto es, manteniéndose la distancia

$$|\delta_k + \delta_{k+1} + \cdots + \delta_{k+n} - (n+1)p|$$

muy pequeña respecto a $n+1$, cuando n es grande.

Todo esto manifiesta, además, que el síndrome de Bernoulli *se mantiene* a lo largo del mismo colectivo, lo cual no es extraño, ya que de la tanda de ejecuciones original, cada ejecución que se tome, digamos k , con las subsecuentes, el colectivo correspondiente

$$\delta_k, \delta_{k+1}, \dots, \delta_{k+n}, \dots$$

es ni más ni menos un colectivo más de resultados de (\mathcal{E}, X_1) ; como se ha establecido, la primeras $k-1$ ejecuciones de \mathcal{E} , no influyen, son independientes, de las que se efectúen posteriormente; algunos autores expresan este hecho en forma dramática, diciendo que el proceso de ejecuciones sucesivas de \mathcal{E} no tiene “memoria”.

De esta forma, puede conseguirse un *colectivo* de colectivos de resultados de (\mathcal{E}, X_1) ,

$$\delta_1, \delta_2, \dots, \delta_n, \dots; \quad \delta_2, \delta_3, \dots, \delta_{n+1}, \dots; \quad \dots; \quad \delta_k, \delta_{k+1}, \dots, \delta_{k+n}, \dots; \quad \dots$$

todos los cuales tienen el síndrome de Bernoulli

Conviene aclarar, antes de continuar, que la pendiente de esa recta, p , es una propiedad intrínseca de \mathcal{E} , en relación con sus posibles resultados, 0 ó 1, y, específicamente aquí, con el resultado 1; en el ejemplo dado al principio resulta que p es la proporción de bolas blancas respecto al número de bolas en el recipiente, sin embargo no se tiene un método general por medio del

cual pueda deducirse p a partir de \mathcal{E} , cuando los colectivos de resultados de (\mathcal{E}, X_1) padezcan, o disfruten, el síndrome de Bernoulli.

Así pues, esa gráfica del colectivo $\delta_1, \delta_2, \dots, \delta_n, \dots$, es prototípica de las gráficas de los colectivos de esta pareja (\mathcal{E}, X_1) cuyos colectivos padecen o disfrutan el síndrome de Bernoulli.

Una corroboración geométrica de esta propiedad se observa *amontonando* las gráficas de un buen número de colectivos de la pareja (\mathcal{E}, X_1) : se obtienen colectivos correspondientes a *tandas* o *corridas* de ejecuciones de \mathcal{E}

<i>Corrida 1:</i>	<i>Ejecución de \mathcal{E}</i>	1, 2, ... , n, ...
	<i>Resultado</i>	$\delta_1^1, \delta_2^1, \dots, \delta_n^1, \dots$
<i>Corrida 2:</i>	<i>Ejecución de \mathcal{E}</i>	1, 2, ... , n, ...
	<i>Resultado</i>	$\delta_1^2, \delta_2^2, \dots, \delta_n^2, \dots$
\vdots	\vdots	\vdots
<i>Corrida N:</i>	<i>Ejecución de \mathcal{E}</i>	1, 2, ... , n, ...
	<i>Resultado</i>	$\delta_1^N, \delta_2^N, \dots, \delta_n^N, \dots$
\vdots	\vdots	\vdots

y se grafica cada uno de éstos, amontonándolos, obteniendo una figura como la siguiente

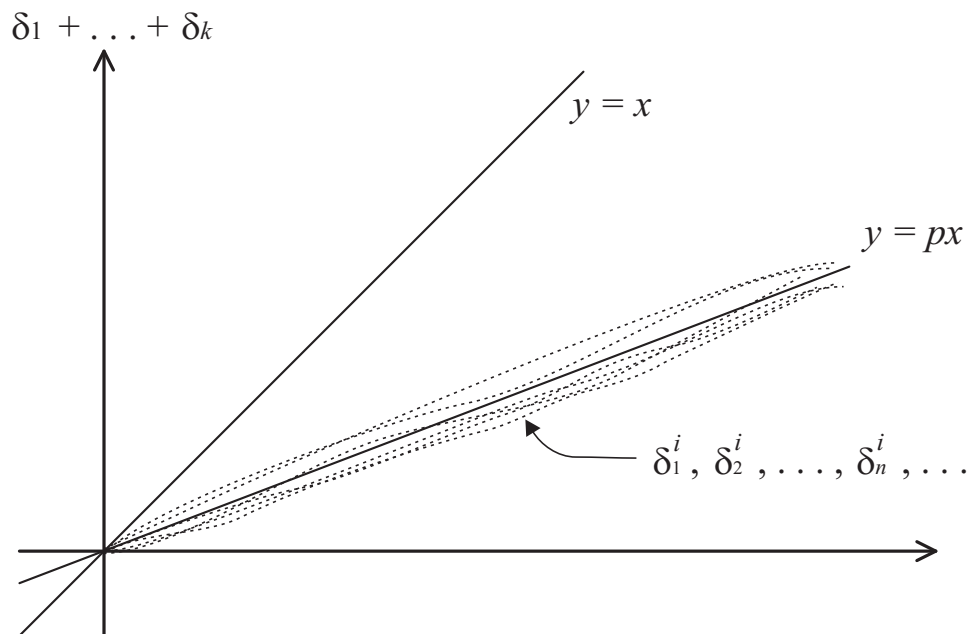


Figura 3: Gráficas de los colectivos de la pareja (\mathcal{E}, X_1)

Las gráficas de esos colectivos así amontonadas conforman una figura similar a la de un cometa, cuya cauda va abriéndose, ampliándose, lenta pero muy lentamente a medida que ésta se va alargando (esto es, a medida que el número k , de ejecuciones de \mathcal{E} en cada tanda, crece) pero siguiendo y desplegándose, simétricamente, a la recta $y = px$. La cauda tiene una forma muy regular, en el sentido de que no hay colectivos cuyas gráficas se disparen de las demás, que se aparten significativamente del eje de simetría $y = px$. Es como si las gráficas siguieran un *mismo* guión previamente establecido; fenómeno extraño este síndrome de Bernoulli, en vista de que cada colectivo es producto de una sucesión de repeticiones independientes de \mathcal{E} con resultados aleatorios.

Y lo cual hace pensar que debe existir una ley física, por cierto aún no descubierta y menos aún establecida, que rijan, pero que además explique, este fenómeno.

Ahora bien, este comportamiento geométrico de ir abriéndose lenta pero muy lentamente de la cauda conformada por el amontonamiento de las gráficas de los colectivos de resultados de (\mathcal{E}, X_1) con el síndrome de Bernoulli, es la (o una) imagen geométrica de la faceta cuantitativa del síndrome de Bernoulli que se expresa analíticamente, como que *la distancia*, y esto para cada $i = 1, 2, \dots, N, \dots$,

$$|\delta_1^i + \dots + \delta_n^i - np|$$

se va haciendo mucho menor que n a medida que n va creciendo. O, puede verse, en sentido contrario: que esta proposición cuantitativa es la (o una) expresión analítica de esa particularidad geométrica del síndrome de Bernoulli.

Esta propiedad cuantitativa del síndrome de Bernoulli es la más conspicua (!claro que se han descubierto otras, aunque no tan primarias como ésta) y la que primero llamó la atención, en tal medida que ha sido uno de los temas creadores, impulsores, originadores, de la teoría matemática de la probabilidad. Así que, en vista de su relevancia, y de los objetivos de la presente monografía, se resaltarán su formulación, además de bautizarla, en la sección siguiente.

4. Ley de los números grandes de Bernoulli

Esta faceta del síndrome de Bernoulli de los colectivos de resultados de ciertas parejas (\mathcal{E}, X_1) , específicamente acerca del *comportamiento del ancho*

de la cauda del cometa, formado por las graficas de sus colectivos, aunado al “*concepto de aleatoriedad*” es, desde la percepción del autor, el *germen*, la *esencia* o el *fenómeno físico central* de las teorías matemáticas de la probabilidad y de los procesos estocásticos. Se aclara de inmediato que tal afirmación es una convicción personal y que con ésa de ninguna manera se intenta demeritar la multitud de resultados o interpretaciones que se han dado, y seguirán dándose en y para esas teorías matemáticas.

Aquel comportamiento merece ser subrayado, sin olvidar al matemático que la intuyó (en una forma diferente aunque equivalente a la aquí expuesta) e intentó formularla matemáticamente por vez primera. Tal comportamiento se denominará *Ley de los números grandes de Bernoulli* (Jakob o James Bernmoulli).

Para formular esta ley o fenómeno, inherente al síndrome de Bernoulli, se echará mano de una notación útil, intuitiva y cotidiana de los físicos: para números reales positivos a y b , el símbolo $a \gg b$ significa que a es mucho mayor que b o que b es mucho menor que a , y esto último también suele denotarse por el símbolo $b \ll a$.

Así pues, el comportamiento de la distancia vertical de la gráfica de cada colectivo de resultados $\delta_1, \delta_2, \dots, \delta_n, \dots$ de la pareja (\mathcal{E}, X_1) , cuyos colectivos tienen el síndrome de Bernoulli, a la recta que persigue conforme n crece, $y = px$, queda expresada en la forma

$$|\delta_1 + \dots + \delta_n - np| \ll n, \forall n \gg 1;$$

así que, para el colectivo de colectivos de resultados que configuran el cometa

de la sección anterior se tendrá:

$$|\delta_1^i + \dots + \delta_n^i - np| \ll n, \forall n \gg 1$$

y para cada $i = 1, 2, \dots, N, \dots$. Y un buen número n de repeticiones de \mathcal{E} a partir del cual $n \gg 1$, según la experiencia, es $n \approx 1000$, n alrededor de 1000.

Es importante ver, notar, que esas expresiones que acotan la distancia del colectivo a la recta *no dependen* del colectivo particular (es la misma para todos los colectivos de (\mathcal{E}, X_1) , sólo dependen de ese número p , que a su vez depende de \mathcal{E}), lo que resaltan, además de expresar, el carácter *uniforme* (“pero no idéntico”), *regular*, del comportamiento de los colectivos con el síndrome de Bernoulli. Y esto es un típico ejemplo de lo que conoce por *regularidad estadística*.

Más aún, éstas relaciones analíticas, también expresan que el número de veces que se obtiene el resultado 1, se *conserva*, se *mantiene*, *casi igual*, *proporcionalmente* a n , el número de ejecuciones de \mathcal{E} , siendo el factor de proporcionalidad p , y esto para cada colectivo de (\mathcal{E}, X_1) y n grande.

Otra forma de expresar esta faceta, peculiaridad, del síndrome de Bernoulli, y equivalente a la de arriba (y más ortodoxa, cabe agregar), ya que expresa lo asentado en el párrafo anterior, puede verse geométricamente: la *pendiente* de la gráfica de cada colectivo (ya que la gráfica *va muy pegada*, o se *parece* mucho observandola desde lejos, a la recta $y = px$ y que la gráfica del colectivo parte de punto $(1, 0)$ o $(1, 1)$) de la ejecución 1 a la n de \mathcal{E} , o del término 1 al n del colectivo, visto como sucesión, es la altura de la gráfica sobre n , es aproximada a p ,

$$\frac{\delta_1^i + \dots + \delta_n^i}{n} \approx p,$$

donde \approx significa aproximado; sin embargo, dividiendo por n ambos miembros, de la expresión de arriba, es plausible, que

$$\left| \frac{\delta_1^i + \dots + \delta_n^i}{n} - p \right| \ll 1, \forall n \gg 1,$$

lo cual corrobora la anterior aproximación, y esto para cada colectivo $i = 1, 2, \dots, N, \dots$.

Esta expresión analítica, y equivalente a la expresión dada más arriba, expresa, así puede interpretarse, que para toda n grande, la proporción de resultados 1 obtenidos en n ejecuciones de \mathcal{E} consecutivas, o la proporción de ejecuciones de \mathcal{E} en las cuales salió 1, respecto al número n de ejecuciones consecutivas efectuadas,

$$\frac{\delta_1^i + \dots + \delta_n^i}{n},$$

es *casi* p .

Así, tanto en el aspecto interpretativo, como el geométrico y analítico, se tiene que ambas expresiones analíticas son equivalentes:

$$|\delta_1^i + \dots + \delta_n^i - np| \ll n, \forall n \gg 1 \Leftrightarrow \left| \frac{\delta_1^i + \dots + \delta_n^i}{n} - p \right| \ll 1, \forall n \gg 1.$$

Resumiendo: si los colectivos de resultados de la pareja (\mathcal{E}, X_1) tienen el síndrome de Bernoulli y la recta que persiguen es $y = px$, $0 < p < 1$, la característica especial de este síndrome, en relación al comportamiento de la distancia vertical entre cada colectivo

$$\begin{array}{l} \text{Ejecución de } \mathcal{E} \quad 1, \quad 2, \quad \dots, \quad n, \quad \dots \\ \text{Resultado} \quad \delta_1, \quad \delta_2, \quad \dots, \quad \delta_n, \quad \dots \quad ; \end{array}$$

y la recta $y = px$,

$$|\delta_1 + \dots + \delta_n - np| \ll n, \forall n \gg 1,$$

o, equivalentemente, al comportamiento de la *pendiente del colectivo* en relación a la pendiente de la recta

$$\left| \frac{\delta_1 + \dots + \delta_n}{n} - p \right| \ll 1, \forall n \gg 1,$$

es llamada *ley de los números grandes de Bernoulli*.

Estas relaciones analíticas, a veces, por brevedad, se denotarán en las formas

$$\delta_1, \dots, \delta_n \approx np, \forall n \gg 1$$

y

$$\frac{\delta_1, \dots, \delta_n}{n} \approx p, \forall n \gg 1,$$

respectivamente.

A las parejas (\mathcal{E}, X_1) cuyos colectivos tengan el síndrome de Bernoulli, se les denotará en la forma $(\mathcal{E}, X_1)_p$, siendo el índice p la pendiente de la recta $y = px$ cuya participación ya se ha descrito; además, a este tipo de parejas se les llamará *experimentos de Bernoulli* y a sus colectivos, *colectivos de Bernoulli*. Así que cuando se utilice la frase “experimento de Bernoulli” se estará refiriendo a alguna pareja compuesta por un experimento u observación \mathcal{E} y un conjunto $X_1 = \{0, 1\}$, cuyos elementos son los resultados que interesan al efectuarse \mathcal{E} , y que además tiene las características C_1 y C_2 y cuyos colectivos de resultados tienen el síndrome de Bernoulli, los cuales persiguen una recta de la forma $y = px$, $0 < p < 1$, teniendo en cuenta que, como *una* de las peculiaridades del síndrome de Bernoulli, se cumple la ley de los números grandes de Bernoulli correspondiente. Y que cuando diga “colectivo de Bernoulli”, se estará uno refiriendo a un colectivo de 0s y 1s de un experimento de Bernoulli.

5. Acerca de p

En esta sección se expondrán algunas propiedades adicionales del papel que juega el parámetro p para los colectivos de resultados de un experimento de Bernoulli $(\mathcal{E}, X_1)_p$ y por medio de o basándose en la ley de los números grandes de Bernoulli, ésta como una faceta analítica, de tipo cuantitativo, del comportamiento de estos colectivos, el síndrome de Bernoulli.

Quizá se repitan aquí algunas observaciones o señalamientos que se hayan hecho al respecto, repetición que no puede ser dañina, al contrario; ver un objeto desde distintos lados ayuda a comprenderlo y a conocerlo más.

1. Los experimentos de Bernoulli se encuentran en las más diversas situaciones físicas, naturales o experimentales, siendo su número p correspondiente una constante física inherente al experimento \mathcal{E} y a su resultado 1, uno de los dos de sus dos posibles resultados, 0 y 1. p dicta o desvela una regularidad, observable, repetible, mensurable y verificable.

¿Hay alguna ley física que ligue \mathcal{E} con su respectivo real p , en relación con su resultado 1? Debe existir, según lo que se corrobora experimentalmente una y otra vez con el comportamiento de sus colectivos. Pero ¿Cómo conocer p a partir de \mathcal{E} ? Sólo en pocos casos se sabe, por ejemplo en los juegos de azar, pero no en esa gran mayoría de situaciones donde se encuentran estos experimentos de Bernoulli; en éstas hay que hacer consideraciones indirectas (simetrías, etc.) pero a fin de cuentas experimentando con multitud de sus colectivos y..., adivinar p .

2. Cada uno de sus colectivos, vistos como sucesiones de 0s y 1s, son singulares, con “personalidad propia”, debido a la aleatoriedad intrínseca (Característica $C2$); en este sentido, dos de estos colectivos son del todo

diferentes, es más no existe un método por medio del cual, a partir de su sucesión, puedan hacerse inferencias del otro, o ni siquiera a partir del tramo de una, lo que pasará después en este mismo colectivo, como sucesión de 0s y 1s : dos tramos largos es rarísimo que sean iguales, *umentando* esto entre mas largos sean los tramos tomados.

Sin embargo, sus gráficas desvelan un comportamiento *similar*, el síndrome de Bernoulli que subyace (y como una particularidad de éste, un aspecto analítico de su comportamiento, expresado como la ley de los números grandes de Bernoulli), se oculta, no es manifiesto, si sólo se consideran como sucesiones de 0s y 1s, y muestran el papel central de p a lo largo de un mismo colectivo:

$$\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1}, \dots, \delta_{n+m}, \dots$$

$$\delta_1 + \dots + \delta_n \approx np \quad \text{ó} \quad \frac{\delta_1 + \dots + \delta_n}{n} \approx p, \quad \forall n \gg 1$$

pero también

$$\delta_1 + \dots + \delta_n + \delta_{n+1}, \dots, \delta_{n+m} \approx (n+m)p \quad \text{ó} \quad \frac{\delta_1 + \dots + \delta_n + \delta_{n+1}, \dots, \delta_{n+m}}{n+m} \approx p,$$

$\forall n \gg 1$;pero, además, en cada colctivo de colectivos (Sección 3)

$$\delta_1^1, \quad \delta_2^1, \quad \dots, \quad \delta_n^1, \quad \dots$$

$$\delta_1^2, \quad \delta_2^2, \quad \dots, \quad \delta_n^2, \quad \dots$$

$$\vdots \quad \vdots \quad \quad \quad \vdots$$

$$\delta_1^N, \quad \delta_2^N, \quad \dots, \quad \delta_n^N, \quad \dots,$$

$$\delta_1^i + \dots + \delta_n^i \approx np \quad \text{ó} \quad \frac{\delta_1^i + \dots + \delta_n^i}{n} \approx p, \quad \forall n \gg 1,$$

(recuérdese el sentido que se le ha dado a \approx).

Estos hechos, el comportamiento geométrico similar y el de las expresiones analíticas \approx , que *no idéntico ni igual* (= por \approx), sino con el carácter de aproximados en general, *carácter* o *regularidad estadística* suele decirse, está resumido en el número p , vía ley de los números grandes de Bernoulli.

3. Otro aspecto que se da con los colectivos de $(\mathcal{E}, X_1)_p$ es el comportamiento *autosimilar* de cada uno de éstos, en el sentido siguiente: a partir de un solo colectivo obtenido mediante una corrida de ejecuciones de \mathcal{E} ,

$$\delta_1, \delta_2, \dots, \delta_n, \dots$$

se forma la sucesión de colectivos

$$\begin{array}{ccccccc} \delta_1, & \delta_2, & \dots & , & \delta_n, & \dots & \\ & & & & & & \\ \delta_2, & \delta_3, & \dots & , & \delta_{n+1}, & \dots & \\ & & & & & & \\ \vdots & \vdots & & & \vdots & & \\ & & & & & & \\ \delta_k, & \delta_{k+1}, & \dots & , & \delta_{k+(n+1)}, & \dots & \\ & & & & & & \\ \vdots & \vdots & & & \vdots & & \end{array}$$

es como si se tomara la ejecución k de aquella corrida de ejecuciones de \mathcal{E} como la primera, para cada $k = 1, 2, \dots$.

Si se grafican estos colectivos, se obtiene que conforman un cometa semejante a los ya descritos: que extiende su cauda a lo largo de y simetricamente a la recta $y = px$, con las particularidades de que cada subcolectivo corretea a la recta, a veces por arriba y a veces por abajo y con la distancia vertical

del colectivo a la recta siguiendo la ley de los números grandes de Bernoulli

$$|\delta_k + \dots + \delta_{k+n} - (n+1)p| \ll 1, \quad \forall n \gg 1$$

o, equivalentemente,

$$\frac{\delta_k + \dots + \delta_{k+n}}{n+1} \ll 1, \quad \forall n \gg 1.$$

Esto es que los colectivos tienen el síndrome de Bernoulli.

Así, el comportamiento de

$$\delta_1, \delta_2, \dots, \delta_n, \dots$$

es similar al de

$$\delta_2, \delta_3, \dots, \delta_{n+1}, \dots$$

y al de

$$\delta_3, \delta_4, \dots, \delta_{n+2}, \dots$$

etc.

En este sentido es que el comportamiento de

$$\delta_1, \delta_2, \dots, \delta_n, \dots$$

es *autosimilar*: *conserva* o se *mantiene* el síndrome de Bernoulli a lo largo de cada colectivo, y en particular el comportamiento de su ley de los números grandes de Bernoulli, y con su misma p .

4.Lo anterior, además, muestra el fenómeno natural sorprendente, del síndrome de Bernoulli, otra faceta de este que hay que recalcar tanto por su aspecto teórico como por sus aplicaciones y es que *la naturaleza*, o el síndrome de Bernoulli, o la sucesión de repeticiones *independientes* de \mathcal{E} en relación

a sus resultados posibles, 0 ó 1, se las arregla para *mantener* un equilibrio, en tramos largos del colectivo, y lo cual se da para más o menos el mismo número de ejecuciones [asentamiento], el número de veces del resultado 1 (y por lo tanto el resultado 0) i.e. la ley de los números grandes de Bernoulli

$$|\delta_k + \dots + \delta_{k+n} - (n+1)p| \ll n, \quad \forall n \gg 1$$

o, equivalentemente,

$$\frac{\delta_k + \dots + \delta_{k+n}}{n+1} \ll 1, \quad \forall n \gg 1.$$

Y esto para la misma n grande, puede decirse para $n \approx 1000$; pero además para cada colectivo

$$\delta_1^i, \delta_2^i, \dots, \delta_n^i, \dots \quad :$$

$$|\delta_1^i + \dots + \delta_n^i - np| \ll n, \quad \forall n \gg 1$$

o, equivalentemente

$$\left| \frac{\delta_1^i + \dots + \delta_n^i}{n} - p \right| \ll 1, \quad \forall n \gg 1.$$

En otros términos, que para cada n grande, sucede

$$\delta_k + \dots + \delta_{k+n} \approx \delta_l + \dots + \delta_{l+n}$$

aquí \approx se toma como aproximado, o

$$\delta_1^i + \dots + \delta_n^i \approx \delta_1^j + \dots + \delta_n^j, \quad i \neq j$$

en cada colectivo de colectivos de resultados. O en términos de proporciones, se mantienen casi iguales:

$$\frac{\delta_k + \dots + \delta_{k+n}}{n+1} \approx \frac{\delta_l + \dots + \delta_{l+n}}{n+1}$$

y

$$\frac{\delta_1^i + \cdots + \delta_n^i}{n} \approx \frac{\delta_l + \cdots + \delta_{l+n}}{n}$$

siempre para valores grandes de n , y un buen valor de n para el cual sucede esto es $n \approx 1000$; esto es: se da el fenómeno de mantener casi la misma proporción de veces que sale el resultado 1, en grandes tramos, numerosas ejecuciones consecutivas de \mathcal{E}

5. Una propiedad adicional que debe ser subrayada, de manera enfática, acerca de los colectivos de Bernoulli y que la muestra su faceta que es la ley de los números grandes, es el hecho de que a lo largo de cada uno de ellos, para *cada* tramo de un número grande de ejecuciones consecutivas de \mathcal{E} , n , se las arregla para que el número de resultados 1 sea un poco más o un poco menos que np :

$$|\delta_{k+1} + \cdots + \delta_{k+n} - np| \ll n,$$

$$\delta_{k+1} + \cdots + \delta_{k+n} \approx np,$$

en otras palabras, que la proporción de veces que se obtiene el resultado 1 o el número de ejecuciones que se produce el resultado 1 es muy proximo, casi p ,

$$\left| \frac{\delta_{k+1} + \cdots + \delta_{k+n}}{n} - p \right| \ll 1$$

$$\frac{\delta_{k+1} + \cdots + \delta_{k+n}}{n} \approx p.$$

Pero esto sucede para el mismo tamaño, más o menos, de n esto es

$$\delta_{k+1}^i + \cdots + \delta_{k+n}^i \approx \delta_{k+1}^j + \cdots + \delta_{k+n}^j$$

o,

$$\frac{\delta_{k+1}^i + \cdots + \delta_{k+n}^i}{n} \approx \frac{\delta_{k+1}^j + \cdots + \delta_{k+n}^j}{n}$$

y esto para cada k : en un mismo colectivo

$$\delta_1, \delta_2, \dots, \delta_n, \dots; \delta_2, \delta_3, \dots, \delta_{n+1}, \dots; \delta_{k+1}, \dots, \delta_{k+n}, \dots$$

o

$$\delta_{k+1}, \dots, \delta_{k+n} \approx \delta_{l+1}, \dots, \delta_{l+n}$$

y esto para cada k y l .

En resumen, en cada colectivo de Bernoulli, en cada tramo de n repeticiones consecutivas de \mathcal{E} se obtendrán que, aproximadamente en np de éstas, saldrá el resultado 1 y en $n = (1-p)$ el resultado 0. Pero ojo, algo importante, como esto sucede en cada

$$\delta_1, \delta_2, \dots, \delta_n, \dots; \delta_2, \delta_3, \dots, \delta_{n+1}, \dots; \delta_{k+1}, \dots, \delta_{k+n}, \dots$$

lo que nos dice, además de ser verificado, que nunca se verán sucesiones de resultados de la forma

$$1, 1, 1, \dots, 1, \dots, 0, 0, 0, \dots, 0, \dots$$

es decir

$$\delta_1 + \dots + \delta_{1000} = 1000 \text{ y } \delta_{1001}, \dots, \delta_{2000} = 0$$

por ejemplo, aunque $p = \frac{1}{2}$: Por eso es tan difícil construir una sucesión de 0s y 1s como las originadas por extracciones consecutivas de bolas, o repeticiones consecutivas de \mathcal{E} .

6. Otra propiedad de los colectivos de Bernoulli es que si de *cualquier manera* se seleccionan los naturales

$$k_1, k_2, \dots, k_m, \dots$$

con $k_i \leq k_{i+1}$ (una subsección de dos naturales) y se obtiene uno de ellos mediante una corrida de ejecuciones sucesivas de \mathcal{E}

$$\delta_1, \delta_2, \dots, \delta_n, \dots$$

y se toman los resultados de las ejecuciones $k_1, k_2, \dots, k_m, \dots,$

$$\delta_{k_1}, \delta_{k_2}, \dots, \delta_{k_m}, \dots$$

resulta que

$$|\delta_{k_1} + \dots + \delta_{k_m} - mp| \ll m$$

o

$$\frac{\delta_{k_1} + \dots + \delta_{k_m}}{m} \approx p, \quad \forall m \gg 1.$$

Todo esto como uno de los aspectos del síndrome de Bernoulli que tienen los colectivos de Bernoulli, o sea que

$$\delta_{k_1}^1, \quad \delta_{k_2}^1, \quad \dots, \quad \delta_{k_m}^1, \quad \dots$$

$$\delta_{k_1}^2, \quad \delta_{k_2}^2, \quad \dots, \quad \delta_{k_m}^2, \quad \dots$$

$$\vdots \quad \quad \quad \vdots$$

$$\delta_{k_1}^M, \quad \delta_{k_2}^M, \quad \dots, \quad \delta_{k_m}^M, \quad \dots$$

$$\vdots \quad \quad \quad \vdots$$

7. Después de estos señalamientos acerca de los colectivos de $(\mathcal{E}, X_1)_p$ es conveniente comentar el término *probabilidad* usado para p . ¿Por qué se usa? ¿Cuando se dice *probabilidad* p , a qué se refiere? ¿En qué contexto es adecuado, prudente, utilizarlo? ¿Cuándo se dice que p es una *medida de probabilidad*, a qué se refiere, cómo se interpreta?.

En primer lugar nótese que el número p , por el papel que juega como *dictador* o, en términos más benignos, p *mantiene el equilibrio*, acerca del número de ejecuciones de \mathcal{E} cuyo resultado es 1 es una sucesión numerosa de ejecuciones de \mathcal{E} : en n ejecuciones de \mathcal{E} , aproximadamente en np de éstas se obtiene el resultado 1; o, normalizado, la proporción de ejecuciones en las que se obtiene el resultado 1 de un gran número de ejecuciones de \mathcal{E} , es aproximadamente p (esto de aproximaciones según la ley de los números grandes).

Así, si p es grande, cercano a 1, el número de resultados 1 en un gran número de ejecuciones sucesivas de \mathcal{E} , es grande; si p es pequeño, cercano a 0, en una sucesión de ejecuciones numerosas de \mathcal{E} , el número de veces que se obtiene el resultado 1, es pequeño; si p es cercano a $\frac{1}{2}$, el número de ejecuciones de \mathcal{E} , cuyo resultado es 1 de un gran número de ejecuciones sucesivas de \mathcal{E} , es aproximadamente $\frac{1}{2}$ de éste. Y así puede seguirse observando según el valor de p en el intervalo $[0, 1)$. Todo esto muestra que el número p de un experimento de Bernoulli, $(\mathcal{E}, X_1)_p$, puede verse, interpretarse, también, como *una medida de qué tan frecuente se obtiene el resultado 1 en una sucesión numerosa de ejecuciones de \mathcal{E}* . (Y, por consecuencia, $1 - p$, una medida respectiva para el resultado 0). Recuérdese, de ejecuciones o *repeticiones*, independientes, de \mathcal{E} .

p *mantiene el equilibrio*, no de manera exacta, rígida, acerca del resultado 1 con respecto al resultado 0 (o como $1 - p$ acerca del resultado 0), sino aproximada según la ley de los números grande de Bernoulli. Y manteniendo este equilibrio a lo *largo* de cada colectivo: a veces un poco más, a veces un poco menos, pero siempre, cercano a p .

Ahora bien, si p es grande (cercano a 1) en una sucesión de ejecuciones de

\mathcal{E} el resultado 1 saldra muy seguido, muy frecuente, ¿cuándo sale 1? ¡quien sabe! ni el diablo, ni por viejo ni por diablo, lo sabe; aunque eso sí manteniéndose el equilibrio, o desequilibrio, del que se habló arriba. Lo cual induce a pensar que en la siguiente ejecución, o en una ejecución de \mathcal{E} , es muy probable que se obtenga el resultado 1, y, como consecuencia, muy improbable que en ésa salga un 0. Y, lo contrario si p es cercano a 0.

Pero si p es cercano a 0,5, por ejemplo 0,51, hay una ligera, pudiera decirse, una probabilidad *mayor* de que el resultado 1 se obtenga en una solá ejecución de \mathcal{E} . Y p es por ejemplo 0,49 la probabilidad de que salga el resultado 1 en una ejecución de \mathcal{E} es ligeramente *inferior* de que salga el resultado 0.

Ahora bien, aunque el experimento de Bernoulli $(\mathcal{E}, X_1)_p$ se da en el contexto fundamental, primario, y por lo tanto tiene sentido y validez en referencia a sucesiones de numerosas repeticiones de \mathcal{E} , además del marco de aleatoriedad, del síndrome de Bernoulli de sus colectivos, de la faceta particular de éste como en la ley de los números grandes de Bernoulli, sería natural, y es común hacerlo, *convenir* en cuantificar esas probabilidades.

Un hecho curioso es que, si uno intenta construir una sucesión de 0s y 1s que cumpla con estas condiciones, es casi imposible hacerlo, sin embargo si lo hace por medio de extracciones al azar de bolas, unas marcadas con 0 y otras marcadas con 1, como en el ejemplo, entonces obtiene lo que se queria.

Una observación adicional: el fenómeno recién descrito o esa propiedad que tienen los colectivos de Bernoulli, correspondientes a $(\mathcal{E}, X_1)_p$, se observa para un número de repeticiones consecutivas, n , no muy grande como lo especifica la ley de los números grandes de Bernoulli, el fenómeno descrito

se observa plausiblemente cuando n anda alrededor de 100; es más, si de un colectivo de Bernoulli se toman por tramos, de digamos de 100 en 100 ejecuciones

$$\delta_1, \dots, \delta_{100}, \delta_{101}, \dots, \delta_{200}, \dots$$

se verá que sus graficas son parecidas ni se dispara el número de resultados 1 ni el de resultados 0. Aunque, cuando las cosas se ven del todo claras, en cuanto n anda alrededor de 100, y de esto en adelante, como lo especifica la ley de los números grandes de Bernoulli.

Es más, hay que decirlo, recalcarlo aunque sea obvio: algo semejante sucede en relación al resultado 0: la *proporción* de veces que se obtiene el resultado 0 en n repeticiones consecutivas se mantiene: $1 - p$

$$\frac{(1 - \delta_1) + \dots + (1 - \delta_n)}{n} \approx 1 - p, \quad \frac{(1 - \delta_2) + \dots + (1 - \delta_{n+1})}{n} \approx 1 - p, \quad \dots$$

$$\dots, \quad \frac{(1 - \delta_{k+1}) + \dots + (1 - \delta_{k+n})}{n} \approx 1 - p$$

o dicho directamente; el número de ejecuciones que dan el resultado 0 en n ejecuciones consecutivas de \mathcal{E} , es

$$(1 - \delta_1) + \dots + (1 - \delta_n) \approx n(1 - p), \quad (1 - \delta_2) + \dots + (1 - \delta_{n+1}) \approx n(1 - p), \quad \dots$$

$$\dots, \quad (1 - \delta_{k+1}) + \dots + (1 - \delta_{k+n}) \approx n(1 - p)$$

Nota: ¿Cuándo el lector ha visto?, que de, digamos simplemente en, 10 volados seguidos (volados con las condiciones del merengero) se obtengan 10 aguilas o 10 soles (como las mencionadas, ejemplificadas arriba) por medio de esa medida de qué tan frecuente se obtiene el resultado 1 en una sucesión numerosa de ejecuciones de \mathcal{E} , *pero* en relación a una sólo ejecución de \mathcal{E} ; esto

es, se dira que p es la *la medida de probabilidad de que en una ejecución de \mathcal{E} se obtenga el resultado 1* y, por lo tanto, $1 - p$ es la *medida de probabilidad de que en ésa se obtenga el resultado 0*.

Se insiste, esto se dirá, pensará, interpretará, solamente en el contexto de un experimento de Bernoulli $(\mathcal{E}, X_1)_p$. Esto no puede afirmarse y menos decirse en relación a cualquier pareja experimento-espacio de resultados $\{0, 1\}$; incluso a las parejas (\mathcal{E}, X_1) con las características $C1$ y $C2$; abundan las de este tipo cuyos colectivos no tienen el síndrome de Bernoulli. O de aquellas cuyos colectivos (por ejemplo para índices de valores en la Bolsa de Valores) no se comportan, ni hay testimonios, ni medidas, de que sea autosimilar, en el síndrome por lo menos de la ley de los números grandes de Bernoulli, de que grandes tramos tengan el síndrome de Bernoulli y menos que se cumpla una ley de los números grandes. Al afirmar que la medida de probabilidad de cierto evento es p al ejecutar \mathcal{E} .

Afirmar que cierto evento como producto de una ejecución de cierto experimento u observación tiene una medida de probabilidad, p , debe sustentarse en el hecho básico que para *otra* ejecución u observación de tal experimento, forzosamente esa medida de probabilidad sea la misma p . Y de una gran cantidad de ejecuciones de \mathcal{E} puesto que la ciencia, las teorías físicas o matemáticas no se hacen para un solo suceso de la naturaleza que no se vuelva a repetir, sino para una gran cantidad. Además, si se afirma que p es la medida de probabilidad de cierto evento, debe sustentarse en bases experimentales o teoricas, y sobre todo que tal acierto deba corroborarse experimentalmente . . . ¿y como se corrobora experimentalmente? Pues efectuando un gran número de observaciones de los resultados de ese experimento y corroborar eventual-

mente, si ese número es la medida de probabilidad de tal evento. Aún más, debe precisarse, ese término de probabilidad o al menos el de medida de probabilidad pero en el mundo experimental ya que en el teórico no es más que una medida normal definida en un álgebra.

Si se insiste en hablar, decir, que la medida de probabilidad de cierto evento es el número p , es por que tal afirmación puede comprobarse de alguna manera, ya sea teórico o experimentalmente y que esa probabilidad es siempre la misma. Pero si usualmente el término probabilidad se refiere a un hecho aislado, solitario, así suceda (o así se le ve), no hay forma de comprobar que su medida de probabilidad es cierto número. Y esto por principio no puede admitirlo la ciencia; la ciencia no trata un hecho que suceda una sola vez.

Estas son las razones por las cuales aquí se sustenta que se hable de *medida de probabilidad de que en una sola ejecución de \mathcal{E} , resulte, salga, el resultado 1*, que se enmarque en un experimento de Bernoulli, $(\mathcal{E}, X_1)_p$. Y es que tiene el síndrome de Bernoulli y en particular con su ley de los números grandes

$$|\delta_1 + \dots + \delta_n - np| \ll n, \quad \forall n \gg 1$$

en el cual puede comprobarse, convencerse ese dicho: *la medida de probabilidad de que al ejecutarse \mathcal{E} resulte, salga, el resultado 1 es p* .

Incluso, por ejemplo, en las encuestas que se reducen al hecho de extraer al azar bolas negras y de colores, de un recipiente. No puede decirse que cierto candidato tiene la probabilidad p de ganar, por ejemplo, cierto puesto. Basta que se hagan encuestas con una semana de diferencia, que un candidato haga cierta declaración pública, para que en otra encuesta obtenga un número p' , $p \neq p'$. Así, lo que aquí se pide, para hablar de medida de probabilidad es que

se cumpla la condición $C1$, situación que en los fenómenos sociales cambia muy rápidamente. Y en nuestros experimentos de Bernoulli se conserva, es lo que pedimos; y estos existen, por ejemplo en la física donde las leyes evolucionan muy pero muy lentamente. O en los juegos de azar.

Así, debe garantizarse de alguna manera que esa probabilidad, ese número sea siempre el mismo, y para que esto sea verificable, medible, repetible corroborable, debe repetirse las cosas, en las mismas condiciones, en lo posible (como puede hacerse en un laboratorio de física, por ejemplo), convincentemente.

Por otro lado, en el mundo de la teoría matemática es inconcebible mezclar la estructura matemática con los términos como experimento, realización, frecuencia, etc., términos conceptos, imposibles e inútiles de matematizar. Sin embargo, el aspecto teórico puede, y debe, tener presente una pareja $(\mathcal{E}, X_1)_p$ pensada (*un experimento de Bernoulli pensado*) que represente en el aspecto experimental esa estructura matemática; es conveniente tener siempre una imagen física, geométrica, de la estructura matemática.

6. Ley basta del logaritmo iterado

La *ley basta del logaritmo iterado* de un experimento de Bernoulli $(\mathcal{E}, X_1)_p$ es una refinación de la ley de los números grandes de Bernoulli; es una ley que comprende y afina a esta última pero manteniendo su esencia.

Agregando las gráficas de las funciones

$$g_1(x) = \sqrt{\frac{x}{2} \ln \ln x}, \quad g_2(x) = -\sqrt{\frac{x}{2} \ln \ln x},$$

$x \geq e$ (e base de los logaritmos naturales) a la gráfica de $y = px$, $x \geq 0$, se

obtiene la Figura 4.

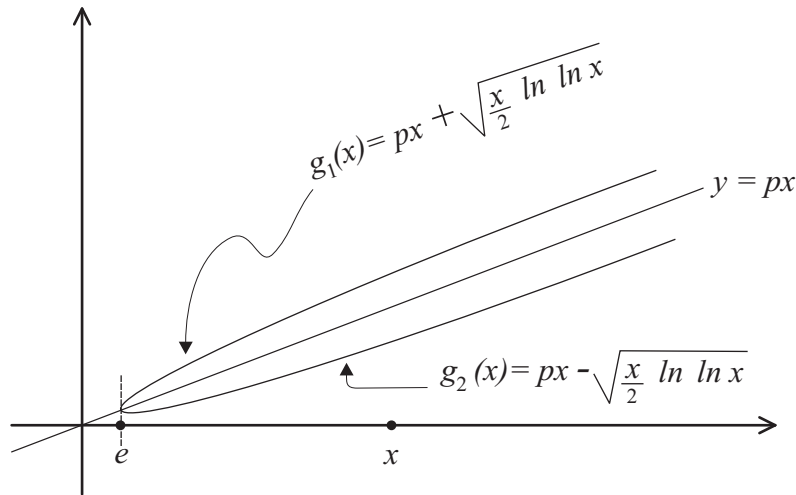


Figura 4: Gráficas de las funciones $g_1(x)$ y $g_2(x)$.

La región de \mathbb{R}^2 comprendida entre las dos nuevas curvas obtenidas

$$K_p = \{(x, y) : px - \sqrt{\frac{x}{2} \ln \ln x} < y < px + \sqrt{\frac{x}{2} \ln \ln x}\}$$

será llamada *zona de Khinchin de parámetro p*. Como puede verse esta zona es simétrica respecto a la recta $y = px$ y se va ensanchando lenta pero muy lentamente conforme crece x .

Por ejemplo, su *ancho vertical* para

i) $x = 10^3$, es $2\sqrt{\frac{10^3}{2} \ln \ln 10^3} = 62,17\dots$; $16,08\dots \times 62,17\dots = 1000$

ii) $x = 10^4$, es $2\sqrt{\frac{10^4}{2} \ln \ln 10^4} = 210,72\dots$; $47,45\dots \times 210,72\dots = 1000$

iii) $x = 10^5$, es $2\sqrt{\frac{10^5}{2} \ln \ln 10^5} = 699,06\dots$; $143,04\dots \times 699,06\dots = 1000$

lo que va mostrando lo lentamente que crece el ancho vertical, $2\sqrt{\frac{x}{2} \ln \ln x}$, la pequeñez de la pendiente entre dos valores alejados x_1 y x_2 de la curva $g(x) = \sqrt{\frac{x}{2} \ln \ln x}$ y cuántas veces x es mayor que $2\sqrt{\frac{x}{2} \ln \ln x}$.

Las relaciones entre x y $\sqrt{\frac{x}{2} \ln \ln x}$ pueden verse en sus gráficas (Fig.5),

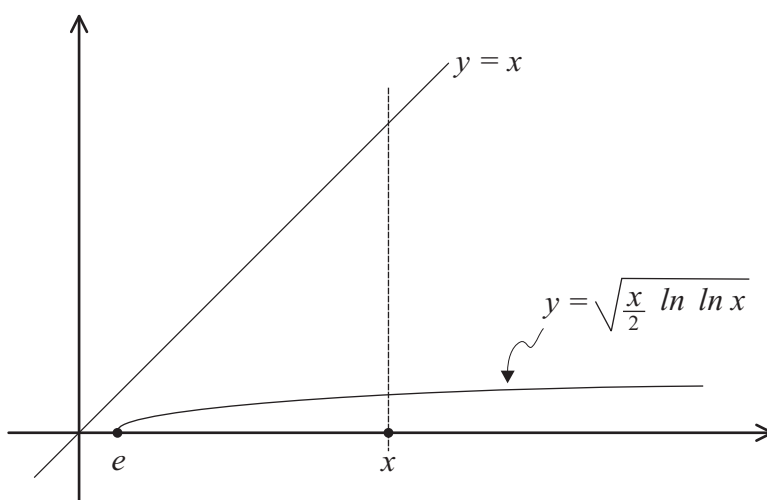


Figura 5: Gráficas de x y $\sqrt{\frac{x}{2} \ln \ln x}$.

y sentirse con esos ejemplos numéricos; aunque

$$\sqrt{\frac{x}{2} \ln \ln x} \xrightarrow{x \rightarrow \infty} +\infty,$$

la función es creciente, pero crece tan lentamente que en tramos grandes, un poco distantes de $x = e$, parece una recta horizontal.

Lo que se tiene con estas dos funciones es que

$$\sqrt{\frac{x}{2} \ln \ln x} \ll x,$$

hecho que proviene de lo más básico y sabido, que

$$\ln \ln x \ll x;$$

viendolo con algunos valores concretos de x :

$$x = e, \ln \ln e = 0 \ll 2e;$$

$$x = 10, \ln \ln 10 = 0,834 \dots \ll 20;$$

$$x = 100, \ln \ln 100 = 1,5227 \dots \ll 200;$$

$$x = 10^3, \ln \ln 10^3 = 1,932 \dots \ll 2000;$$

$$x = 10^6, \ln \ln 10^6 = 2,625 \dots \ll 2 \times 10^6;$$

etc.; con qué lentitud tan desesperante crece $y = \ln \ln x$; de la impresión de ser una semirecta que parte de $x = e$ y tiene una pendiente pequeñísima: por ejemplo de $x = e$ a $x = 10^4$ apenas ha crecido $2,625 \dots$ unidades en un intervalo de un millón de unidades, practicamente. En fin, lo que se necesita destacar entre las funciones es la expresión de arriba,

$$\sqrt{\frac{x}{2} \ln \ln x} \ll x, \forall x \geq e$$

o sea que

$$\frac{\sqrt{\frac{x}{2} \ln \ln x}}{x} \ll 1, \forall x \geq e$$

o equivalentemente

$$1 \ll \frac{x}{\sqrt{\frac{x}{2} \ln \ln x}}, \forall x \geq e.$$

Como se ve hay una gran semejanza geométrica entre la zona de Khinchin K_p y el cometa conformado por las gráficas de los colectivos de $(\mathcal{E}, X_1)_p$

así que no es de extrañar que se establezca la *Ley basta del logaritmo iterado*:
cada colectivo de resultados de un experimento de Bernoulli $(\mathcal{E}, X_1)_p$,

$$\begin{array}{l} \text{Ejecución de } \mathcal{E} \quad 1, \quad 2, \quad \dots, \quad n, \quad \dots \\ \text{Resultado} \quad \delta_1, \quad \delta_2, \quad \dots, \quad \delta_n, \quad \dots \end{array}$$

tiene la propiedad de que

$$|\delta_1 + \dots + \delta_n - np| < \sqrt{\frac{n}{2} \ln \ln n}, \quad \forall n \gg 1,$$

o, equivalentemente,

$$\left| \frac{\delta_1 + \dots + \delta_n}{n} - p \right| < \sqrt{\frac{\ln \ln n}{2n}}, \quad \forall n \gg 1.$$

Geoméricamente, esta ley basta del logaritmo iterado está afirmando que la gráfica de cada colectivo en alguna ejecución de \mathcal{E} entra en la zona de Khinchin K_p y permanece adentro para ya no volver a salir en las ejecuciones posteriores de \mathcal{E} . O que a partir de cierta ejecución k , $\delta_1 + \dots + \delta_k$, el cual es un número aleatorio, difiere del número determinístico kp , de la recta $y = px$, en menos que $\sqrt{\frac{k}{2} \ln \ln k}$, y que se mantiene esto para cada $n \geq k$ (Fig.6),

$$|\delta_1 + \dots + \delta_n - np| < \sqrt{\frac{n}{2} \ln \ln n}.$$

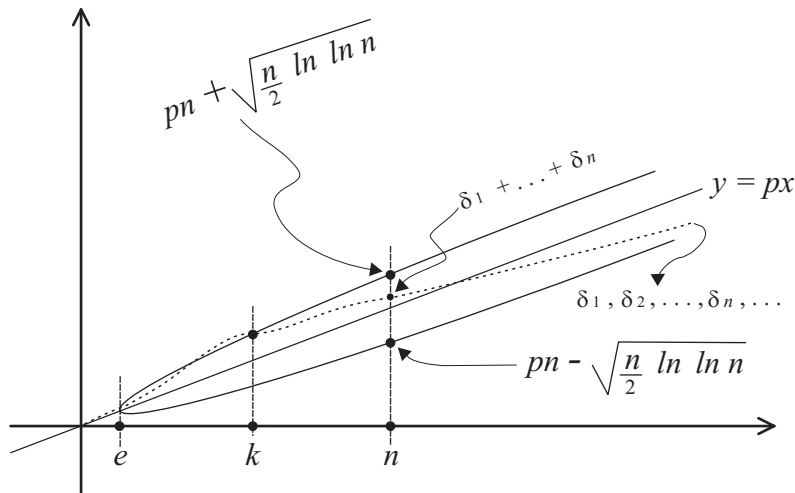


Figura 6: Gráfica de cada colectivo en alguna ejecución de \mathcal{E} .

Como puede verse, esta ley basta del logaritmo iterado, comprende y refina la tan entrañable ley de los números grandes de Bernoulli; en efecto, ya que aquélla afirma que el valor absoluto

$$|\delta_1 + \dots + \delta_n - np|$$

es menor que

$$\sqrt{\frac{n}{2} \ln \ln n},$$

para toda n grande, como se vio anteriormente, y este último número es mucho menor que su correspondiente n , que es lo que afirma la ley de Bernoulli. Así pues, se tiene una ley más fina acerca de ese comportamiento cuantitativo del síndrome de Bernoulli; es más, de esta ley del logaritmo iterado se desprende la importante afirmación que

$$\frac{\delta_1 + \dots + \delta_n}{n} \xrightarrow{n \rightarrow \infty} p,$$

en vista de que

$$\left| \frac{\delta_1 + \dots + \delta_n}{n} - p \right| < \sqrt{\frac{\ln \ln n}{2n}} \xrightarrow{n \rightarrow \infty} 0.$$

Así pues, es en este sentido que esta ley basta refina y comprende la ley de los números grandes de Bernoulli: no sólo afirma que aquellas distancias son mucho menores que n , sino que las *cuantifica* por medio de una expresión analítica sencilla y fácil de evaluar para los diversos valores de n .

Entonces, hay un *comportamiento común* a todos los colectivos generados por las *mismas causa*, el síndrome de Bernoulli; y como una faceta cuantitativa de éste, la ley basta del logaritmo iterado que, en cierto aspecto es *equivalente* a la ley de los números grandes de Bernoulli, aunque, cabe agregar, con aquella se tiene una expresión que permite calcular la *velocidad*, por decirlo así, *de aproximación* de $\delta_1 + \dots + \delta_n$ a np , o la velocidad de crecimiento de su distancia, o equivalentemente, la *velocidad* de convergencia a p de

$$\frac{\delta_1 + \dots + \delta_n}{n}.$$

Cabe, desde luego, una explicación en cuanto a la ley basta expuesta aquí; esta ley es una adaptación de una parte de un teorema finísimo acerca de los números reales, descubierto por el gran matemático soviético A.I. Khinchin, en 1924, acerca del comportamiento de la distribución de los 0s y 1s en los números reales, expresados en binario, del intervalo $[0, 1]$ en el soporte de la medida normal o medida de probabilidad μ_p teorema que a su vez refina el teorema de Borel, el cual en su parte esencial dice que *para casi todo número real, respecto a μ_p , $\delta_1 \delta_2 \dots \delta_n \dots$, se tiene que*

$$\frac{\delta_1 + \dots + \delta_n}{n} \xrightarrow{n \rightarrow \infty} p.$$

El teorema de Khinchin dice más acerca de estos números, describe con una fineza exhaustiva el comportamiento de las sucesiones, correspondientes a esos números,

$$\delta_1, \delta_2, \dots, \delta_n, \dots,$$

en cuanto cómo se van reflejando estos colectivos en las zonas de Khinchin, más detalladas y menos rígidas de la que aquí se presenta. Sin embargo no todos los números reales en binario que obedecen el teorema de Khinchin son colectivos de Bernoulli, colectivos de $(\mathcal{E}, X_1)_p$, ya que puede tenerse que el real

$$\underbrace{0 \dots 0}_{10^6} \delta_{10^6+1} \delta_{10^6+2} \dots$$

sea de Khinchin, o que

$$\frac{0 + \dots + 0 + \delta_{10^6+1} \delta_{10^6+2} + \dots + \delta_n}{n} \xrightarrow[n \rightarrow \infty]{} p$$

pero el colectivo

$$0, 0, \dots, 0, \delta_{10^6+1}, \delta_{10^6+2}, \dots$$

no es colectivo de Bernoulli (por ejemplo

$$|\underbrace{0 + \dots + 0}_{10^3} - 10^3 \cdot p| < \sqrt{\frac{10^3}{2} \ln \ln 10^3}$$

no se cumple, o lo mismo con la proporción de 1s en los 10^3 primeros resultados de ejecutar \mathcal{E}). Nunca se observa u observará tal colectivo.

Sin embargo si

$$\delta_1, \delta_2, \dots, \delta_n, \dots$$

es un colectivo de $(\mathcal{E}, X_1)_p$, el número real $0.\delta_1\delta_2 \dots \delta_n \dots$ satisfará las condiciones en el teorema de Khinchin. Esto es, los números reales *generados* por corridas de repeticiones sucesivas de \mathcal{E} son de Khinchin.

Las zonas de Khinchin en su teorema, además de depender de otros parámetros, depende de $(1 - p)$, dentro del radical, lo cual no se ha incluido en la presentada aquí, sino se ha sustituido por una cota superior, $1/4$; en aras de la sencillez, de la corroboración directa [Arnaldo].

Así pues, aquí no se ha tomado en cuenta la posición de p en el intervalo abierto $(0, 1)$, ya sea cercano a 0 ó a 1 o cercano a $1/2$, lo que redundaría en que, para los colectivos de Bernoulli la zona de Khinchin K_p es muy sobrada si p es cercano a 0 ó a 1, pero un poco justa si p está cerca de $1/2$. De cualquier manera los colectivos que se obtiene quedan siempre dentro de la zona de Khinchin para toda n grande. Véase, por ejemplo, los experimentos observados por E.Arnaldo [Tesis].

En cuanto a los aspectos aleatorios en relación con la ley basta del logaritmo iterado, éstos se manifiestan claramente cuando, como antes, se ejecutan multitud de corridas de repeticiones sucesivas de \mathcal{E} ; así, si éstos son

$$\begin{array}{ccccccc} \delta_1^1 & \delta_2^1, & \dots & \delta_n^1, & \dots & & \\ & \delta_1^2 & \delta_2^2, & \dots & \delta_n^2, & \dots & \\ & & & \vdots & & & \\ & \delta_1^N & \delta_2^N, & \dots & \delta_n^N, & \dots & \\ & & & \vdots & & & \end{array}$$

y si bien para cada colectivo i se tiene, determinísticamente que

$$|\delta_1^i + \dots + \delta_n^i - np| < \sqrt{\frac{n}{2} \ln \ln n} \quad \forall n \gg 1,$$

sin embargo

- i*. la ejecución $n_0(i)$ a partir de la cual la gráfica del colectivo i entra definitivamente a la zona de Khinchin K_p no es siempre la misma para

$i = 1, 2, \dots, N \dots$; esa entrada es aleatoria respecto a la misma causa, aunque estos números generalmente son pequeños. Y esta es, pudiérase decir, *lo razonable* de la afirmación de la ley basta, al no especificar, y pueda especificarse con base en las especificaciones generales del experimento de Bernoulli $(\mathcal{E}, X_1)_p$, cuál es la ejecución $n_0(i)$ de entrada definitiva, ya que sólo afirma que sucederá para toda n grande; viéndolo bien $n_0(i)$ depende de la *cola* del colectivo, algo así como

$$n_0(i) = n_0(\delta_k^i, \delta_{k+1}^i, \dots)$$

para el cual

$$|\delta_i^i + \dots + \delta_n^i - np| < \sqrt{\frac{n}{2} \ln \ln n} \quad \forall n \geq n_0(i)$$

¡el tiempo de entrada definitiva de un colectivo depende de su futuro!
lo cual por lo menos en el plano real es absurdo.

Así que con base en $(\mathcal{E}, X_1)_p$ y la corrida de repeticiones sucesivas de \mathcal{E} , no se determina esa n_0 . Aunque puede plantear la pregunta si esa n_0 depende de algo más, si no hay por ahí una variable oculta, una componente que aún no se ha descubierto y que no se haya tomado en cuenta.

Aunque esto negaría la existencia real de la aleatoriedad en la naturaleza y retornaría las cosas, la visión de la naturaleza al principio determinístico Laplaciano visión superada por innumerables testimonios y razonamientos.

No obstante esos naturales aleatorios $n_0(1), n_0(2), \dots, n_0(N), \dots$ se distribuyen de la misma forma, siguiendo el *mismo patrón para cada*

serie de corridas de repeticiones sucesivas de \mathcal{E} que se efectúe. Y este hecho determinístico es la regularidad estadística del comportamiento de esos números naturales aleatorios.

ii. Otra regularidad estadística es la ya mencionada acerca de la distribución, siempre la misma para cada n , de los números

$$\delta_1^1 + \cdots + \delta_n^1, \quad \delta_1^2 + \cdots + \delta_n^2, \quad \delta_1^N + \cdots + \delta_n^N, \quad \dots$$

en el intervalo $[0, n]$.

Y si n es grande esa distribución se acumula alrededor de los naturales cercanos a np , como lo manda la ley de los números grandes de Bernoulli, aunque, ahora se lo tiene en un sentido más preciso, con la ley basta del logaritmo iterado

$$|\delta_1^i + \cdots + \delta_n^i - np| < \sqrt{\frac{n}{2} \ln \ln n},$$

aunque esa *precisión* aportada por la ley basta causa un ligero inconveniente, el cual era de esperarse aunque se reduce al aumentar n ; el inconveniente consiste en que no todos esos número aleatorios cumplen la desigualdad, esto es que habrá algunos $\delta_1^j + \cdots + \delta_n^j$ tales que

$$|\delta_1^j + \cdots + \delta_n^j - np| \geq \sqrt{\frac{n}{2} \ln \ln n},$$

geométricamente, que para la ejecución n , parte del colectivo j no está en la zona de Khinchin. O viéndolo en el intervalo $[0, n]$, $\delta_1^j + \cdots + \delta_n^j$ cae fuera del intervalo abierto

$$\left(np - \sqrt{\frac{n}{2} \ln \ln n} \ , \ np + \sqrt{\frac{n}{2} \ln \ln n} \right)$$

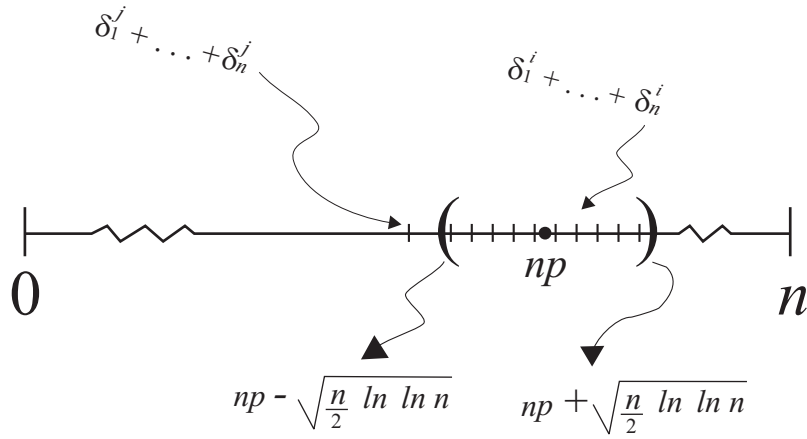


Figura 7: Intervalo abierto $(np - \sqrt{\frac{n}{2} \ln \ln n}, np + \sqrt{\frac{n}{2} \ln \ln n})$.

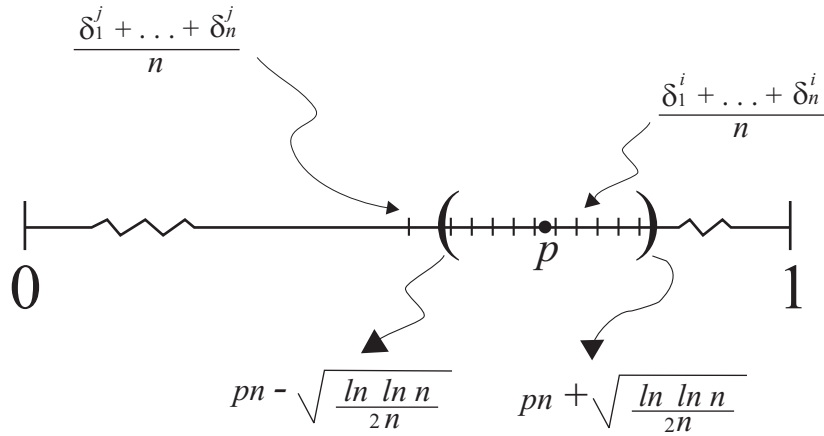


Figura 8: Intervalo abierto $(p - \sqrt{\frac{\ln \ln n}{2n}}, p + \sqrt{\frac{\ln \ln n}{2n}})$.

contenido en $[0, n]$, Figura 7. O equivalentemente, lo que dice la Figura 8.

Sin embargo, por la ley de los números grandes de Bernoulli, se tiene que

$$|\delta_1^j + \dots + \delta_n^j - np| \ll n,$$

aunque $\delta_1^j + \dots + \delta_n^j$ esté fuera de aquella vecindad.

Por ejemplo, si $p = 2/3$ y $n = 1000$ repeticiones sucesivas de \mathcal{E} se obtendrá

a) Por la ley de los números grandes de Bernoulli,

$$\left| \delta_1 + \dots + \delta_{1000} - 1000 \times \frac{2}{3} \right| \ll 1000$$

y

b) *Seguramente, muy probablemente*, por la ley basta del logaritmo iterado, que

$$1000 \times \frac{2}{3} - \sqrt{500 \ln \ln 1000} < \delta_1 + \dots + \delta_{1000} < 1000 \times \frac{2}{3} + \sqrt{500 \ln \ln 1000}$$

O, equivalentemente

$$-31,08 \dots = -\sqrt{500 \ln \ln 1000} < \delta_1 + \dots + \delta_{1000} - 1000 \times \frac{2}{3} < \sqrt{500 \ln \ln 1000} = 31,08 \dots$$

o que

$$636 \leq \delta_1 + \dots + \delta_{1000} \leq 697$$

o

$$0,636 \leq \frac{\delta_1 + \dots + \delta_{1000}}{1000} \leq 0,697$$

Entonces, *es muy probable* que en $n = 1000$ ejecuciones de \mathcal{E} se obtenga *un número de 1s* que está en el intervalo $[636, 697]$.

La *seguridad* de que $\delta_1 + \dots + \delta_n$ caiga en el intervalo $[636, 697]$ no es absoluta, supóngase que el resultado $\delta_1 + \dots + \delta_{1000}$ caiga en el número 634, entonces se tiene que

$$\left| \delta_1 + \dots + \delta_{1000} - 1000 \times \frac{2}{3} \right| \geq \sqrt{500 \ln \ln 1000},$$

sin embargo

$$|634 - 666,66\dots| = 32,66\dots$$

que, como se ve, $32,66\dots \ll 1000$, como lo afirma la ley de los números grandes.

Volviendo al “inconveniente” mencionado arriba, se tiene otra regularidad estadística: si bien no todos los números aleatorios obtenidos hasta la ejecución n , n grande, en esas tandas de repeticiones de \mathcal{E} ,

$$\delta_1^1 + \dots + \delta_n^1, \quad \delta_1^2 + \dots + \delta_n^2, \quad \delta_1^N + \dots + \delta_n^N, \quad \dots$$

caen en el subintervalo de $[0, n]$ dictado por la ley basta del logaritmo iterado, sí sucede con la gran mayoría cuando N , el número de esas corridas, es grande. Y por tal razón se dice que *es muy probable*, que en una sola corrida de n ejecuciones sucesivas de \mathcal{E} el resultado aleatorio $\delta_1 + \dots + \delta_n$ caiga en el intervalo

$$\left(np - \sqrt{\frac{n}{2} \ln \ln n} \quad , \quad np + \sqrt{\frac{n}{2} \ln \ln n} \right).$$

Es más, ese *muy probable* es cuantificable debido a que la proporción de el hecho que las sumas que caen en tal intervalo, respecto a N , *siempre, determinísticamente*, es un número cercano a 1 cuando N es grande, y entre más grande sea n , tal cercanía a 1 será mayor. Y esto no es de extrañar, ya que con esto se tiene otro experimento de Bernoulli: el experimento consiste en efectuar n repeticiones sucesivas de \mathcal{E} y los resultados son 1, sí $\delta_1 + \dots + \delta_n$ cae dentro del intervalo $(np - \sqrt{\frac{n}{2} \ln \ln n} \quad , \quad np + \sqrt{\frac{n}{2} \ln \ln n})$ y 0 si $\delta_1 + \dots + \delta_n$ cae fuera.

En alguno de los capítulos se expondrá tal cuantificación (aunque para ser sinceros, esa cuantificación teórica, da un número menor que el observado

experimentalmente; puede verse la tesis profesional de Arnaldo.)

Para concluir este capítulo se propone un ejercicio muy ilustrativo de todo lo expuesto en este capítulo, y es el siguiente:

1. Obténgase cuatro dígitos al azar x_1, x_2, y_1, y_2 .
2. Fórmese el punto, de \mathbb{R}^2 , $(\bullet x_1 x_2, \bullet y_1 y_2)$, en $[0, 1] \times [0, 1]$.
3. Si $(\bullet x_1 x_2, \bullet y_1 y_2)$ cae en el cuadrante del disco dado por $x^2 + y^2 \leq 1$, $A = \{(x, y) : x^2 + y^2 \leq 1, 0 \leq x \leq 1, 0 \leq y \leq 1\}$ (Fig.9), póngase 1 y si

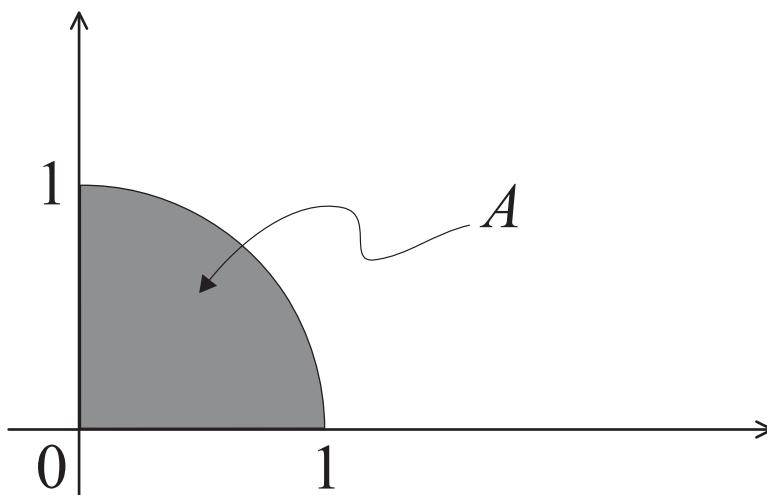


Figura 9: $A = \{(x, y) : x^2 + y^2 \leq 1, 0 \leq x \leq 1, 0 \leq y \leq 1\}$.

no, póngase 0.

4. Entonces se tiene el experimento \mathcal{E} , que consiste en tomar cuatro dígitos al azar y formar con ellos un punto de $[0, 1] \times [0, 1]$; los resultados de

efectuar \mathcal{E} que interesan son 1, si el punto obtenido cae en A , o si el punto obtenido no cae en A .

5. Repítase \mathcal{E} 1000 veces y regístrese en una tabla los resultados obtenidos

<i>Ejecución de \mathcal{E}</i>	1, 2, ..., 1000
<i>Resultado</i>	$\delta_1, \delta_2, \dots, \delta_{1000}$;

- i) gráfíquese el colectivo obtenido en una hoja de papel milimétrico
- ii) gráfíquese en ésta la recta $y = \frac{\pi}{4}x$ y la zona de Khinchin correspondiente
- iii) verifíquese que la gráfica del colectivo tiene el síndrome de Bernoulli y que en particular para alguna ejecución de \mathcal{E} el colectivo se meta en la zona de Khinchin $K_{\frac{\pi}{4}}$. Esto es, $(\mathcal{E}, X_1 = \{0, 1\})$ es un experimento de Bernoulli

Nota. Un procedimiento para obtener esos cuatro dígitos al azar es tomando dos columnas de números telefónicos de un directorio, tómesese como x_1^1 y x_2^1 los dos últimos dígitos del renglón de la primera columna y como y_1^1 y y_2^1 los dos últimos dígitos del número telefonico que aparece en el primer renglón de la segunda columna. Y así sucesivamente, así hasta obtener la cuarteta $x_1^{1000}, x_2^{1000}, y_1^{1000}, y_2^{1000}$ y con todos éstos formar los puntos

$$(\bullet x_1^1 x_2^1, \bullet y_1^1 y_2^1), \dots, (\bullet x_1^{1000} x_2^{1000}, \bullet y_1^{1000} y_2^{1000})$$

y éstos son puntos que se distribuyen al azar por el cuadro $[0, 1] \times [0, 1]$; bueno, mas bien se distribuyen al azar por los puntos, de este cuadro, de la forma $(\bullet d_1 d_2, \bullet d_3 d_4)$, con

$$d_i = 0, 1, \dots, 9 \quad , \quad i = 1, 2, 3, 4.$$

Capítulo 2

Distribuciones iguales estadísticamente

2.1. Colectivos de resultados aleatorios

Por \mathcal{E} se denotará un experimento u observación que se *repite*, o puede repetirse a voluntad de un experimentador, *en las mismas*, o *casi las mismas*, condiciones experimentales o materiales, un gran número de veces y esto en un gran número de ocasiones; tal tipo de ejecuciones de \mathcal{E} son llamadas *independientes*. Ya sea que \mathcal{E} sea una observación o un experimento, aquí se le llamará *experimento*.

De las ejecuciones (que siempre serán independientes) de \mathcal{E} , interesa cierta categoría de resultados, los cuales son representados por elementos de un conjunto X : cuando se ejecuta \mathcal{E} , se obtiene un resultado de esta categoría, el cual está representado por cierto elemento x de X y, entonces, se dirá que el *resultado* de esta ejecución de \mathcal{E} fue x . Interesarán aquellas categorías de al menos dos resultados, razón por la cual sólo se considerarán conjuntos X con al menos dos elementos resultados, esto es con $\#(X) \geq 2$. Una vez especificados la categoría o tipo de resultados que interesa observar al efectuarse \mathcal{E} y sus representaciones por elementos del conjunto X , se llamará a éste *espacio de resultados* de \mathcal{E} . En general, por restricciones de carácter matemático (aunque a fin de cuentas resulta cómodo y no afecta mayormente los objetivos que se persiguen) el espacio de resultados suele ser “sobrado” en relación a la categoría de resultados que interesan, en el sentido de que cada resultado de \mathcal{E} está en X aunque cada elemento en X no es, necesariamente, un resultado de \mathcal{E} . Lo ideal sería que los elementos de X y los resultados de la categoría de interés estuvieran en una relación biunívoca, sin embargo esto no es posible debido al carácter de las mediciones que pueden hacerse de las observaciones, por ejemplo, o de complicaciones innecesarias de carácter matemático, etc.; no obstante, cuando la categoría de resultados es poco numerosa tal correspondencia es posible.

Ahora bien, interesarán aquellas parejas *experimento* (u *observación*) \mathcal{E} -

espacio de resultados X , con las características

- C1. \mathcal{E} se repite o puede repetirse un gran número de veces y esto en un gran número de ocasiones.
- C2. Las condiciones en que se efectúa \mathcal{E} no determinan su resultado (en X): los resultados de \mathcal{E} son aleatorios.

(Esta última característica es una versión del principio clásico del indeterminismo: a *iguales causas*, ejecución de \mathcal{E} , *diferentes efectos*, resultado en X).

En lo que sigue, este tipo de parejas se denotará por el arreglo de símbolos (\mathcal{E}, X) .

Lo que interesará siempre de estas parejas será en primer lugar, sus *colectivos de resultados* correspondientes a las diferentes corridas de ejecuciones sucesivas y numerosas de \mathcal{E} ; al efectuar una corrida de ejecuciones sucesivas de \mathcal{E} se registran los resultados correspondientes (en X):

<i>Ejecución de \mathcal{E}</i>	1,	2,	...	i,	...
<i>Resultado</i>	$x_1,$	$x_2,$...	$x_i,$...

la sucesión de resultados $x_1, x_2, \dots, x_n, \dots$ es el *colectivo* (o *serie de tiempo*) obtenido en esa sucesión de ejecuciones de \mathcal{E} : el resultado de x_1 , en X , se obtuvo en la ejecución 1; el resultado x_2 , en X , se obtuvo en la ejecución 2 de \mathcal{E} ; ... ; el resultado de x_n , en X , se obtuvo en la ejecución de n de \mathcal{E} ; ... Este colectivo es una sucesión de elementos de X *suministrados aleatoriamente*, pero no por cualquier tipo de aleatoriedad sino por la peculiar de la pareja (\mathcal{E}, X) .

Otra peculiaridad de éste, y de cada colectivo que se obtenga, es que no hay una relación de causa efecto en la ejecución i de \mathcal{E} y el resultado correspondiente x , ya lo dice C2. Aunque puede ser confusa la expresión, algunas veces se dirá que el colectivo $x_1, x_2, \dots, x_n, \dots$ es una *sucesión aleatoria* de elementos de X , pero teniendo en cuenta que lo es en el contexto (\mathcal{E}, X) .

2.2. Descripción cuantitativa del colectivo

Al graficar un gran número de resultados sucesivos de un colectivo de una pareja (\mathcal{E}, X) en una representación geométrica de X , se observa un *conglomerado* o *constelación* de puntos distribuidos por X , una imagen geométrica de la distribución de esos resultados por X . Así, puede verse, que en tal región la concentración de puntos, resultados, es más *intensa* que en tal otra región, y que en aquella otra zona los puntos, los resultados, son *escasos*, como si estos evitaran *caer* en esta zona y *preferieran* aquélla, por decirlo así; y en este tenor puede seguirse describiendo la forma del conglomerado, más o menos detalladamente: en términos de mayor, menor, intermedia o nula concentración de puntos por diversas regiones de X .

Esta manera de describir el conglomerado de puntos, la distribución de resultados, se está haciendo en términos numéricos: cuantificando la mayor, intermedia o menor concentración de puntos en diversas regiones, asignando a éstas la *intensidad*, *frecuencia* o *medida de concentración* de puntos o resultados.

Una técnica más apropiada de esa medición, será utilizando la aritmética: una vez graficados los resultados x_1, x_2, \dots, x_n , para n grande, del colectivo en cuestión, se eligen diversas regiones de X , genéricamente denotadas por R ; para cada R se obtiene el promedio aritmético, la proporción respecto a n del número de resultados que *cayeron* en R , lo cual da una *medida* de la *intensidad*, *grado de concentración* o proporción de resultados en R .

Utilizar esta técnica tan simple, desvelará aspectos sutiles e inimaginables acerca de los diferentes colectivos de (\mathcal{E}, X) , y de sus relaciones, que no son patentes de las figuras geométricas de los conglomerados de puntos o distribuciones de sus resultados, por un lado, y por otro ni se imagina uno la magna teoría matemática en que desembocará esa idea tan natural y elemental de medir esos grados de concentración.

Para una primera formulación práctica de esta técnica se considerará:

- i) Una región R de X , ni “muy grande” ni “muy pequeña”, práctica y fácilmente descriptible, y que permita considerar o construir otra u otras de su tipo;
- ii) la función característica de R ,

$$\delta_R : X \rightarrow \{0, 1\} \ni$$

$$\delta_R(x) = \begin{cases} 1 & \text{si } x \in R \\ 0 & \text{si } x \in R^c, \end{cases}$$

con objeto de registrar si al ejecutarse \mathcal{E} su resultado x *cae* (está) o *no cae* (no está) en R :

- si x *cae* en R , como $x \in R$, entonces $\delta_R(x) = 1$,
- si x *no cae* en R , como $x \in R^c$, entonces $\delta_R(x) = 0$;

o, en otros términos, pero equivalentes:

al ejecutarse \mathcal{E}

R *se realiza*, si el resultado x *cae* en R , $\delta_R(x) = 1$,

R *no se realiza*, si el resultado x *cae* en R^c , $\delta_R(x) = 0$.

Ahora bien, la mayor, intermedia o menor concentración de puntos o resultados en una región R se aprecia por medio de una cuantificación, por lo general inconciente, que se puede expresar por la suma

$$\delta_R(x_1) + \dots + \delta_R(x_n) = \sum_{i=1}^n \delta_R(x_i),$$

el número de los n resultados x_1, x_2, \dots, x_n que *caen* en R , aunque es más realista comparar esta suma respecto a n , esto es, considerar el promedio aritmético

$$\frac{1}{n} \sum_{i=1}^n \delta_R(x_i) \quad ;$$

promedio aritmético de los resultados x_1, x_2, \dots, x_n que *cayeron* en R ; proporción de los resultados obtenidos en esas primeras n ejecuciones de $\mathcal{E}, x_1, x_2, \dots, x_n$

distribuidos en X que caen en la región R . Tal promedio aritmético es una *medida* de la frecuencia con que aquéllos cayeron en ésta o la *medida de concentración* de los puntos de la configuración geométrica (de la distribución de resultados) en esta región; o, como puede verse, también es una medida de *que tan frecuentemente* se realizó R en estas n ejecuciones sucesivas de \mathcal{E} .

En relación a la descripción geométrica del conglomerado de puntos x_1, x_2, \dots, x_n en X , ésta puede ser, y lo será, descrito por la colección de *medidas*

$$\frac{1}{n} \sum_{i=1}^n \delta_R(x_i)$$

para diferentes regiones de R de X , substituyendo así, o complementando, la descripción cualitativa por, o con, una *descripción mensurable*.

Pero teniendo en consideración el carácter aleatorio de los resultados de \mathcal{E} de una ejecución a otra, esas proporciones o medidas (de *qué tan frecuentemente cayeron* resultados, de las n primeras ejecuciones de \mathcal{E} , en las diferentes R ; o, lo que es equivalente, son *medidas* de *que tan frecuentemente* se realizó R) obtenidas en cualesquiera n ejecuciones sucesivas de \mathcal{E} , son también resultados, números, aleatorios.

Así pues, con estas proporciones se describirán las diferentes o iguales distribuciones de resultados de los colectivos de una pareja (\mathcal{E}, X) y, por lo tanto, sus colectivos, en primer lugar y después, los colectivos o, lo que es equivalente, las distribuciones de sus resultados, de interés fundamental para las teorías matemáticas de la probabilidad y de los procesos estocásticos, y que son generados sólo por cierto tipo de parejas (\mathcal{E}, X) .

2.3. Igualdad estadística de distribuciones

La descripción cuantitativa de un conglomerado de puntos o de una distribución de resultados, presentada en la sección anterior, será utilizada para detectar y finalmente caracterizar lo que se entenderá por *igualdad estadística* de conglomerados de puntos, de distribuciones de resultados o de colectivos.

Antes conviene aclarar que el término *estadística* o *estadístico* lo es en el sentido original y más generalizado; en primer lugar se refiere a una gran cantidad de datos (mediciones, resultados, números) empíricos y, en segundo lugar, que todos estos datos son de carácter aleatorio en relación a la fuente, experimento u observación, de la cual se obtengan.

Quizá sea más conveniente mostrar cuándo manifiestamente dos distribuciones de resultados son evidentemente diferentes; si de una cierta pareja (\mathcal{E}, X) los resultados de uno de sus colectivos

$$\begin{array}{ll} \text{Ejecución de } \mathcal{E} & 1, \quad 2, \quad \dots, \quad i, \quad \dots \\ \text{Resultado} & x_1, \quad x_2, \quad \dots, \quad x_i, \quad \dots; \end{array}$$

se distribuyen por X en la forma como se ilustra en a) de la figura 2.1 y los del otro

Ejecución de \mathcal{E}	1, 2, ..., i , ...
Resultado	$y_1, y_2, \dots, y_i, \dots$;

se distribuyen por X en la forma como se ilustra en b) de la figura 2.1,

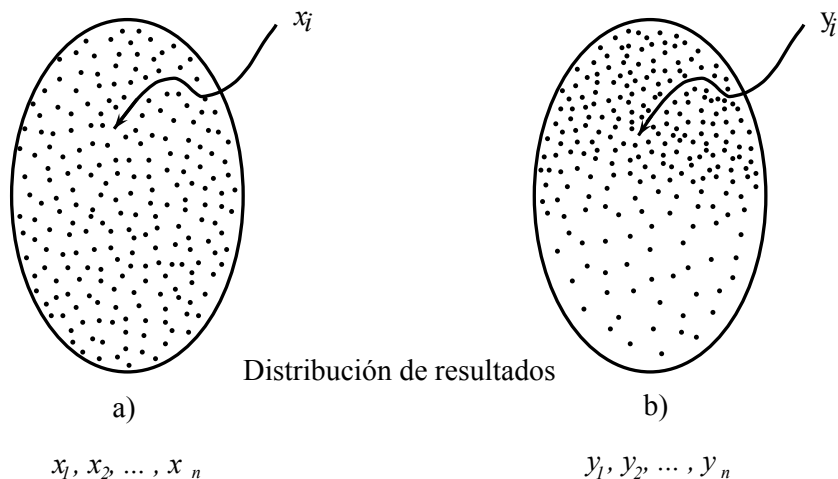


Figura 2.1:

para n grande, es evidente que ambas distribuciones de resultados, como figuras geométricas, como conglomerados de puntos, son notoriamente diferentes; mientras que los resultados del primer colectivo se distribuyen por X en forma aparentemente *uniforme*, ya que no se observa alguna zona de *preferencia* donde se acumulen, donde su grado de concentración, de su frecuencia, sea notoriamente más acentuado que en alguna otra, los resultados del segundo colectivo *tienden a preferir* la parte *superior* de X , aglomeración de resultados evidente que va disminuyendo notablemente conforme se va hacia *abajo* en la representación de X : en la región superior la intensidad, frecuencia o medida de concentración de resultados o puntos de la configuración es más grande que en la región inferior. Explícitamente, en términos cuantitativos, para una misma región R de X , como se bosquejó éste anteriormente, figura 2.2, en cada copia de X , esas diferencias geométricas se manifiestan en que las medidas de concentración de resultados de uno y otro colectivo,

$$\frac{1}{n} \sum_{i=1}^n \delta_R(x_i), \quad \frac{1}{n} \sum_{i=1}^n \delta_R(y_i)$$

son significativamente diferentes en la escala del 0 al 1, y esto para diferentes R , otras zonas, de X .

No es raro encontrar parejas (\mathcal{E}, X) para las cuales las distribuciones de sus colectivos sean notoriamente diferentes, y más aún cuando las condiciones

que rodean las ejecuciones de \mathcal{E} cambian, particularmente con el transcurrir del tiempo, como suele ocurrir con las actividades relacionadas con la economía, con los asuntos financieros e incluso con los de la biología.

Por otro lado, si de cierta pareja (\mathcal{E}, X) los resultados de dos colectivos

$$\begin{array}{l} \text{Ejecución de } \mathcal{E} \quad 1, \quad 2, \quad \dots, \quad i, \quad \dots \\ \text{Resultado} \quad \quad x_1, \quad x_2, \quad \dots, \quad x_i, \quad \dots; \end{array}$$

y

$$\begin{array}{l} \text{Ejecución de } \mathcal{E} \quad 1, \quad 2, \quad \dots, \quad n, \quad \dots \\ \text{Resultado} \quad \quad y_1, \quad y_2, \quad \dots, \quad y_n, \quad \dots; \end{array}$$

se distribuyen aleatoriamente por X en las formas que se ilustran en la figura 2.3, para n grande.

Se estará de acuerdo en que a simple vista ambos conglomerados de puntos tienen *algo* que los asemeja, si bien, punto por punto son diferentes en su carácter de colectivos de puntos, de distribuciones de resultados esparcidos aleatoriamente por X , en ambos hay una gran concentración de puntos en la parte más o menos *central* de X , concentración que va disminuyendo radialmente conforme se avanza hacia la *frontera* del espacio X . O analizándolos desde una perspectiva opuesta: como conglomerados de puntos, como distribuciones de resultados ¿qué los diferencia? Al parecer es difícil, sino imposible (¿será?), dar una respuesta.

Estando así las cosas, no es insensato afirmar que ambas distribuciones de resultados, como conglomerados de puntos, no de resultado o punto a resultado o punto, *son iguales*.

Más aún, lo que los hace parecer *iguales* a simple vista, son los aparentes mismos grados de concentración de puntos de ambos conglomerados en la misma

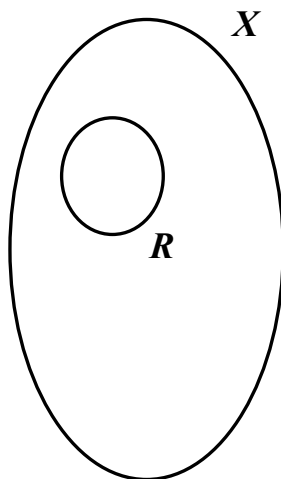


Figura 2.2:

región de X considerada, percibidos inconcientemente e intuitivamente por el observador, razones por las cuales se hace necesario, complementariamente, efectuar un análisis cuantitativo de las medidas que ya se tienen para aquellos conceptos, esto es, comparar las respectivas proporciones

$$\frac{1}{n} \sum_{i=1}^n \delta_R(x_i), \quad \frac{1}{n} \sum_{i=1}^n \delta_R(y_i)$$

para cada $n \gg 1$ y diferentes regiones significativas R de X . Y si éstas son acentuadamente próximas, en el sentido de que

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_R(x_i) - \frac{1}{n} \sum_{i=1}^n \delta_R(y_i) \right| \ll 1, \quad \forall n \gg 1$$

y para las diferentes R , se dirá que ambas distribuciones de resultados o conglomerados de puntos, *son iguales estadísticamente*. O más básicamente, que los colectivos $x_1, x_2, \dots, x_n, \dots$ y $y_1, y_2, \dots, y_n, \dots$, la *materia prima* de (\mathcal{E}, X) , que es a fin de cuentas el producto a analizar, son *iguales estadísticamente*.

2.4. Distribuciones iguales estadísticamente

En esta sección se describirá fenomenológicamente cierta regularidad física propia de los colectivos de *algunas* parejas (\mathcal{E}, X) . Tal regularidad natural es notablemente nítida en multitud de fenómenos de la física relacionados con millones de entes, como pueden ser los átomos, las moléculas, fotones, por ejemplo: en las desintegraciones de los átomos de un elemento radioactivo; en el comportamiento dinámico de las moléculas de un gas; en los cambios de niveles de energía de los electrones de los átomos de un elemento evidenciados por su espectro óptico, el cual es siempre el *mismo*, paradigma éste de la regularidad

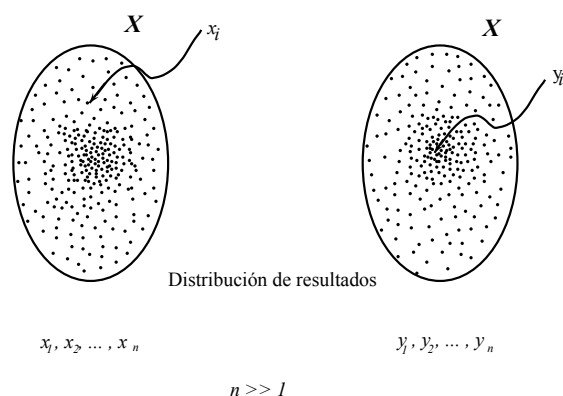


Figura 2.3:

que está por describirse. Tales regularidades, propias de estos fenómenos, son las bases experimentales a partir de las cuales, previa síntesis en forma probabilística, se construyeron, en particular la Mecánica estadística y la Mecánica cuántica. Más aún, tal regularidad es patente en los juegos de azar, como puede corroborarse en los ejemplos expuestos al final de esta sección, o en otros que el lector descubrirá.

Y el fenómeno o la regularidad a que se alude es, en términos generales la *igualdad estadística* de las distribuciones de los resultados de los colectivos de cada una de esas parejas.

Especificando: de una de estas parejas (\mathcal{E}, X) , se tiene u obtiene una multitud de colectivos correspondientes a sendas corridas de ejecuciones de \mathcal{E}

<i>Corrida 1</i>	<i>Ejecución de \mathcal{E}</i>	1,	2,	...	i ,	...
	<i>Resultado</i>	x_1^1 ,	x_2^1 ,	...	x_i^1 ,	...
	⋮			⋮		
<i>Corrida N</i>	<i>Ejecución de \mathcal{E}</i>	1,	2,	...	i ,	...
	<i>Resultado</i>	x_1^N ,	x_2^N ,	...	x_i^N ,	...
	⋮			⋮		

Aclaraciones:

- a) el número asignado a cada corrida no corresponde necesariamente al orden en el que se efectuó o efectúa;
- b) el arreglo que se ha hecho no excluye que los resultados se hayan obtenido de un solo colectivo

$$x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}, x_{2n+1}, \dots, x_{3n}, \dots$$

- i) tomando los *colectivos truncos*

$$\begin{array}{ccc} x_1, & \dots, & x_n \\ x_{n+1}, & \dots, & x_{2n} \\ \vdots & & \\ x_{(N-1)n+1}, & \dots, & x_{(N-1)n+n} \\ & & \vdots \end{array}$$

con n grande y N aún más grande, o

- ii) se tomen los *colectivos desplazados*

$$\begin{array}{ccc} x_1, & \dots, & x_n, & \dots \\ x_2, x_3, & \dots, & x_{n+1}, & \dots \\ \vdots & & & \\ x_N, x_{N+1}, & \dots, & x_{N+n}, & \dots \\ & & \vdots & \end{array}$$

seguramente no se apreciará ningún patrón común en el comportamiento de las sucesiones de resultados, pero si se empieza utilizando el primer método de análisis para estas cuestiones, como es el de graficar un gran número de los resultados sucesivos de cada colectivo en una correspondiente copia de la representación geométrica escogida de X , lo que se observa es que las distribuciones obtenidas, como conglomerados de puntos, no se diferencian unas de otras; no se manifiesta alguna peculiaridad en alguno de los conglomerados que lo distinga de los otros, no obstante que son diferentes uno de otro punto a punto o de resultado a resultado; excepto por estos *accidentes* de carácter aleatorio, en *esencia* las distribuciones son iguales.

El conglomerado de puntos, o distribución de resultados, del colectivo con índice j puede intercambiarse con el de índice k y no se notará ninguna alteración en la colección de gráficas hechas en un principio: no hay ningún criterio que permita decidir cuál es la distribución de resultados, como conglomerados de puntos, correspondientes al colectivo con índice 1, del colectivo con índice 2, ...

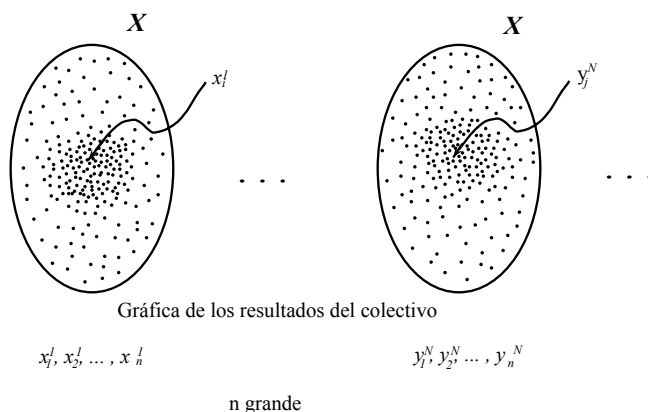


Figura 2.4:

Es más, si se superponen N copias de X , dado que éstas sean transparentes o, si se quiere, graficando los nN puntos de los resultados en una misma representación de X , lo único que se obtiene es un conglomerado de puntos más intenso, tupido o apretado, pero de la misma forma y con las mismas cualidades que las individuales.

Es imposible caracterizar el conglomerado de manera diferente al correspondiente del colectivo k , a menos que se recurra a los detalles de precisar las posiciones de cada uno de los puntos de cada conglomerado, labor ingente pero además inútil ya que el fenómeno que atrae la atención y que interesa es el concerniente a las *distribuciones* de los resultados de los colectivos de esta pareja (\mathcal{E}, X) , al que se aludía al principio de esta sección: la *igualdad* entre esas distribuciones, no su identidad de resultado a resultado, aspecto imposible debido a la aleatoriedad, sino a su *igualdad estadística* como colectividades.

En otras palabras, el fenómeno relevante y distintivo de regularidad de estas distribuciones de resultados de cada colectivo, y que es el que interesa tanto desde el punto de vista físico como del matemático, no son tanto los detalles, *accidentes* de carácter aleatorio que las distinguen, sino los aspectos que las igualan, como lo es su forma, la *esencia*, en tanto conglomerados de puntos.

La situación planteada aparentemente es contradictoria, y desconcertante: para cada repetición de \mathcal{E} se obtiene un resultado aleatorio; para cada corrida de ejecuciones sucesivas de \mathcal{E} se obtiene una sucesión de resultados aleatorios, que a su vez es aleatoria: *a causas iguales, efectos diferentes*. Sin embargo, y esto es lo sorprendente del fenómeno pero que además despeja esa contradicción, las distribuciones de los resultados de cada sucesión de repeticiones de \mathcal{E} , como colectividades numerosas, *determinísticamente* iguales: *a causas iguales, efectos iguales*.

Al efectuar una análisis cuantitativo de esta clase de *igualdad geométrica*, entre las distribuciones de los resultados de los colectivos de (\mathcal{E}, X) , con los recursos de conteo y medición expuestos, se desvelan las siguientes propiedades:

1. Para cada región, R de X , como las que se han bosquejado, las medidas de intensidad con que se *concentran* los resultados de cada colectivo en R o las frecuencias con que los resultados de cada colectivo *caen* en R , son extremadamente próximas:

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_R(x_i^j) - \frac{1}{n} \sum_{i=1}^n \delta_R(x_i^k) \right| \ll 1, \quad \forall n \gg 1,$$

hecho que explica o refleja, según quiera interpretarse, esa *igualdad geométrica esencial* revelada.

2. Los colectivos de 0s y 1s correspondientes a los colectivos de (\mathcal{E}, X) , en cuanto a las realizaciones o no de una región R ,

<i>Corrida 1</i>	<i>Ejecución de \mathcal{E}</i>	1,	2,	...	i ,	...
	<i>Realización de R</i>	$\delta_R(x_1^1)$,	$\delta_R(x_2^1)$,	...	$\delta_R(x_i^1)$,	...
	\vdots	\vdots	\vdots		\vdots	
<i>Corrida N</i>	<i>Ejecución de \mathcal{E}</i>	1,	2,	...	i ,	...
	<i>Realización de R</i>	$\delta_R(x_1^N)$,	$\delta_R(x_2^N)$,	...	$\delta_R(x_i^N)$,	...
	\vdots	\vdots	\vdots		\vdots	

tienen el síndrome de Bernoulli. Estos nuevos colectivos asociados corresponden a la pareja formada por el experimento \mathcal{E} y por el espacio de resultados $\{0, 1\}$: al efectuarse \mathcal{E} , su resultado x , en X , *cae* en R , *realiza* a R , $\delta_R(x) = 1$ o *no cae* en R , *no realiza* a R , $\delta_R(x) = 0$. Así pues, a cada colectivo de (\mathcal{E}, X) y una región R les corresponde un colectivo de la nueva pareja *experimento \mathcal{E} -espacio de resultados $\{0, 1\}_R$* (R , como índice,

es la referencia a la región considerada). Pero además, las características C1 y C2 de (\mathcal{E}, X) se heredan a la pareja $\mathcal{E} - \{0, 1\}_R$, por lo que, según lo convenido en el capítulo I, esta nueva pareja se representará en la forma $(\mathcal{E}, \{0, 1\}_R)$.

Resumidamente: cada pareja $(\mathcal{E}, \{0, 1\}_R)$, correspondiente a (\mathcal{E}, X) y una región R de X , es un experimento de Bernoulli.

Y este es el fenómeno físico, que con los instrumentos analíticos de medición y conteo y el geométrico de graficación, se descubre en un sinnúmero de parejas *experimento-espacio de resultados* con las características C1 y C2, y el cual es la realidad física que subyace en las teorías matemáticas de la probabilidad y los procesos estocásticos.

Así pues, sólo interesarán aquí aquellas parejas (\mathcal{E}, X) para las cuales cada pareja derivada $(\mathcal{E}, \{0, 1\}_R)$, correspondiente a cada región R de la colección \mathcal{R} de todas aquellas regiones de X , es un experimento de Bernoulli

Ahora bien, si (\mathcal{E}, X) es una de estas parejas y \mathcal{R} es la colección de regiones de X , para R de \mathcal{R} , fija, entonces la gráfica de un colectivo de $(\mathcal{E}, \{0, 1\}_R)$,

Ejecución de \mathcal{E}	1,	2,	...	i ,	...
Resultado en X	x_1 ,	x_2 ,	...	x_i ,	...
Resultado en $\{0, 1\}_R$	$\delta_R(x_1)$,	$\delta_R(x_2)$,	...	$\delta_R(x_i)$,	...

persigue, como uno de los síntomas de su síndrome de Bernoulli, una recta $y = p_R x$, con $p_R \in (0, 1)$, a una distancia vertical dada por la condición o regla (o ley) de los números grandes de Bernoulli

$$\left| \sum_{i=1}^n \delta_R(x_i) - np_R \right| \ll 1, \quad \forall n \gg 1,$$

desigualdades equivalentes a las

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_R(x_i) - p_R \right| \ll 1, \quad \forall n \gg 1,$$

(por brevedad, en algunas ocasiones, a estas desigualdades se les llamará *desigualdades de Bernoulli*), asociándose así el número p_R , de manera unívoca, a la región R (en vista del comportamiento, siempre el mismo, esencialmente, de la distribución de los resultados de cada colectivo de (\mathcal{E}, X)). En otras palabras, p_R es una constante física inherente a (\mathcal{E}, X) y la región R , la cual es, según puede leerse en las últimas desigualdades, la medida de qué tan frecuentemente los resultados de cada colectivo de (\mathcal{E}, X) *caen en o realizan la región R* o, equivalentemente, es la *medida de qué tan frecuentemente se realiza R* en cada corrida de ejecuciones sucesivas de \mathcal{E} . O siguiendo la nomenclatura oficial, o usual, p_R es la *medida de probabilidad* de la región R , aunque, debe advertirse, que éste es sólo un término técnico y que de ser usado, únicamente lo será como

alterno y con el mismo significado de los anteriores. Y esto para cada región R en \mathcal{R} .

La asociación natural que se da, debido al fenómeno natural de las distribuciones de los resultados de los colectivos de (\mathcal{E}, X) ,

$$R \rightarrow p_R, \quad \forall R \in \mathcal{R}$$

es la semilla de un concepto teórico fundamental para la construcción de la teoría matemática de la probabilidad, y de los procesos estocásticos, como se verá más adelante, pero que por ahora servirá para hacer la síntesis que a continuación se describe.

Tal correspondencia, natural y empírica, formalmente, independiente del contexto del que surge, es ni más ni menos que una función conjuntista con valores en el intervalo $(0, 1)$, la cual se establecerá como

$$\begin{aligned} \varphi : \mathcal{R} &\rightarrow (0, 1) \ni \\ \varphi(R) &= p_R \quad \forall R \in \mathcal{R} \end{aligned}$$

o, en la notación equivalente, a la manera de sucesiones,

$$(p_R)_{R \in \mathcal{R}}$$

Esta función será llamada *distribución estadística* (o *distribución esencial*) en referencia, claro está, a cómo se distribuyen por X los resultados de cada colectivo de (\mathcal{E}, X) . Y será una síntesis del fenómeno descrito: por un lado resume el hecho determinístico de que todas esas distribuciones son *iguales esencialmente* o *iguales estadísticamente* (implícitamente queda *salvo accidentes aleatorios*), dado por las desigualdades de Bernoulli anteriores, y por otro, describe en términos cuantitativos, por medio de los valores de la función en cada elemento de \mathcal{R} , $\varphi(R) = p_R$, la forma geométrica de la distribución esencial.

Y con objeto de explicar que las distribuciones de los resultados de los colectivos de (\mathcal{E}, X) son *iguales estadísticamente* y además descritas o regidas (según quiera concebirse) esencialmente por la función $\varphi : \mathcal{R} \rightarrow (0, 1)$, su distribución estadística o esencial, se usará la notación

$$(\mathcal{E}, X; (p_R)_{R \in \mathcal{R}}).$$

2.5. Acerca de la distribución $(p_R)_{R \in \mathcal{R}}$

Se tiene un experimento-espacio de resultados del tipo $(\mathcal{E}, X; (p_R)_{R \in \mathcal{R}})$; en lo que respecta únicamente a su distribución $\varphi : \mathcal{R} \rightarrow (0, 1)$, no es, ni más ni menos, que sólo una tabla de valores extraídos del fenómeno: a cada elemento R de \mathcal{R} le corresponde el número $\varphi(R)$ en $(0, 1)$. Desde el enfoque científico, la investigación del tipo de fenómeno físico presentado, la información contenida en esta tabla, por sí misma, es insuficiente. Hay que convertir esa información en conocimiento.

Tanto desde el punto de vista teórico como del experimental es de interés conocer la estructura interna de esa tabla, de esa distribución, pero por ahora considerada única y exclusivamente como función: cómo se relacionan los diferentes números de p_R de los diferentes subconjuntos R , de X , pertenecientes al dominio \mathcal{R} , cuando con éstos se efectúan operaciones conjuntísticas; es necesario encontrar un patrón de comportamiento de éstos números en relación al comportamiento de las regiones a las que corresponden.

Por ejemplo, si se analiza la tabla se encontrará:

- i) Siempre que $R \subset R'$ se tiene que $\varphi(R) \leq \varphi(R')$; pero si, por otro lado, se observa que $\varphi(R) \leq \varphi(R')$ ¿sucederá que $R \subset R'$? Con aquéllo como con esto último ¿por qué se da? ¿a qué se debe?
- ii) Para cualesquiera dos números de la tabla, p_R y $p_{R'}$, se encuentra que

$$\varphi(R) + \varphi(R') = \varphi(R \cup R') + \varphi(RR'),$$

claro está que cuando $R \cup R'$ y RR' resulten ser elementos de \mathcal{R} , o que

$$\varphi(R) + \varphi(R') = \varphi(R + R')$$

si $RR' \neq \emptyset$.

Estos ejemplos son indicios de que existe una estructura interna, pero aun oculta, de la función $\varphi : \mathcal{R} \rightarrow (0, 1)$ de la tabla de valores $(p_R)_{R \in \mathcal{R}}$, consistente en relaciones entre el valor φ de un elemento \mathcal{R} compuesto por otros, por medio de uniones intersecciones o diferencias, elementos de \mathcal{R} , y las sumas, multiplicaciones o diferencias de los valores numéricos de φ de estos elementos componentes. ¿Cuáles son éstas? ¿Por qué se dan? ¿A qué se deben? ¿A partir de qué surgen?

Para responder algunas de estas preguntas hay que regresar al fenómeno de interés acerca de la igualdad estadística de las distribuciones de los resultados de los colectivos de $(\mathcal{E}, X; (p_R)_{R \in \mathcal{R}})$; esas propiedades de la distribución $\varphi : \mathcal{R} \rightarrow (0, 1)$ pueden *deducirse* por medio de argumentos empíricos pero plausibles, partiendo de la desigualdad usual de conteo

$$\sum_{i=1}^n \delta_R(x_i) \leq \sum_{i=1}^n \delta_{R'}(x_i)$$

para i), y para ii) la igualdad familiar de conteo

$$\sum_{i=1}^n \delta_{R \cup R'}(x_i) = \sum_{i=1}^n \delta_R(x_i) + \sum_{i=1}^n \delta_{R'}(x_i) - \sum_{i=1}^n \delta_{RR'}(x_i)$$

para cualquier colección de puntos x_1, \dots, x_n , de X ; pasando luego a un colectivo de la pareja, utilizando después la condición (o ley) de los números grandes de Bernoulli en relación a las regiones participantes, y finalmente con argumentos plausibles acerca de las aproximaciones, y verificables experimentalmente, se

explican las relaciones i) y ii). Aunque no *demostraciones*, ya que no se está en el seno de una estructura matemática.

Así pues, la función $\varphi : \mathcal{R} \rightarrow (0, 1)$ tiene una estructura interna, debida a las características del fenómeno de interés. A partir de esas propiedades i) y las igualdades de ii) pueden *deducirse* otras que aumentan el conocimiento acerca de la distribución, o tabla de valores $(p_R)_{R \in \mathcal{R}}$, y, por lo tanto, acerca de la distribución esencial de las distribuciones de los resultados de los colectivos de $(\mathcal{E}, X; (p_R)_{R \in \mathcal{R}})$, esto por un lado, pero por otro se tiene el aspecto operativo de la función, distribución, $\varphi : \mathcal{R} \rightarrow (0, 1)$, pues, como se dijo antes, se tienen, recetas, algoritmos, para calcular el valor de φ en una región compuesta por uniones, intersecciones o diferencias de otras regiones, en términos de operaciones aritméticas de los valores de φ en estas regiones componentes.

Es para sorprender cómo aquellas regularidades estadísticas repercuten, o se reflejan, en la función $\varphi : \mathcal{R} \rightarrow (0, 1)$ en forma de propiedades, reglas de operación, muy particulares, como lo son a) que es *monótona*, propiedad i) anterior, y b) que es *aditiva*, propiedades ii) anteriores.

Ahora bien, cabe señalar que este tipo de funciones conjuntistas con valores reales no negativos, con tales propiedades, se encuentran frecuentemente en diversos campos de la matemática, basta con mencionar el concepto de longitud en \mathbb{R} , el concepto de área en \mathbb{R}^2 , el concepto de volumen en \mathbb{R}^3, \dots . Y es tal su don de ubicuidad que existe una rama de la matemática dedicada exclusivamente a su estudio, la *Teoría de la medida de conjuntos*.

Surge, entonces, la pregunta ¿qué propiedades, qué aspectos, qué hechos, concernientes únicamente a la función $\varphi : \mathcal{R} \rightarrow (0, 1)$ como tal, independientes del contexto experimental del que se obtuvo, son la causa de su estructura interna, manifestada, hasta ahora, en las dos características establecidas? La pregunta es pertinente ya que, como se mencionó arriba, existen este tipo de funciones en ámbitos muy ajenos y diferentes totalmente del que de aquí se obtuvo.

No sobra mencionar que se tiene que recurrir a aquella teoría matemática para dar respuestas a las preguntas hechas; más específicamente, a la *teoría de la medida normal* (o de *probabilidad*), y esto en vista de otra propiedad de la función $\varphi : \mathcal{R} \rightarrow (0, 1)$ que no se ha atendido, dado que X no se ha considerado como elemento de \mathcal{R} y es que necesariamente debe tenerse que

$$\varphi(X) = 1$$

propiedad que proviene básicamente del conteo

$$\sum_{i=1}^n \delta_X(x_i) = n$$

para cualesquiera n puntos x_1, \dots, x_n de X . Y, aprovechando el viaje, se puede agregar que debe tenerse, también, que,

$$\varphi(\phi) = 0$$

ya que de inicio sucede que ningún x_i está en el ϕ , así que

$$\sum_{i=1}^n \delta_{\phi}(x_i) = 0.$$

Ambas propiedades se presentan en relación a cualquier pareja experimento-espacio de resultados, haya o no las regularidades estadísticas establecidas, en vista de que siempre son válidas ambas sumas.

2.6. Ejemplos

Entre la multiplicidad de parejas (\mathcal{E}, X) cuyos colectivos son *iguales estadísticamente*, están los presentados en el capítulo I:

Ejemplo 1.

En relación por la extracción de bolas de un recipiente, se tiene la pareja $(\mathcal{E}, X_1 = \{0, 1\})$; ahora, considerando la familia de regiones $\mathcal{R} = \{\{0\}, \{1\}\}$ de $X_1 = \{0, 1\}$, resulta que las parejas $(\mathcal{E}, \{0, 1\}_{\{0\}})$ y $(\mathcal{E}, \{0, 1\}_{\{1\}})$ son experimentos de Bernoulli, esto es, que para cada colectivo de $(\mathcal{E}, X_1 = \{0, 1\})$

Ejecución de \mathcal{E}	1,	2,	...	i ,	...
Resultado en X	x_1 ,	x_2 ,	...	x_i ,	...

los colectivos correspondientes

$$\begin{array}{ccccccc} \delta_{\{0\}}(x_1), & \delta_{\{0\}}(x_2), & \dots, & \delta_{\{0\}}(x_i), & \dots & & \\ \delta_{\{1\}}(x_1), & \delta_{\{1\}}(x_2), & \dots, & \delta_{\{1\}}(x_i), & \dots & & \end{array}$$

(Nota: con la notación usada en el capítulo I, δ_i es x_i y se tiene que

$$\delta_{\{0\}}(x_i) = 1 - \delta_i, \quad \delta_{\{1\}} = \delta_i \quad)$$

tienen su propio síndrome de Bernoulli, los cuales se diferencian sólo en el hecho de que la gráfica correspondiente a $(\mathcal{E}, \{0, 1\}_{\{0\}})$ persigue a la recta $y = (1 - p)x$, manteniéndose sus distancias verticales según la condición de los números grandes de Bernoulli

$$\left| \sum_{i=1}^n \delta_{\{0\}}(x_i) - n(1 - p) \right| \ll n, \quad \forall n \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_{\{0\}}(x_i) - (1 - p) \right| \ll 1, \quad \forall n \gg 1 \quad ;$$

mientras que la gráfica del colectivo de $(\mathcal{E}, \{0, 1\}_{\{1\}})$ persigue a la recta $y = px$ a las distancias verticales sometidas a las desigualdades de Bernoulli,

$$\left| \sum_{i=1}^n \delta_{\{1\}}(x_i) - np \right| \ll n, \quad \forall n \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_{\{1\}}(x_i) - p \right| \ll 1, \quad \forall \quad n \gg 1.$$

Así pues, la función

$$\begin{aligned} \varphi : \mathcal{R} &\rightarrow (0, 1) \ni \\ \varphi(\{0\}) &= 1 - p, \quad \varphi(\{1\}) = p \end{aligned}$$

es la *distribución* (estadística) que rige o describe, *esencialmente*, la distribución de los resultados de cada colectivo de $(\mathcal{E}, X_1 = \{0, 1\})$ por $X_1 = \{0, 1\}$, en el sentido de que el valor de φ en R de \mathcal{R} , $\varphi(R)$, es la medida de qué tan frecuentemente resultados del colectivo *caen* en la región R o, equivalentemente, la medida de qué tan frecuentemente se *realiza* el subconjunto R en cada corrida de ejecuciones sucesivas de \mathcal{E} ; pero esto, salvo los *accidentes aleatorios*, en el sentido de que las frecuencias relativas

$$\frac{1}{n} \sum_{i=1}^n \delta_R(x_i)$$

y la medida $\varphi(R)$ difieren según las desigualdades, o condiciones de los números grandes, de Bernoulli:

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_R(x_i) - \varphi(R) \right| \ll 1, \quad \forall \quad n \gg 1$$

para cada R de \mathcal{R} .

Cabe agregar que, en el lenguaje usual, como una síntesis del fenómeno de interés acerca del comportamiento de los colectivos de $(\mathcal{E}, X_1 = \{0, 1\})$, $\varphi(R)$ es la medida de probabilidad de R , $R \in \mathcal{R}$.

Ejemplo 2.

El experimento \mathcal{E} es *extraer* cuatro dígitos al azar, en la forma descrita al final del capítulo I; la categoría de resultados que interesan a este experimento es la de las parejas ordenadas $(.d_1d_2, .d_3d_4)$, construidas a partir de las cuartetas de dígitos, también ordenadas, (d_1, d_2, d_3, d_4) , obtenidas al ejecutarse \mathcal{E} . El espacio de resultados será, entonces, el conjunto

$$X = \{(.x_1x_2, .y_1y_2) : x_i = 0, 1, \dots, 9, \forall \quad i = 1, 2; \quad y_j = 0, 1, \dots, 9, \forall \quad j = 1, 2\}$$

Esta pareja *experimento* \mathcal{E} -*espacio de resultados* X tiene las características C1 y C2, razón por la cual se escribirá en la forma (\mathcal{E}, X) .

Como representación geométrica de X se considerará el cuadrado unitario $[0, 1) \times [0, 1)$ de \mathbb{R}^2 , que a su vez puede representarse por un cuadrado de 10cm x 10cm de una hoja de papel milimétrico, el cual tiene la ventaja de que los elementos de X quedan en correspondencia biunívoca con los puntos de la cuadrícula de este cuadrado.

Las distribuciones de los resultados de los colectivos de (\mathcal{E}, X) son iguales estadísticamente, con la característica adicional de que son uniformes o, equivale a decir, que sus distribuciones son uniformes estadísticamente.

Al ir graficando, en la hoja de papel milimétrico, los resultados de un colectivo de (\mathcal{E}, X)

<i>Ejecución de \mathcal{E}</i>	1, 2, ..., i , ...
<i>Resultado</i>	$\bar{x}_1, \bar{x}_2, \dots, \bar{x}_i, \dots$

con $\bar{x}_i = (.x_1^i x_2^i, .y_1^i y_2^i)$, estos van *cayendo* o *apareciendo* aleatoriamente en la hoja y cuando n es grande, p. ej. $n = 1000$, la distribución de resultados es notablemente uniforme, en el sentido de que el número de resultados que cayeron, por ejemplo, en cada una de las regiones $R_{(k,l)} = [.k, .(k+1)) \times [.l, .(l+1))$, $k, l = 0, 1, \dots, 9$, de $[0, 1) \times [0, 1)$, es más o menos el mismo o que el número de resultados que cayeron en uno de estos cuadritos no difiere escandalosamente del número de los que cayeron en otro; fenómeno de emparejamiento que se acentúa conforme el número de resultados que se grafican, n , crece.

Expresándolo con los símbolos de conteo, para ser más precisos: para n resultados, n grande, los $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$ se distribuyen por $[0, 1) \times [0, 1)$ en forma más o menos *equitativa* entre los 100 cuadritos $R_{(k,l)}$, es decir, $\frac{n}{100}$ resultados, aproximadamente, en cada cuadrito,

$$\sum_{i=1}^n \delta_{(k,l)}(\bar{x}_i) \approx \frac{n}{100} = na(R_{(k,l)})$$

siendo $\delta_{(k,l)}$ “ δ índice $R_{(k,l)}$ ” y $a(R_{(k,l)}) = \text{área de } R_{(k,l)}$. Y si R es una región *decente* de $[0, 1) \times [0, 1)$, una función que pueda *aproximarse* en su forma, tanto por *dentro* como por *fuera*, con uniones de cuadritos $R_{(k,l)}$, como, por ejemplo

$$R_1 = \{(x, y) : x^2 + y^2 \leq 1\}, \quad R_2 = \{(x, y) : x \leq y\}, \quad R_3 = \{(x, y) : x^2 \geq y\}$$

entonces

$$\sum_{i=1}^n \delta_R(\bar{x}_i) \approx na(R)$$

o, lo que es equivalente,

$$\frac{1}{n} \sum_{i=1}^n \delta_R(\bar{x}_i) \approx na(R), \quad \forall n \gg 1 :$$

la proporción de resultados de n ejecuciones sucesivas de \mathcal{E} que caen en la región R es aproximada al área de R ; expresiones que son la primera traducción a términos cuantitativos del fenómeno sistemático que se presenta en relación a la *forma geométrica* de las distribuciones de los resultados de los colectivos de (\mathcal{E}, X) , la de su *uniformidad*.

Sin embargo, esas expresiones son una consecuencia de un comportamiento más sutil del comportamiento de los colectivos de (\mathcal{E}, X) , el cual es que las

parejas asociadas $(\mathcal{E}, \{0, 1\}_R)$ son experimentos de Bernoulli, para cada región R como las descritas anteriormente; a costa de repetir: $(\mathcal{E}, \{0, 1\}_R)$ es la pareja *experimento \mathcal{E} -espacio de resultados* $\{0, 1\}$ de la cual al efectuar \mathcal{E}

su resultado \bar{x} cae en o realiza a R , entonces $\delta_R(\bar{x}) = 1$,

su resultado \bar{x} no cae en o no realiza a R , entonces $\delta_R(\bar{x}) = 0$,

siendo sus colectivos derivados de los de (\mathcal{E}, X) ; así, para el colectivo anterior de ésta, el colectivo asociado de $(\mathcal{E}, \{0, 1\}_R)$ es

$$\delta_R(\bar{x}_1), \delta_R(\bar{x}_2), \dots, \delta_R(\bar{x}_n), \dots$$

Pues bien, este tipo de colectivos tienen el síndrome de Bernoulli, siendo la recta que persiguen sus gráficas

$$y = a(R)x$$

y manteniéndose las distancias verticales entre ésta y aquéllas según las condiciones (o ley) de los números grandes de Bernoulli; así, por ejemplo, para el colectivo anterior, se tiene que

$$\left| \sum_{i=1}^n \delta_R(\bar{x}_i) - na(R) \right| \ll n, \quad \forall \quad n \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_R(\bar{x}_i) - a(R) \right| \ll 1, \quad \forall \quad n \gg 1.$$

Resumiendo: el fenómeno de interés, que sucede, en relación a las distribuciones de los resultados de los colectivos de la pareja (\mathcal{E}, X) , es que

- i) son iguales estadísticamente
- ii) la *forma* de las distribuciones es la que se denomina *distribución uniforme*; los resultados de aquéllos se distribuyen por X *uniformemente*.

Ahora bien, en cuanto a la distribución estadística asociada a este fenómeno, se tiene, en primer lugar que la constante física intrínseca a (\mathcal{E}, X) y la región R de X (o de $[0, 1] \times [0, 1]$) es el número $a(R)$, área de R , lo cual da lugar a la función conjuntista

$$\begin{aligned} \varphi : \mathcal{R} &\rightarrow (0, 1) \ni \\ \varphi(R) &= a(R), \quad \forall \quad R \in \mathcal{R} \end{aligned}$$

o, en la notación equivalente,

$$(a(R))_{R \in \mathcal{R}};$$

siendo \mathcal{R} la colección de todas aquellas regiones de X (o de $[0, 1] \times [0, 1]$) descritas anteriormente.

Y esta función es la *distribución estadística* de las distribuciones de los resultados de cada colectivo de (\mathcal{E}, X) , *esencialmente*, salvo los *accidentes aleatorios*, en el sentido descrito y cuantificado por las desigualdades de Bernoulli.

Todo esto queda condensado, simbólicamente, escribiendo $(\mathcal{E}, X; (a(R))_{R \in \mathcal{R}})$ en lugar de sólo (\mathcal{E}, X) .

Por último, no debe dejarse pasar la oportunidad de aludir a una técnica (conocida como método de Monte Carlo) muy interesante que usa esta propiedad determinística de las distribuciones de resultados de cada colectivo de (\mathcal{E}, X) , para calcular, de manera aproximada, valores de integrales de funciones, por ejemplo continuas

$$f : [0, 1] \rightarrow [0, 1],$$

$$\int_0^1 f(x) dx,$$

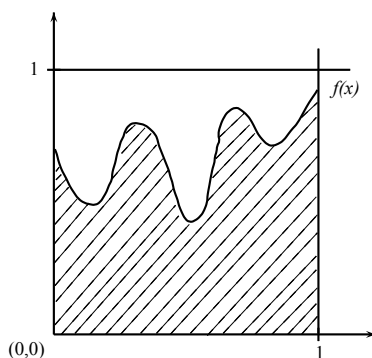


Figura 2.5:

ya que si

$$R = \{(x, y) : 0 \leq y \leq f(x), \quad x \in [0, 1]\}$$

en una región de $[0, 1] \times [0, 1]$ (o de $[0, 1] \times [0, 1]$) “decente”, en vista de que para cada colectivo de 0’s y 1’s, derivado de (\mathcal{E}, X) , para R , por ejemplo del que se tiene al principio,

$$\delta_R(\bar{x}_1), \delta_R(\bar{x}_2), \dots, \delta_R(\bar{x}_n), \dots$$

tiene el síndrome de Bernoulli y, en especial por tener la propiedad (o ley) de los números grandes de Bernoulli,

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_R(\bar{x}_i) - \int_a^b f(x) dx \right| \ll 1, \quad \forall n \gg 1,$$

se obtiene aquella aproximación que se mencionó al cuantificar la uniformidad,

$$\frac{1}{n} \sum_{i=1}^n \delta_R(\bar{x}_i) \approx \int_a^b f(x) dx, \quad \forall n \gg 1;$$

esto es, el número

$$\frac{1}{n} \sum_{i=1}^n \delta_R(\bar{x}_i)$$

obtenido experimentalmente, cuando n es grande, es muy aproximado a esa integral.

En particular, si

$$f(x) = \sqrt{1-x^2}, \quad x \in [0, 1]$$

se tiene que

$$\int_0^1 f(x) dx = \frac{\pi}{4},$$

así que para la región

$$R = \{(x, y) : 0 \leq x^2 + y^2; 0 \leq x \leq 1, 0 \leq y \leq 1\}$$

la proporción

$$\frac{1}{n} \sum_{i=1}^n \delta_R(\bar{x}_i), \quad n \gg 1$$

es muy próxima a

$$\frac{\pi}{4},$$

dándose así el hecho sorprendente de calcular π , aproximadamente, por medio de extracciones al azar de dígitos, por métodos experimentales de carácter estadístico. (Como puede verse, este es un método mucho más sencillo que el método famoso de la Aguja del Conde Buffon).

¿Qué tan próxima es una proporción de las anteriores, correspondiente al área de la región, dependiendo de n ? La respuesta de carácter teórico se dará más adelante, cuando se traten las desigualdades de Tchebishef y del logaritmo iterado para los colectivos de Bernoulli.

Capítulo 3

Medidas normales y distribuciones (o medidas de probabilidad)

3.1. Medidas normales.

La *medida normal* es el concepto fundamental de la teoría matemática acerca de las distribuciones estadísticas. Aquí se tratará cierto tipo de medidas normales, aunque lo suficientemente general para comprender diversos aspectos relacionados a las regularidades estadísticas de las parejas $(\mathcal{E}, X; (p_R)_{R \in \mathcal{R}})$.

Se comienza con un universo X que conste de al menos dos elementos

I. Se *despedaza* en un número finito de subconjuntos, *pedazos*, P_1, \dots, P_M tales que

1. $\emptyset \neq P_i \neq X$, $i = 1, \dots, M$;
2. $P_i P_j = \emptyset$, $i \neq j$ (AB es otra forma de denotar la intersección $A \cap B$); y
3. $\sum_{i=1}^M P_i = X$; ($\sum_{i=1}^M P_i = P_1 + \dots + P_M$ es $P_1 \cup \dots \cup P_M$, pero se escribirá como *suma* cuando los conjuntos que se *unan* sean ajenos);

esta colección particular de subconjuntos de X , $\mathcal{P} = \{P_1, \dots, P_M\}$ se denomina *partición* (finita) de X y a sus elementos, P_i , se les llamará *pedazos*.

II. A cada pedazo P_i , perteneciente a \mathcal{P} , se le asocia un número positivo p_i , arbitrario, lo cual define una función conjuntista,

$$m : \mathcal{P} \rightarrow (0, +\infty) \ni$$

$$m(P_i) = p_i \quad :$$

el valor de la función en el elemento P_i de \mathcal{P} es p_i ; a esta función se le llama *medida básica*. Pero las que interesan aquí serán las *medidas básicas normales*, aquellas medidas básicas para las cuales la suma de los valores p_1, \dots, p_M asignados a los correspondientes pedazos P_1, \dots, P_M , suman 1:

$$\begin{aligned} m(P_1) + \dots + m(P_M) &= p_1 + \dots + p_M \\ &= 1. \end{aligned}$$

Es más, a partir de una medida básica

$$m : \mathcal{P} \rightarrow (0, +\infty)$$

siempre puede obtenerse una medida básica normal, *normalizándola*: por medio de la asociación

$$P_i \rightarrow \frac{p_i}{p_1 + \dots + p_M}, \quad \forall i = 1, \dots, M.$$

Con objeto de tener claridad acerca de lo que se está haciendo; quizás sea conveniente dar unos cuantos ejemplos concretos y que, además, serán los que interesarán para el futuro desarrollo del tema principal.

Ejemplos

1. Medida básica de Bernoulli.

Si el universo del que se parte es $X_1 = \{0, 1\}$, su única partición es

$$\mathcal{P}_1 = \{\{0\}, \{1\}\}.$$

Ahora, tomando cualquier p del intervalo $(0, 1)$ se construye la función conjuntista

$$\begin{aligned} m_1 : \mathcal{P}_1 &\rightarrow (0, 1) \ni \\ m_1(\{1\}) &= p, \quad m_1(\{0\}) = 1 - p; \end{aligned}$$

la cual resulta ser una medida básica normal, que se denominará *medida básica (normal) de Bernoulli*.

2. Para el universo $X = [0, 1) \times [0, 1)$, considérense sus cien subconjuntos $[\cdot d_1, \cdot(d_1 + 1) \times [\cdot d_2, \cdot(d_2 + 1))$, con $d_i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$, para $i = 1, 2$; la colección de estos subconjuntos de X ,

$$\mathcal{P} = \{[\cdot d_1, \cdot(d_1 + 1) \times [\cdot d_2, \cdot(d_2 + 1) : d_i = 0, \dots, 9; i = 1, 2\}$$

es una partición de X .

Si a cada pedazo de la partición se le asocia su área, la función que se obtiene

$$m : \mathcal{P} \rightarrow (0, 1) \ni$$

$$\begin{aligned} m([\cdot d_1, \cdot(d_1 + 1) \times [\cdot d_2, \cdot(d_2 + 1)]) &= (\cdot(d_1 + 1) - \cdot d_1)(\cdot(d_2 + 1) - \cdot d_2) \\ &= \frac{1}{10} \times \frac{1}{10} \\ &= 10^{-2} \end{aligned}$$

es una medida básica normal, del tipo *cuasiuniforme*.

3. Para cada número natural n se tiene el universo

$$X_n = \{(\delta_1, \dots, \delta_n) : \delta_i = 0 \text{ ó } 1, \forall i = 1, \dots, n\};$$

una representación geométrica de este conjunto es la colección de las gráficas de las sucesiones o vectores $(\delta_1, \dots, \delta_n)$ en el plano \mathbb{R}^2 , como se hizo en el Cap. I: la gráfica de $(\delta_1, \dots, \delta_n)$ es la sucesión de puntos $(k, \delta_1 + \dots + \delta_k)$, $k = 1, 2, \dots, n$; e inversamente a cada sucesión de puntos \mathbb{R}^2

$$(1, n_1), (2, n_2), \dots, (n, n_n),$$

con las condiciones

$$0 \leq n_1 \leq 1$$

y

$$n_k \leq n_{k+1} \leq n_k + 1, \quad \forall k = 1, \dots, n-1,$$

siendo cada n_k un entero no negativo pero además $n_k \leq k$, le corresponde un y sólo un elemento de X_n .

Con todo esto se establece una correspondencia biunívoca entre los elementos de X_n y las gráficas del tipo descrito.

Por otro lado, la partición de X_n que se tomará será

$$\mathcal{P}_n = \{P_0, P_1, \dots, P_n\}$$

de la cual cada pedazo es el subconjunto

$$P_k = \{(\delta_1, \dots, \delta_n) : \delta_1 + \dots + \delta_n = k\}, \quad k = 0, 1, \dots, n;$$

P_k consta de todas aquellas sucesiones de X_n que tienen solamente k 1s, y su representación geométrica es la colección aquellas sucesiones de puntos, gráficas correspondientes a sus elementos, que parten de $(1, 0)$ ó $(1, 1)$ y terminan en (n, k) .

Entre la infinidad de métodos de asignación de reales positivos que pueden asociarse a los pedazos de la partición \mathcal{P}_n , hay uno especialmente importante y es aquel que se hace mediante la llamada *función binomial* de parámetros n y p , siendo n número natural y p un real $0 < p < 1$:

$$b_n(p; k) := \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n;$$

debido a que se tiene

a)

$$0 < \binom{n}{k} p^k (1-p)^{n-k} < 1, \quad \forall k = 0, 1, \dots, n,$$

recordando que

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

y

b)

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

se construye la *medida (normal) básica binomial*

$$m_n : \mathcal{P}_n \rightarrow (0, 1) \ni$$

$$m_n(P_k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k = 0, 1, \dots, n.$$

El ejemplo 1 es un caso particular del presentado aquí, cuando en éste $n = 1$.

Es suficiente, por ahora, con estos ejemplos concretos de diferentes medidas normales básicas; posteriormente aparecerán, de forma natural, otras, al avanzar en el desarrollo de la teoría que aquí se presentará.

El paso siguiente es la *extensión* de la medida básica $m : \mathcal{P} \rightarrow (0, 1)$.

III. Se construye lo que se llama en la teoría de la medida de conjuntos, el *álgebra generada* por \mathcal{P} :

$$a(\mathcal{P}) = \left\{ \sum_{i \in N} P_i : \forall N \subset \{1, \dots, M\} \right\};$$

$a(\mathcal{P})$ es la colección de todas las sumas (uniones de conjuntos ajenos) de pedazos pertenecientes a \mathcal{P} .

Entre las propiedades que interesa resaltar, están

1. Como $\{i\} \subset \{1, \dots, M\}$, $\forall i = 1, \dots, M$, entonces

$$\sum_{j \in \{i\}} P_j = P_i$$

$$\therefore P_i \in a(\mathcal{P}), \quad \forall i = 1, \dots, M$$

$$\therefore \mathcal{P} \subset a(\mathcal{P}).$$

2. Dado que \emptyset es subconjunto de $\{1, \dots, M\}$, es razonable definir

$$\sum_{j \in \emptyset} P_j = \emptyset,$$

$$\therefore \emptyset \in a(\mathcal{P}).$$

3. Como $\{1, \dots, M\} \subset \{1, \dots, M\}$ y

$$X = \sum_{j=1}^M P_j = \sum_{j \in \{1, \dots, M\}} P_j,$$

entonces

$$X \in a(\mathcal{P});$$

4. Si $A, B \in a(\mathcal{P})$, sucede que

i) $A \cup B \in a(\mathcal{P})$,

ii) $AB \in a(\mathcal{P})$,

iii) $A - B \in a(\mathcal{P})$,

iv) $A^c \in a(\mathcal{P})$.

Como A, B pertenecen a $a(\mathcal{P})$, entonces son de la forma

$$A = \sum_{i \in N} P_i, \quad B = \sum_{i \in M} P_i,$$

siendo N, M subconjuntos del conjunto de índices $\{1, \dots, M\}$,

$$\therefore A \cup B = \left(\sum_{i \in N} P_i \right) \cup \left(\sum_{j \in M} P_j \right) = \sum_{k \in N + (M - N)} P_k,$$

siendo claro que $N \cup M = N + (M - N) \subset \{1, \dots, M\}$.

$$AB = \left(\sum_{i \in N} P_i \right) \left(\sum_{j \in M} P_j \right) = \sum_{k \in NM} P_k,$$

siendo $NM \subset \{1, \dots, M\}$;

$$A - B = \sum_{i \in N} P_i - \sum_{j \in M} P_j = \sum_{k \in N - M} P_k$$

con $N - M \subset \{1, \dots, M\}$;

$$A^c = \left(\sum_{i \in N} P_i \right)^c = \sum_{k \in \{1, \dots, M\} - N} P_k,$$

con $\{1, \dots, M\} - N \subset \{1, \dots, M\}$.

5. Si A_1, A_2, \dots, A_k son elementos de $a(\mathcal{P})$, entonces

$$A_1 \cup A_2 \cup \dots \cup A_k \in a(\mathcal{P})$$

y

$$A_1 A_2 \dots A_k \in a(\mathcal{P}).$$

No está de más observar que estas propiedades, y otras que se usarán posteriormente, pueden deducirse sólo de las propiedades i) y iv) o de ii) y iv); es más, una familia de subconjuntos \mathcal{a} de un universo Y es un *álgebra de subconjuntos* si

$$\text{a) } \forall A \in \mathcal{a} \Rightarrow A^c \in \mathcal{a}$$

y

$$\text{b) } \forall A, B \in \mathcal{a} \Rightarrow A \cup B \in \mathcal{a}$$

o, la definición equivalente

$$\text{a') } \forall A \in \mathcal{a} \Rightarrow A^c \in \mathcal{a}$$

y

$$\text{b') } \forall A, B \in \mathcal{a} \Rightarrow AB \in \mathcal{a},$$

así que la colección de subconjuntos $a(\mathcal{P})$ es una álgebra de subconjuntos del universo X y es llamada el *álgebra generada por \mathcal{P}* , la cual, además será el dominio de la extensión de $m : \mathcal{P} \rightarrow (0, 1)$.

IV. A partir de la función $m : \mathcal{P} \rightarrow (0, 1)$, se construye la función

$$\mu : a(\mathcal{P}) \rightarrow [0, 1] \ni$$

$$\begin{aligned} \mu \left(\sum_{i \in N} P_i \right) &:= \sum_{i \in N} m(P_i) \\ &= \sum_{i \in N} p_i, \quad \forall \sum_{i \in N} P_i \in a(\mathcal{P}), \end{aligned}$$

La propiedad característica de esta nueva función es la llamada *Aditividad (finita)*: para dos elementos de $a(\mathcal{P})$, A y B , ajenos, se cumple la igualdad

$$\mu(A + B) = \mu(A) + \mu(B),$$

(Como

$$A = \sum_{i \in I} P_i, \quad I \subset \{1, \dots, M\},$$

$$B = \sum_{j \in J} P_j, \quad J \subset \{1, \dots, M\}$$

y

$$IJ = \emptyset,$$

entonces

$$\begin{aligned} A + B &= \sum_{k \in I+J} P_k, \\ \therefore \mu(A + B) &= \sum_{k \in I+J} m(P_k) \\ &= \sum_{k \in I} m(P_k) + \sum_{k \in J} m(P_k) \\ &= \mu(A) + \mu(B). \end{aligned}$$

a partir de lo cual se obtiene que, para un número arbitrario de elementos de $a(\mathcal{P})$, A_1, A_2, \dots, A_k , ajenos entre sí,

$$\mu(A_1 + A_2 + \dots + A_k) = \mu(A_1) + \mu(A_2) + \dots + \mu(A_k).$$

Otras propiedades de la función $\mu : a(\mathcal{P}) \rightarrow [0, 1]$ que desde ahora conviene mencionar y que son usadas frecuentemente son

$$1. \mu(\emptyset) \leq \mu(A) \leq \mu(X) \quad \forall A \in a(\mathcal{P}).$$

(Como $A \in a(\mathcal{P})$, entonces para algún $N \subset \{1, \dots, M\}$ debe tenerse

$$A = \sum_{i \in N} P_i$$

$$\therefore \mu(A) = \sum_{i \in N} m(P_i),$$

pero

$$X = \sum_{j \in \{1, \dots, M\}} P_j,$$

entonces

$$\mu(X) = \sum_{j \in \{1, \dots, M\}} m(P_j) \geq \sum_{i \in N} m(P_i) = \mu(A)$$

y como

$$\mu(A) = \sum_{i \in N} m(P_i) \geq \sum_{j \in \emptyset} m(P_j) = \mu(\emptyset),$$

se obtiene esa cadena de desigualdades.)

$$2. \mu(\emptyset) = 0$$

(Como $\emptyset \in a(\mathcal{P})$ y $\emptyset + \emptyset = \emptyset$, entonces

$$\mu(\emptyset + \emptyset) = \mu(\emptyset),$$

y como la función es aditiva,

$$\mu(\emptyset + \emptyset) = \mu(\emptyset) + \mu(\emptyset)$$

$$\therefore \mu(\emptyset) + \mu(\emptyset) = \mu(\emptyset),$$

pero debido a que los valores de la función son números reales, esta igualdad implica que

$$\mu(\emptyset) = 0 \quad)$$

3. Si $A \subset B$, siendo A y B elementos de $a(\mathcal{P})$, entonces

i) $\mu(B - A) = \mu(B) - \mu(A)$;

ii) la función es *monótona*:

$$\mu(B) \geq \mu(A) .$$

(Como $A \subset B$, entonces

$$B = A + (B - A),$$

$$\therefore \mu(B) = \mu(A) + \mu(B - A)$$

$$\therefore \mu(B) - \mu(A) = \mu(B - A) ;$$

como $\mu(B - A) \geq 0$, y por estando operando con números reales, se obtiene

$$\mu(B) \geq \mu(A) .$$

4. Para dos elementos arbitrarios de $a(\mathcal{P})$, A y B , se tiene la igualdad

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(AB).$$

(Como siempre, entre conjuntos de un mismo universo se tiene la igualdad conjuntista

$$A \cup B = A + (B - AB)$$

y, en el caso particular en que se está operando, todos los conjuntos que en ésta participan pertenecen a $a(\mathcal{P})$; por la propiedad aditiva de la función se tiene

$$\mu(A \cup B) = \mu(A) + \mu(B - AB).$$

Y como $B \supset AB$, por i) del inciso 3 anterior se logra

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(AB). \quad)$$

5. $\mu(A^c) + \mu(A) = 1$.

Debe notarse que las propiedades 2, 3, 4 y 5 de la función $\mu : a(\mathcal{P}) \rightarrow [0, 1]$ son deducidas del hecho, solamente, de ser ésta aditiva y no de las formas particulares de los elementos de $a(\mathcal{P})$ y de los valores de la función en éstos; en otras palabras, si en un universo arbitrario $Y \neq \emptyset$, se tiene

i) Un álgebra de subconjuntos $a \subset \mathcal{P}(Y)$

y

ii) Una *medida*, esto es, una función

$$\nu : \mathcal{a} \rightarrow [0, +\infty)$$

aditiva, lo cual quiere decir que para A, B , ajenos entre sí, elementos de \mathcal{a} , se cumple que

$$\nu(A + B) = \nu(A) + \nu(B),$$

propiedad de la cual se obtienen las siguientes relaciones entre el valor de la función de una expresión algebraico-conjuntista de elementos de \mathcal{a} y operaciones aritméticas de los valores de la función en estos elementos:

a) $\nu(\emptyset) = 0$.

b) Si $A, B \in \mathcal{a}$ y $A \subset B$, entonces

i) $\nu(B - A) = \nu(B) - \nu(A)$

ii) *monotonía*: $\nu(B) \geq \nu(A)$.

c) Para cualesquiera dos elementos de \mathcal{a} , A y B

$$\nu(A \cup B) + \nu(AB) = \nu(A) + \nu(B).$$

d)

$$0 \leq \nu(A) \leq \nu(X), \quad \forall A \in \mathcal{a}.$$

Así pues, la función con valores en los reales no negativos

$$\mu : \mathcal{a}(\mathcal{P}) \rightarrow [0, 1]$$

es una *medida*, pero con la particularidad de ser *normal*, esto es, que

6. $\mu(X) = 1$

(Como $X = P_1 + \dots + P_M$, entonces $X \in \mathcal{a}(\mathcal{P})$ y

$$\begin{aligned} \mu(X) &= m(P_1) + \dots + m(P_M) \\ &= p_1 + \dots + p_M \\ &= 1 \end{aligned} \quad)$$

y, por lo tanto, con el auxilio de la propiedad 1,

7. $0 \leq \mu(A) \leq 1, \quad \forall A \in \mathcal{a}(\mathcal{P}),$

lo cual justifica que como codominio de esta medida se haya puesto el intervalo cerrado $[0, 1]$ en lugar de $[0, +\infty)$.

Ahora se presentará un teorema acerca de la unicidad de la extensión $\mu : \mathcal{a}(\mathcal{P}) \rightarrow [0, 1]$ de $m : \mathcal{P} \rightarrow (0, 1)$.

La medida normal $\mu : \mathcal{a}(\mathcal{P}) \rightarrow [0, 1]$ es la única medida normal cuyos valores en los elementos de \mathcal{P} , $\mu(P_i)$, son iguales a los valores $m(P_i)$.

(La demostración que se hará mostrará otros aspectos que cobrarán un significado más amplio: supóngase que se tiene una medida normal $\tilde{\mu} : \mathcal{a} \rightarrow [0, 1]$ tal que

i) \mathcal{a} es un álgebra de subconjuntos de X que contiene a \mathcal{P}

y

ii) $\tilde{\mu}(P_i) = m(P_i)$.

Por la propiedad b) de la definición de álgebra, las uniones de los elementos de \mathcal{P} pertenecen a \mathcal{a} , pero éstas son los elementos de $a(\mathcal{P})$, así que $a(\mathcal{P}) \subset \mathcal{a}$.

Como

$$\tilde{\mu} \left(\sum_{i \in N} P_i \right) = \sum_{i \in N} \tilde{\mu}(P_i) = \sum_{i \in N} m(P_i)$$

resulta que

$$\tilde{\mu} \left(\sum_{i \in N} P_i \right) = \mu \left(\sum_{i \in N} P_i \right), \quad \forall \sum_{i \in N} P_i \in a(\mathcal{P}).$$

∴ a) el dominio de la medida $\tilde{\mu} : \mathcal{a} \rightarrow [0, 1]$, que contiene a \mathcal{P} , contiene también al álgebra generada por esta partición, $a(\mathcal{P})$;

b) la única *medida extensión* de la medida básica $m : \mathcal{P} \rightarrow (0, 1)$ al dominio más amplio y rico de subconjuntos de X que \mathcal{P} , $a(\mathcal{P})$, y que coincide en sus valores en cada P_i de \mathcal{P} con los de $m : \mathcal{P} \rightarrow (0, 1)$, es $\mu : a(\mathcal{P}) \rightarrow [0, 1]$.

Ejemplos correspondientes a los expuestos anteriormente.

1. Para la medida básica de Bernoulli

$$m_1 : \{\{0\}, \{1\}\} \rightarrow (0, 1) \ni$$

$$m_1(\{1\}) = p, m_1(\{0\}) = 1 - p$$

se tiene que el álgebra de subconjuntos de $X_1 = \{0, 1\}$ generada por \mathcal{P}_1 es

$$a(\mathcal{P}_1) = \{\emptyset, \{0\}, \{1\}, X_1\}$$

y la extensión de la medida básica a este dominio es la *medida normal de Bernoulli*

$$\mu_1 : a(\mathcal{P}_1) \rightarrow [0, 1] \ni$$

$$\mu_1(\emptyset) = 0, \quad \mu_1(\{0\}) = 1 - p, \quad \mu_1(\{1\}) = p$$

y

$$\begin{aligned} \mu_1(X) &= \mu_1(\{0\} + \{1\}) \\ &= m_1(\{0\}) + m_1(\{1\}) \\ &= (1 - p) + p \\ &= 1. \end{aligned}$$

2. El álgebra generada por \mathcal{P} , $a(\mathcal{P})$, es la colección de todas las figuras formadas por uniones de cuadrados pertenecientes a \mathcal{P} : por comodidad se usará la notación siguiente

$$I(i, j) = [i, (i+1) \times [j, (j+1)), \quad i, j = 0, 1, \dots, 9;$$

entonces, cada elemento de $a(\mathcal{P})$ será de la forma

$$\sum_{(i,j) \in N} I(i, j), \quad N \subset \{(i, j) : i, j = 0, 1, \dots, 9\}$$

y su medida

$$\begin{aligned} \mu\left(\sum_{(i,j) \in N} I(i, j)\right) &= \sum_{(i,j) \in N} m(I(i, j)) \\ &= \sum_{(i,j) \in N} 10^{-2} \\ &= \frac{\#(N)}{10^2} \end{aligned}$$

la cual no es otra cosa que el *área* del conjunto

$$\sum_{(i,j) \in N} I(i, j)$$

y, en particular,

$$\begin{aligned} \mu(X) &= \frac{\#\{(i, j) : i, j = 0, 1, \dots, 9\}}{10^2} \\ &= \frac{100}{10^2} \\ &= 1. \end{aligned}$$

Esta medida es un tipo de *medida de Peano y Jordan*, como la considerada en la Sección del Cap. II.

3. Para la medida básica binomial

$$m_n : \mathcal{P}_n \rightarrow (0, 1)$$

se obtiene la *medida (normal) binomial*

$$\mu_n : a(\mathcal{P}_n) \rightarrow [0, 1] \ni$$

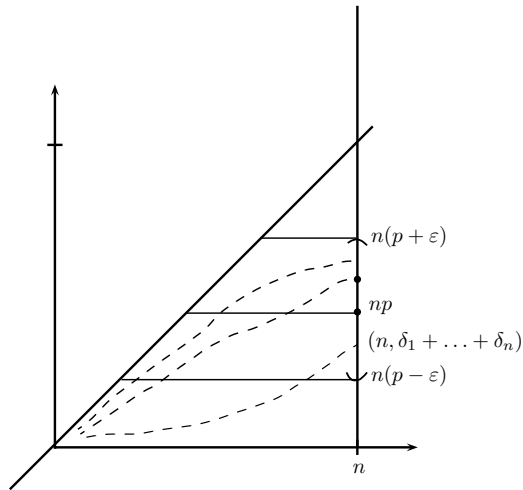
$$\mu_n\left(\sum_{k \in N} P_k\right) = \sum_{k \in N} \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall \sum_{k \in N} P_k \in a(\mathcal{P}_n).$$

Entre la multitud de subconjuntos de X_n que pertenecen al álgebra $a(\mathcal{P}_n)$, los más relevantes son los de la siguiente forma: tomando $\epsilon > 0$, se define el subconjunto de $\{0, 1, \dots, n\}$.

$$N(\epsilon) = \left\{ k : \left| \frac{k}{n} - p \right| < \epsilon \right\}$$

y como las condiciones analíticas que deben cumplir sus elementos pueden expresarse en otras formas equivalentes, se tiene que

$$\begin{aligned} N(\epsilon) &= \left\{ k : p - \epsilon < \frac{k}{n} < p + \epsilon \right\} \\ &= \left\{ k : -n\epsilon < k - np < n\epsilon \right\} \\ &= \left\{ k : n(p - \epsilon) < k < n(p + \epsilon) \right\}. \end{aligned}$$



Entonces el elemento de $a(\mathcal{P}_n)$

$$\sum_{k \in N(\epsilon)} P_k$$

es el conjunto de todas las sucesiones de X_n cuyas proporciones de 1s respecto a n , $\frac{k}{n}$, es *cercana* a p , en menos que ϵ o, equivalentemente, a las sucesiones cuyo número de 1s es *cercano* a np , en relación a n , en menos que $n\epsilon$ o, graficamente, el conjunto de todas las gráficas que partiendo de $(1, 0)$ ó $(1, 1)$ terminan en el segmento $\{n\} \times (n(p - \epsilon), n(p + \epsilon))$, Fig.3.1.

Y las negaciones correspondientes para el conjunto

$$\left(\sum_{k \in N(\epsilon)} P_k \right)^c = \sum_{k \in N(\epsilon)^c} P_k,$$

$$\begin{aligned} N(\epsilon)^c &= \{k : |k - np| \geq \epsilon\} \\ &= \{k : k \geq n(p + \epsilon) \text{ ó } k \leq n(p - \epsilon)\}; \end{aligned}$$

el conjunto de todas las sucesiones de X_n cuyas gráficas parten de los puntos $(1, 0)$ ó $(0, 1)$ y terminan en $\{n\} \times ([0, n(p - \epsilon)] + [n(p + \epsilon), n])$.

Sus respectivas medidas serán entonces

$$\begin{aligned} \mu_n \left(\sum_{k \in N(\epsilon)} P_k \right) &= \sum_{k \in N(\epsilon)} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{n(p-\epsilon) < k < n(p+\epsilon)} \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

y

$$\begin{aligned} \mu_n \left(\left(\sum_{k \in N(\epsilon)} P_k \right)^c \right) &= 1 - \sum_{k \in N(\epsilon)} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k \leq n(p-\epsilon)} \binom{n}{k} p^k (1-p)^{n-k} + \sum_{k \geq n(p+\epsilon)} \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

3.2. De la distribución esencial a la medida normal.

Se tiene una pareja experimento-espacio de resultados con las características C_1 y C_2 , de la cual las distribuciones de los resultados de cada uno de sus colectivos son iguales estadísticamente, $(\mathcal{E}, X; (p_R)_{R \in \mathcal{R}})$; de la familia de regiones \mathcal{R} de X se selecciona una partición $\mathcal{P} = \{P_1, \dots, P_N\}$ de X , con sus constantes $p_{(1)}, \dots, p_{(N)}$ correspondientes a los pedazos seleccionados ($p_{(i)}$ es $\varphi(P_i)$ o "p índice P_i "); la asociación

$$P_k \rightarrow p_{(k)}$$

puede formularse como la *medida normal básica*

$$m : \mathcal{P} \rightarrow (0, 1) \ni$$

$$\begin{aligned} m(P_k) &= \varphi(P_k) \\ &= p_{(k)}, \end{aligned}$$

$$k = 1, \dots, N$$

ya que, además,

$$m(P_1) + \dots + m(P_N) = 1.$$

(La función característica de P_k se denotará por $\delta_{(k)}$; para cada

$$x \in X = \sum_{k=1}^N P_k$$

se tiene

$$\delta_{(1)}(x) + \cdots + \delta_{(N)}(x) = 1;$$

y para cualesquiera n puntos x_1, \dots, x_n de X ,

$$\sum_{i=1}^n \delta_{(1)}(x_i) + \cdots + \sum_{i=1}^n \delta_{(N)}(x_i) = n$$

$$\therefore \frac{1}{n} \sum_{i=1}^n \delta_{(1)}(x_i) + \cdots + \frac{1}{n} \sum_{i=1}^n \delta_{(N)}(x_i) = 1.$$

Y si en particular $x_1, x_2, \dots, x_n, \dots$ es un colectivo de (\mathcal{E}, X) , para cada $n \gg 1$, los números

$$\frac{1}{n} \sum_{i=1}^n \delta_{(k)}(x_i), p_{(k)},$$

para cada $k = 1, \dots, N$, son muy próximos, según las desigualdades de Bernoulli -debido a que $(\mathcal{E}, \{0, 1\}_p), \forall P \in \mathcal{P}$, es un experimento de Bernoulli-, siendo pausable concluir que también

$$p_{(1)} + \cdots + p_{(N)} = 1 \quad ;$$

la argumentación anterior no es una demostración como las usuales en cualquier rama de la matemática, ya que no se está argumentando en una estructura matemática; sólo es un razonamiento convincente).

Esta medida normal básica, independizada del contexto del cual ha sido extraída, es un concepto primordial a partir del cual se construirá, utilizando técnicas de la teoría de la medida de conjuntos, la estructura matemática que se utilizará en el futuro; esta medida normal básica está *obligada, necesariamente*, a cumplir las leyes y consecuencias que las medidas normales básicas formales y que, por lo tanto, igual como se demostró en la sección anterior, genera una y sólo una *medida normal*

$$\mu : a(\mathcal{P}) \rightarrow [0, 1] \ni$$

$$\begin{aligned} \mu \left(\sum_{k \in \mathbb{N}} P_k \right) &= \sum_{k \in \mathbb{N}} m(P_k) \\ &= \sum_{k \in \mathbb{N}} p_{(k)}, \quad \forall \sum_{k \in \mathbb{N}} P_k \in a(\mathcal{P}), \end{aligned}$$

la cual es su extensión a la colección de subconjuntos más rica $a(\mathcal{P})$, el álgebra generada por la partición \mathcal{P} .

Adicionalmente a lo puramente formal, esta medida normal está asociada al fenómeno físico de la igualdad estadística de las distribuciones de resultados de cada colectivo (\mathcal{E}, X) : es una ampliación del conocimiento acerca de la distribución esencial de aquéllas, con más contenido que la simple medida normal básica: el valor de aquella función en A , de $a(\mathcal{P})$, $\mu(A)$, es la medida de *qué tan frecuentemente* resultados en cada colectivo de (\mathcal{E}, X) caen en el subconjunto A de X o de qué tan *frecuentemente* se realiza A en cada sucesión de ejecuciones de \mathcal{E} ; esto es, en términos cuantitativos, para cada colectivo de (\mathcal{E}, X)

Ejecución de \mathcal{E}	1,	2,	...	n ,	...
Resultado	x_1 ,	x_2 ,	...	x_n ,	...

el colectivo correspondiente al experimento de Bernoulli $(\mathcal{E}, \{0, 1\}_A)$

$$\delta_A(x_1), \delta_A(x_2), \dots, \delta_A(x_n), \dots$$

sigue la propiedad (o ley) de los números grandes de Bernoulli,

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_A(x_i) - \mu(A) \right| \ll 1, \quad \forall n \gg 1,$$

dado, también, el *carácter aleatorio* de los resultados o de las realizaciones de A en cada ejecución de \mathcal{E} (a excepción de X , para quien $\mu(X) = 1$ y

$$\frac{1}{n} \sum_{i=1}^n \delta_X(x_i) = 1, \forall n = 1, 2, \dots$$

Una medida normal como ésta, surgida, consecuencia, del fenómeno físico de las regularidades estadísticas, será llamada *distribución*; al parecer esta nominación es más afortunada que su nombre usual, (*medida de*) *probabilidad*, cuya permanencia se sustenta sólo en la tradición pero oculta el fenómeno físico de fondo, la igualdad estadística de las distribuciones de los resultados de los colectivos de una pareja (\mathcal{E}, X) , además de inducir a la mente a pensar a $\mu(A)$ como algo relacionado con la adivinación, la suerte, la fortuna o causando confusión en el estudiante al tratar de encontrar la relación entre éstas y aquel número. Así, al encontrarse la frase " $\mu(A)$ es la (medida de) probabilidad del evento A " lo más aconsejable es quitar esa palabra controvertida, "probabilidad", y pensar o interpretar ese número como una constante física relacionada con el fenómeno dicho o, simplemente, como la medida de A o, más esencial, como el valor de la función en A .

Por último, debe aclararse que la medida normal, la distribución y

$$\mu : a(\mathcal{P}) \rightarrow [0, 1],$$

es una de las posibles descripciones de la distribución esencial en cuestión, fina o basta, dependiendo del número y *tamaño* de los pedazos de la partición \mathcal{P} .

Y, posiblemente, susceptible de ser afinada, sobre todo cuando $\#(X)$ es infinito; no obstante lo cual, muestra o describe de cierta manera la estructura geométrica de esa distribución esencial. Aquí, sin embargo, interesarán sólo

aquellas parejas (\mathcal{E}, X) cuyos espacios de resultados tienen un número finito de elementos y que, por lo tanto, sus particiones siempre serán finitas, lo cual, como se verá, será suficiente para empezar a construir la teoría matemática, la estructura matemática, que representará teóricamente esa problemática acerca de los colectivos de Bernoulli.

3.3. Tesis de correspondencia

Un *espacio de medida normal* es un universo $X \neq \emptyset$ con una medida normal $\mu : \mathcal{a} \rightarrow [0, 1]$, de la cual su dominio \mathcal{a} es un álgebra de subconjuntos de X , que se denotará por el arreglo $(X, \mu : \mathcal{a} \rightarrow [0, 1])$. Un *espacio con distribución* (o *espacio de medida de probabilidad*) es, ni más ni menos, un espacio de medida normal $(X, \mu : \mathcal{a} \rightarrow [0, 1])$, con X , constando de al menos dos elementos; a X se le nombra *espacio de resultados* y a la medida normal *distribución* (o *medida de probabilidad*).

Conceptualmente no hay diferencias entre uno y otro espacio, sólo cambian los nombres de sus componentes. Es de notarse que si se utiliza el nombre de espacio de medida de probabilidad, en su definición no hay ningún elemento que tenga relación con la *probabilidad*, entendiéndose por ésta lo que se quiera.

Sin embargo, el espacio de medida normal, cuyo universo consta de al menos dos elementos, es el concepto fundamental a partir del cual se construye toda la estructura matemática en la cual se representan fidedigna y teóricamente las diversas regularidades de carácter estadístico inherentes a aquellas parejas (\mathcal{E}, X) cuyas distribuciones de resultados de sus colectivos son iguales estadísticamente.

Por tal motivo, cuando un espacio de medida normal $(X, \mu : \mathcal{a} \rightarrow [0, 1])$ es pensado como un espacio con distribución (o como espacio de medida de probabilidad) es usual, como método heurístico, asociarlo mentalmente a una pareja (\mathcal{E}, X) de la cual los resultados de cada uno de sus colectivos

<i>Ejecución de \mathcal{E}</i>	1,	2,	...	n ,	...
<i>Resultado</i>	x_1 ,	x_2 ,	...	x_n ,	...

se *distribuyen aleatoriamente*, por el espacio de resultados X , pero *esencialmente* según la función $\mu : \mathcal{a} \rightarrow [0, 1]$ en el sentido de que el valor *determinado* de ésta en A , subconjunto de X perteneciente a \mathcal{a} , $\mu(A)$, *representa* la medida de qué tan frecuentemente resultados del colectivo *caen* en o *realizan* a A , no determinísticamente sino aleatoriamente, aunque siguiendo, determinísticamente, las *regularidades estadísticas* representadas analíticamente por las desigualdades, condiciones o propiedades (o ley) de los números grandes de Bernoulli:

$$\left| \frac{1}{n} \sum_{i=1}^n \delta_A(x_i) - \mu(A) \right| \ll 1, \forall n \gg 1.$$

O, equivalentemente, el número $\mu(A)$ puede ser imaginado como la pendiente de la recta que pasa por el origen y es *perseguida* por las gráficas de los

correspondientes colectivos de Bernoulli

$$\delta_A(x_1), \delta_A(x_2), \dots, \delta_A(x_n), \dots$$

del experimento de Bernoulli $(\mathcal{E}, \{0, 1\}_A)$ asociado a (\mathcal{E}, X) respecto o en relación a A , manteniéndose determinísticamente, las distancias verticales, aleatorias entre esas gráficas y la recta $y = \mu(A)x$, según las desigualdades de Bernoulli,

$$\left| \sum_{i=1}^n \delta_A(x_i) - n\mu(A) \right| \ll n, \forall n \gg 1.$$

(Desde luego que puede existir A en \mathcal{a} , $A \neq \emptyset$ de $\mu(A) = 0$ o, equivalentemente, $A^c \in \mathcal{a}$, $A^c \neq X$ con $\mu(A^c) = 1$, situaciones que no fueron consideradas en el Cap. II, sin embargo esto no cambia la interpretación frecuencial de $\mu(A)$ o $\mu(A^c)$: el colectivo, su gráfica,

$$\delta_A(x_1), \delta_A(x_2), \dots, \delta_A(x_n), \dots$$

persigue a la $y = 0x$ según

$$\sum_{i=1}^n \delta_A(x_i) \ll n, \forall n \gg 1 :$$

rara pero muy rara vez se realiza A o cae algún x_i en A ; y la gráfica del colectivo

$$\delta_{A^c}(x_1), \delta_{A^c}(x_2), \dots, \delta_{A^c}(x_n), \dots$$

persigue a la recta $y = x$ según

$$0 \leq n - \sum_{i=1}^n \delta_{A^c}(x_i) \ll n, \forall n \gg 1 :$$

correspondientemente a lo anterior, rara, pero muy rara vez no se realiza A^c o no cae algún x_i en A^c).

Más aún, dado un espacio de medida normal $(X, \mu : \mathcal{a} \rightarrow [0, 1])$ concreto no es raro que pueda elaborarse un programa de computadora que con una buena tabla de dígitos extraídos al azar (Por ejemplo, las tablas compiladas por M. G. Kendall y B. Babington Smith *Tables of random sampling numbers*, Cambridge, University Press, 1971. Tales tablas se encuentran en la dirección electrónica www.dynamics.unam.edu/LaboratorioAleatoriedad) a su disposición, genere colectivos de resultados en X , los grafique en éste siguiendo en buena medida la distribución $\mu : \mathcal{a} \rightarrow [0, 1]$; genere y grafique colectivos de Bernoulli correspondientes a $(\mathcal{E}, \{0, 1\}_A)$, $A \in \mathcal{a}$, etc. Ejemplos concretos de estas realizaciones se encuentran en la Tesis profesional *Ampliación de tablas de dígitos al azar* de E.E. Arnoldo O., Facultad de Ciencias de la U.N.A.M., 2005).

Las conclusiones de la sección anterior y las anteriores de la presente sección, son testimonios suficientes para establecer el principio heurístico

Tesis de correspondencia

Toda pareja (\mathcal{E}, X) cuyas distribuciones de resultados de sus colectivos son iguales estadísticamente, tiene asociada un espacio de medida normal

$$(X, \mu : \mathcal{a} \rightarrow [0, 1]),$$

cuya medida normal describe a la distribución esencial de aquélla;

Todo espacio de medida normal $(X, \mu : \mathcal{a} \rightarrow [0, 1])$ tiene asociado una pareja (\mathcal{E}, X) , cuyas distribuciones por X de los resultados de cada uno de sus colectivos, son descritas esencialmente por la función $\mu : \mathcal{a} \rightarrow [0, 1]$.

Capítulo 4

Experimentos de Bernoulli truncos

Entre la multitud de parejas (\mathcal{E}, X) cuyas distribuciones de resultados de cada uno de sus colectivos son iguales estadísticamente entre sí, y que en particular interesan en este libro, están aquellas que se denominarán *experimentos de Bernoulli truncos*: se tienen un experimento de Bernoulli $(\mathcal{E}, X_1 = \{0, 1\})_p$ y un número natural n , fijo; se considera el experimento (u observación) \mathcal{E}_n , el cual la ejecución de \mathcal{E} n veces, y de quien interesan las sucesiones de n términos, cada uno en $X_1 = \{0, 1\}$, correspondientes a aquéllas, esto es, los resultados del tipo $\delta_1, \delta_2, \dots, \delta_n$: δ_1 es el resultado, en X_1 , de la ejecución 1 de \mathcal{E}, \dots y δ_n es el resultado, en X_1 , de la ejecución n de \mathcal{E} . Estas sucesiones se representarán en la forma de vectores de n componentes, $(\delta_1, \dots, \delta_n)$ y que por ser tramos de un colectivo de Bernoulli de $(\mathcal{E}, X_1)_p$ serán llamados *colectivos de Bernoulli truncos* (de rango n).

El conjunto de estos resultados, la cateroría de resultados que interesan al efectuarse \mathcal{E}_n , los colectivos de Bernoulli truncos de rango n , se representarán como elementos del conjunto

$$X_n = \{(\delta_1, \dots, \delta_n) : \delta_i = 0 \text{ ó } 1, i = 1, 2, \dots, n\} ,$$

el cual se tomará como el *espacio de resultados* de \mathcal{E}_n , y como representación geométrica del conjunto X_n la descrita ya en anteriores capítulos: cada sucesión de n términos $(\delta_1, \dots, \delta_n)$, $\delta_i = 0$ ó 1 , está en correspondencia biunívoca con cada sucesión de puntos $(k, \delta_1 + \dots + \delta_k)$, $k = 1, 2, \dots, n$, de \mathbb{R}^2 .

Aquí se tomarán las repeticiones de \mathcal{E}_n *independientes* unas de las otras: una vez efectuado \mathcal{E}_n , efectuándose \mathcal{E} n veces, se efectúa \mathcal{E} otras n veces, sin que en éstas se incluya alguna de las n anteriores.

Esta condición se aclara por medio de una sucesión de ejecuciones (independientes) de \mathcal{E} , del colectivo obtenido,

$$x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}, x_{2n+1}, \dots, x_{3n}, \dots$$

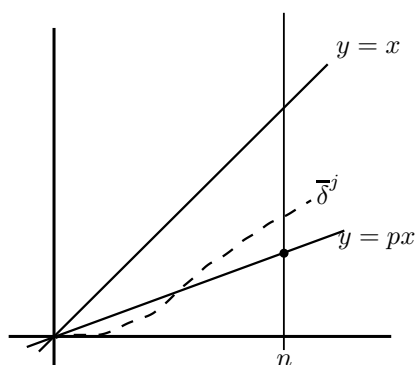
y considerando la ejecución 1 de \mathcal{E}_n aquella que dio los resultados x_{n+1}, \dots, x_{2n} , y así sucesivamente. Contrastando: tómesese la ejecución 1 de \mathcal{E}_n como antes, pero la ejecución 2 de \mathcal{E}_n , como las ejecuciones 2, 3, \dots y $n+1$ de \mathcal{E} , obteniéndose en este caso los resultados x_2, x_3, \dots, x_{n+1} ; la ejecución 3 de \mathcal{E}_n como las ejecuciones 3, 2, \dots y $n+2$ de \mathcal{E} , obteniéndose los resultados x_3, x_4, \dots, x_{n+3} , etc.; estas ejecuciones de \mathcal{E}_n como se ve, son dependientes unas de otras.

Esta nueva clase de parejas, *experimento \mathcal{E}_n - espacio de resultados X_n* , tomadas, surgidas, construidas o consideradas a partir de un experimento de Bernoulli $(\mathcal{E}, X = \{0, 1\})_p$ y de un número natural n , hereda las características C1 y C2 de aquel, razón por la cual, como ya se convino, se escribirá en la forma (\mathcal{E}_n, X_n) y se denominará *experimento de Bernoulli trunco de rango n* .

La propiedad más relevante que se observa, o se preve, de cada colectivo de (\mathcal{E}_n, X_n)

$$\begin{array}{l} \text{Ejecución de } \mathcal{E}_n \quad 1, \quad 2, \quad \dots, \quad N, \dots \\ \text{Resultado} \quad \quad \quad \bar{\delta}^1, \quad \bar{\delta}^2, \quad \dots, \quad \bar{\delta}^N, \dots, \end{array}$$

(con $\bar{\delta}^i = (\delta_1^i, \dots, \delta_n^i)$, resultado de la *repetición i* de \mathcal{E}_n y δ_i^j el resultado de la *repetición i* de \mathcal{E} en la tanda j de n *repeticiones* de \mathcal{E}) es cuando n es grande (p. ej. $n \geq 500$); sus resultados se distribuyen aleatoriamente por X_n de forma muy singular, pero la gran mayoría se *aglomeran alrededor* del segmento de la recta $y = px$ que va del punto $(0, 0)$ al (n, pn) de \mathbb{R}^2 , hecho que no es de extrañar en vista de que cada $\bar{\delta}^j$ es un tramo de n términos de un colectivo de Bernoulli de $(\mathcal{E}, X_1)_p$ (un colectivo de Bernoulli trunco de rango n), suficientes para manifestarse el síndrome de Bernoulli de sus colectivos, acentuándose este fenómeno al tomar valores de n más y más grandes.



Aunque esta aglomeración se da cuando n es *pequeña*, por ejemplo $n \leq 10$, además sucede que cada elemento de X_n aparece en cada colectivo de (\mathcal{E}, X_n) , lo que no acontece incluso cuando n no es grande, por ejemplo $n = 50$; en este caso el número de elementos de X_{50} es $\#(X_{50}) = 2^{50} = 10^{15}$, número inmenso; pero aun si se repitiera \mathcal{E}_{50} 10^{15} veces o más (con fotones, con electrones; con moléculas de un gas) es prácticamente imposible que todos los elementos de X_{50} *salgan* en el correspondiente colectivo, es más, para cada $(\delta_1, \dots, \delta_{50})$, que se

obtenga al efectuarse \mathcal{E}_{50} , puede asegurarse, incluso apostando la vida propia que

- i) los 0s y 1s aparecen *desordenados*,
- ii) *cae cerca* de la sucesión de puntos $(1, p), (2, 2p), \dots, (50, 50p)$ de la recta $y = px$.

4.1. Distribución básica de (\mathcal{E}_n, X_n)

Una descripción de la distribución esencial de (\mathcal{E}_n, X_n) , suficientemente elocuente, la mejor que se tiene sin ser oneroso sino todo lo contrario y que coincide sorprendente e increíblemente con los datos numéricos experimentales es la *distribución (o medida normal) binomial*, Cap III, que se construye a partir de la *medida básica binomial*

$$m_n : \mathcal{P}_n \rightarrow (0, 1),$$

de la cual su dominio $\mathcal{P}_n = \{P_0, P_1, \dots, P_n\}$ es la partición de X_n , ya presentada anteriormente,

$$P_k = \{(\delta_1, \dots, \delta_n) : \delta_1 + \dots + \delta_n = k\}, \quad k = 0, 1, \dots, n$$

y el valor de la función en cada pedazo P_k , la medida de P_k , es

$$m_n(P_k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Los valores $m_n(P_k)$, $k = 0, 1, \dots, n$, describen cuantitativa cómo se *reparten, esencialmente, salvo accidentes aleatorios*, los resultados de cada colectivo de (\mathcal{E}_n, X_n) entre los subconjuntos P_0, P_1, \dots, P_n de X_n (pero no cómo se distribuyen por cada P_k). Así, para cada colectivo

<i>Ejecución de \mathcal{E}_n</i>	1,	2,	...	N,	...
<i>Resultado</i>	$\bar{\delta}^1$,	$\bar{\delta}^2$,	...	$\bar{\delta}^N$,	...

el colectivo correspondiente a las realizaciones o no de P_k

$$\delta_{(k)}(\bar{\delta}^1), \delta_{(k)}(\bar{\delta}^2), \dots, \delta_{(k)}(\bar{\delta}^N), \dots$$

es un colectivo de Bernoulli del experimento de Bernoulli $(\mathcal{E}_n, \{0, 1\}_{(k)})$ correspondiente a (\mathcal{E}_n, X_n) y el pedazo $P_k \in \mathcal{P}_n$, que entre otras propiedades, sigue la condición o propiedad (o ley) de los números grandes de Bernoulli

$$\left| \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - N m_n(P_k) \right| \ll N, \quad \forall N \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{N} \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - m_n(P_k) \right| \ll 1, \forall N \gg 1,$$

siendo, entonces,

$$m_n(P_k) = \binom{n}{k} p^k (1-p)^{n-k}$$

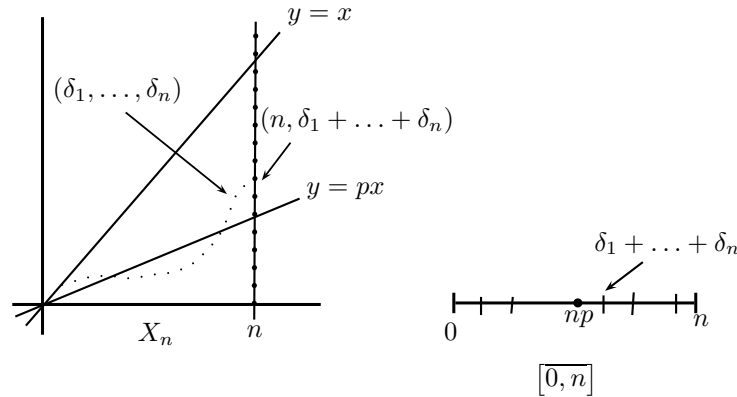
la medida de qué tan frecuentemente resultados del colectivo *caen*, salvo accidentes aleatorios, en el subconjunto P_k de X_n o, equivalentemente, la medida de qué tan frecuentemente se realiza, salvo accidentes aleatorios, el pedazo P_k en cada sucesión de repeticiones de \mathcal{E}_n . Y esto para cada $P_k \in \mathcal{P}_n$.

4.2. Experimento $(\mathcal{E}_n, \overline{[0, n]})$

Otra pareja *experimento - espacio de resultados* con las características $C1$ y $C2$ para la cual las distribuciones de los resultados de sus colectivos son iguales estadísticamente es la derivada de un experimento de Bernoulli trunco (\mathcal{E}_n, X_n) (que a su vez lo es de un experimento de Bernoulli $(\mathcal{E}, X)_p$) : cada vez que se efectúa \mathcal{E}_n , se obtiene un resultado aleatorio $(\delta_1, \dots, \delta_n)$, en X_n , pero, además, el número entero $\delta_1 + \dots + \delta_n$, también aleatorio, respecto a las condiciones en que se efectúa \mathcal{E}_n ; este entero es el número de 1s que se obtiene en n ejecuciones de \mathcal{E} , el número de 1s en aquella sucesión o el número, de esas n ejecuciones, en las cuales *salió* 1.

Éstos son los resultados que ahora interesan al efectuarse \mathcal{E}_n , obteniéndose así la pareja \mathcal{E}_n con la categoría de resultados que interesa, el conjunto de enteros $\{0, 1, 2, \dots, n\}$, el cual se denotará por $\overline{[0, n]}$; esta nueva pareja *experimento - espacio de resultados* $\overline{[0, n]}$ hereda las características $C1$ y $C2$ del experimento de Bernoulli trunco (\mathcal{E}_n, X_n) y que, por lo ya convenido, se escribirá en la forma $(\mathcal{E}_n, \overline{[0, n]})$.

Este nuevo espacio de resultados, $\overline{[0, n]}$, es una representación simple y equivalente del subconjunto $\{n\} \times \{0, 1, 2, \dots, n\}$ de \mathbb{R}^2 , el cual es el conjunto de puntos extremos, derechos, $(n, \delta_1 + \dots + \delta_n)$ de la representación geométrica que se ha venido usando para los elementos de X_n .

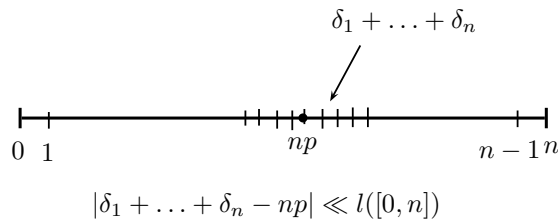


Puede obtenerse una primera idea de cómo se distribuyen por $\overline{[0, n]}$ los resultados de un colectivo $(\mathcal{E}_n, \overline{[0, n]})$ cuando n es grande (por ejemplo, $n \approx 1000$); como esta pareja proviene del experimento de Bernoulli $(\mathcal{E}, X_1)_p$, cada uno de sus colectivos $\delta_1, \delta_2, \dots, \delta_n, \dots$ tiene el síndrome de Bernoulli, que entre sus propiedades está la de que la distancia vertical entre sus gráficas y la recta $y = px$ sigue las desigualdades de Bernoulli

$$\left| \sum_{i=1}^n \delta_i - np \right| \ll n, \forall n \gg 1;$$

y como $(\delta_1, \dots, \delta_n)$ es un resultado del experimento de Bernoulli trunco de rango n , (\mathcal{E}_n, X_n) , $\delta_1 + \dots + \delta_n$, en $\overline{[0, n]}$, es un resultado de la pareja $(\mathcal{E}_n, \overline{[0, n]})$, el cual dista *poco* del real np , perteneciente al intervalo cerrado $[0, n]$ de \mathbb{R} , en el sentido de que la distancia entre los reales $\delta_1 + \dots + \delta_n$ y np en $[0, n]$ es *pequeña* respecto a la longitud n de este intervalo.

Por tal hecho, los resultados del experimento $(\mathcal{E}_n, \overline{[0, n]})$, para n grande, se acumulan en los enteros alrededor de np , en el intervalo $[0, n]$.



Para tener una idea de esos *poco*, *pequeña*, "se acumulan..." puede usarse la ley basta del logaritmo iterado: para $n = 1000$, por ejemplo,

$$|\delta_1 + \dots + \delta_{1000} - 1000p| < \sqrt{\frac{1000}{2} \ln \ln 1000} = 31.08 \dots$$

$$\therefore -31.08 \dots < \delta_1 + \dots + \delta_{1000} - 1000p < 31.08 \dots :$$

los resultados de $(\mathcal{E}_{1000}, \overline{[0, 1000]})$ se acumulan, *salvo accidentes aleatorios*, en la vecindad de np

$$(1000p - 31.08 \dots, 1000p + 31.08 \dots),$$

cuyo radio, $31.08 \dots$, es muy pequeño respecto a $l([0, 1000]) = 1000$.

Para n pequeña ($n \leq 15$ o 20 , por ejemplo) sucede casi lo mismo aunque en estos casos *casi* hay una correspondencia biunívoca entre los resultados de $(\mathcal{E}_n, \overline{[0, n]})$ y los elementos de $\overline{[0, n]}$, acentuándose esto cuando n es más pequeña, pero siempre dándose una concentración notable, de esos resultados, alrededor de np (sin embargo no deja de causar suspicacia que en 15 extracciones de bolas, en ralación al ejemplo 1 del Cap. 1, se obtenga sólo bolas blancas o que en 15 volados sucesivos efectuados según las condiciones del merengero, se obtengan 15 águilas seguidas)

Bien ¿pero cómo se distribuyen por $\overline{[0, n]}$ los resultados de cada colectivo $(\mathcal{E}_n, \overline{[0, n]})$?

Una descripción de esa distribución puede obtenerse de la *medida básica* $m_n : \mathcal{P}_n \rightarrow (0, 1)$ de los resultados de cada colectivo de (\mathcal{E}_n, X_n) :

i) se emplea la partición $\mathcal{P}_n^{(1)} = \{\{0\}, \{1\}, \dots, \{n\}\}$ del espacio de resultados $\overline{[0, n]}$;

ii) se utiliza que la condición

$$\bar{\delta} = (\delta_1, \dots, \delta_n) \in P_k \Leftrightarrow \sum_{i=1}^n \delta_i = k$$

es equivalente a la condición

$$\bar{\delta} \in P_k \Leftrightarrow \sum_{i=1}^n \delta_i \in \{k\};$$

iii) usando la función característica del pedazo $\{k\}$ de $\overline{[0, n]}$

$$\delta_{\{k\}} : \overline{[0, n]} \rightarrow \{0, 1\} \quad \ni$$

$$\delta_{\{k\}}(i) = \begin{cases} 1 & \text{si } i \in \{k\} \quad (\Leftrightarrow i = k) \\ 0 & \text{si } i \notin \{k\} \quad (\Leftrightarrow i \neq k) \end{cases}$$

así se llega a que

$$\delta_{(k)}(\bar{\delta}) = \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i \right),$$

y esto para cada $k = 0, 1, \dots, n$ y cada $\bar{\delta} = (\delta_1, \dots, \delta_n) \in X_n$;

iv) y ahora, en relación a una sucesión de ejecuciones de \mathcal{E}_n se obtienen los colectivos

$$\begin{array}{llll} \text{Ejecución de } \mathcal{E}_n & 1, & 2, & \dots, \quad N, \dots \\ \text{Resultado en } X_n & \bar{\delta}^1, & \bar{\delta}^2, & \dots, \quad \bar{\delta}^N, \dots \\ \text{Resultado en } \overline{[0, n]} & \sum_{i=1}^n \delta_i^1, & \sum_{i=1}^n \delta_i^2, & \dots, \quad \sum_{i=1}^n \delta_i^N, \dots, \end{array}$$

siendo $\bar{\delta}^j = (\delta_1^j, \dots, \delta_n^j) \forall j = 1, 2, \dots, N, \dots$, como

a) para cada $k = 0, 1, \dots, n$ y $N = 1, 2, \dots$ sucede que

$$\sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) = \sum_{j=1}^N \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^j \right)$$

y

b) substituyendo en las desigualdades de Bernoulli, en relación al colectivo de (\mathcal{E}_n, X_n) ,

$$\left| \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - N \binom{n}{k} p^k (1-p)^{n-k} \right| \ll N, \forall N \gg 1$$

las sumas que aparecen en éstas por las sumas de arriba, a los que son iguales, se obtiene que

$$\left| \sum_{j=1}^N \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^j \right) - N \binom{n}{k} p^k (1-p)^{n-k} \right| \ll N, \forall N \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{N} \sum_{j=1}^N \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^j \right) - \binom{n}{k} p^k (1-p)^{n-k} \right| \ll 1, \forall N \gg 1,$$

para el colectivo de $(\mathcal{E}_n, \overline{[0, n]})$, y para cada $\{k\} \in \mathcal{P}_n^{(1)}$.

Así pues, la constante física que asocia a cada subconjunto $\{k\}$, en relación al comportamiento de los resultados de los colectivos de la pareja $(\mathcal{E}_n, \overline{[0, n]})$, es el número $\binom{n}{k} p^k (1-p)^{n-k}$, asociación que origina la función

$$\begin{aligned} m_n^{(1)} : \mathcal{P}_n^{(1)} &\rightarrow (0, 1) \ni \\ m_n^{(1)}(\{k\}) &= \binom{n}{k} p^k (1-p)^{n-k}, \forall \{k\} \in \mathcal{P}_n^{(1)}, \end{aligned}$$

siendo cada una de estas constantes físicas la medida de qué tan frecuentemente resultados de cada colectivo de esta pareja *caen* en la región $\{k\}$ de $\overline{[0, n]}$ o, equivalentemente, la medida de qué tan frecuentemente se realiza el pedazo $\{k\}$ en cada sucesión de ejecuciones de \mathcal{E}_n , tal y como lo expresan estas últimas desigualdades de Bernoulli. Consecuentemente, los resultados de cada colectivo de la pareja $(\mathcal{E}_n, \overline{[0, n]})$ se distribuyen, *esencialmente*, entre los subconjuntos de $\overline{[0, n]}$ pertenecientes a $\mathcal{P}_n^{(1)}$ según la *medida normal básica*

$$m_n^{(1)} : \mathcal{P}_n^{(1)} \rightarrow (0, 1),$$

salvo accidentes aleatorios, conforme a esa condición, propiedad (o ley) de los números grandes de Bernoulli.

4.3. Pareja $(\mathcal{E}_n, [0, 1]_n)$.

Otra pareja *experimento-espacio de resultados* con las características $C1$ y $C2$ y de la cual las distribuciones de los resultados de sus colectivos son iguales estadísticamente, proviene del siguiente experimento: de un experimento de Bernoulli $(\mathcal{E}, X_1 = \{0, 1\})_p$ y para cierto número natural n , \mathcal{E} se efectúa n veces, obteniéndose los correspondientes resultados $\delta_1, \delta_2, \dots, \delta_n$ en $X_1 = \{0, 1\}$, pero también el número de 1s o el número de esas ejecuciones en la que *salió* el 1, $\delta_1 + \dots + \delta_n$, y la proporción de éstos respecto a n

$$\frac{\delta_1 + \dots + \delta_n}{n};$$

estas proporciones son los resultados que ahora interesan cada vez que se efectúan n ejecuciones de \mathcal{E} . Estos números, proporciones, son aleatorios en cuanto a resultados de esas n ejecuciones de \mathcal{E} , razón por la cual interesa saber cómo se distribuyen por el intervalo $[0, 1]$ de \mathbb{R} .

Inscribiendo esta problemática en el contexto del experimento de Bernoulli trunco de rango n , $(\mathcal{E}_n, X_n; (m_n(P_k))_{0 \leq k \leq n})$, del experimento de Bernoulli de arriba o, si se quiere, más directamente, en el de la pareja derivada de ésta, $(\mathcal{E}_n, \overline{[0, n]}; (m_n^{(1)}(\{k\})_{0 \leq k \leq n}))$, queda transcrita a los términos siguientes

a) \mathcal{E} se efectúa n veces es, ni más ni menos que se efectúa \mathcal{E}_n

y

b) el espacio de resultados que interesa es el conjunto de números $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$, el cual se denotará por $[0, 1]_n$, ($[0, 1]_n \subset [0, 1]$) obteniéndose así la nueva pareja *experimento \mathcal{E}_n -espacio de resultados* $[0, 1]_n$ y que, por herencia de sus predecesores en cuanto a sus características $C1$ y $C2$, se denotará como $(\mathcal{E}_n, [0, 1]_n)$.

¿Y cómo se distribuyen por el espacio $[0, 1]_n$ los resultados de sus colectivos? Para cada sucesión de ejecuciones se obtienen los colectivos

<i>Ejecución de \mathcal{E}_n</i>	1,	2,	...	N, \dots
<i>Resultado en X_n</i>	$\bar{\delta}^1,$	$\bar{\delta}^2,$...	$\bar{\delta}^N, \dots$
<i>Resultado en $\overline{[0, n]}$</i>	$\sum_{i=1}^n \delta_i^1,$	$\sum_{i=1}^n \delta_i^2,$...	$\sum_{i=1}^n \delta_i^N, \dots$
<i>Resultado en $[0, 1]_n$</i>	$\frac{1}{n} \sum_{i=1}^n \delta_i^1,$	$\frac{1}{n} \sum_{i=1}^n \delta_i^2,$...	$\frac{1}{n} \sum_{i=1}^n \delta_i^N, \dots$

(Con $\bar{\delta}^j = (\delta_1^j, \dots, \delta_n^j) \forall j = 1, 2, \dots, N, \dots$), de los cuales ya se sabe cómo se distribuyen los dos primeros por X_n y por $\overline{[0, n]}$, respectivamente.

Ahora, con la partición $\mathcal{P}_n^{(2)} = \{\{0\}, \{\frac{1}{n}\}, \{\frac{2}{n}\}, \dots, \{1\}\}$ de $[0, 1]_n$ se tiene que para $k = 0, 1, \dots$ ó n ,

$$\bar{\delta} = (\delta_1, \dots, \delta_n) \in P_k \Leftrightarrow \delta_1 + \dots + \delta_n \in \{k\} \Leftrightarrow \frac{\delta_1 + \dots + \delta_n}{n} \in \left\{ \frac{k}{n} \right\}$$

para cada $\bar{\delta} \in X_n$.

Y utilizando la función característica de cada subconjunto $\{\frac{k}{n}\}$ de $[0, 1]_n$

$$\begin{aligned} \delta_{\{\frac{k}{n}\}} : [0, 1]_n &\rightarrow (0, 1) \ni \\ \delta_{\{\frac{k}{n}\}} \left(\frac{l}{n} \right) &= \begin{cases} 1 & \text{si } \frac{l}{n} \in \left\{ \frac{k}{n} \right\} \quad (\Leftrightarrow l = k) \\ 0 & \text{si } \frac{l}{n} \notin \left\{ \frac{k}{n} \right\} \quad (\Leftrightarrow l \neq k) \end{cases} \end{aligned}$$

se obtiene que

$$\delta_{(k)}(\bar{\delta}) = \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i \right) = \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i \right)$$

con lo que, en relación a los resultados de los colectivos anteriores, se llega a

$$\sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) = \sum_{j=1}^N \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^j \right) = \sum_{j=1}^N \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^j \right);$$

entonces

$$\begin{aligned} \left| \sum_{j=1}^N \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^j \right) - N \binom{n}{k} p^k (1-p)^{n-k} \right| &= \\ \left| \sum_{j=1}^N \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^j \right) - N \binom{n}{k} p^k (1-p)^{n-k} \right| &= \\ \left| \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - N \binom{n}{k} p^k (1-p)^{n-k} \right|, & \end{aligned}$$

y dado que los números de la derecha en estas igualdades son muy pequeñas respecto a $N \gg 1$, se obtiene que

$$\left| \sum_{j=1}^N \delta_{\left\{\frac{k}{n}\right\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^j \right) - N \binom{n}{k} p^k (1-p)^{n-k} \right| \ll N, \forall N \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{N} \sum_{j=1}^N \delta_{\left\{\frac{k}{n}\right\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^j \right) - \binom{n}{k} p^k (1-p)^{n-k} \right| \ll 1, \forall N \gg 1,$$

y esto para cada $\left\{\frac{k}{n}\right\} \in \mathcal{P}_n^{(2)}$.

Una lectura de estas últimas desigualdades es que el número

$$\binom{n}{k} p^k (1-p)^{n-k}$$

es la medida de cuán frecuentemente resultados de cada colectivo de $(\mathcal{E}_n, [0, 1]_n)$ caen en la región, pedazo $\left\{\frac{k}{n}\right\}$ del espacio de resultados $[0, 1]_n$ o, equivalentemente, que tan frecuentemente se realiza el subconjunto $\left\{\frac{k}{n}\right\}$ en cada sucesión de ejecuciones de \mathcal{E}_n . En el contexto de los colectivos de $(\mathcal{E}_n, [0, 1]_n)$, tal número es la constante física asociada siempre al subconjunto $\left\{\frac{k}{n}\right\}$, obteniéndose así la *medida normal básica*

$$m_n^{(2)} : \mathcal{P}_n^{(2)} \rightarrow (0, 1) \ni$$

$$m_n^{(2)} \left(\left\{ \left\{ \frac{k}{n} \right\} \right\} \right) = \binom{n}{k} p^k (1-p)^{n-k}, \forall \left\{ \frac{k}{n} \right\} \in \mathcal{P}_n^{(2)},$$

que describe (o rige, según se quiera ver) la *distribución esencial* de los resultados de cada colectivo de $(\mathcal{E}_n, [0, 1]_n)$ entre los pedazos de $[0, 1]_n$ pertenecientes a $\mathcal{P}_n^{(2)}$, *salvo accidentes aleatorios*, según aquellas desigualdades de Bernoulli.

Es claro que estas distribuciones de resultados, iguales estadísticamente, son las mismas que las distribuciones de los resultados de cada colectivo de $(\mathcal{E}_n, \overline{[0, n]})$, sólo que en *pequeño* ya que lo único que se ha hecho es alterar la escala de los espacios de resultados, mientras se mantiene la vertical

$$m_n^{(1)}(\{k\}) = m_n^{(2)} \left(\left\{ \frac{k}{n} \right\} \right).$$

Cabe añadir, por último, que cada pareja $(\mathcal{E}_n, \{0, 1\}_{\frac{k}{n}})$ asociada a $(\mathcal{E}_n, [0, 1]_n)$ es un experimento de Bernoulli, pues cada colectivo

$$\delta_{\left\{\frac{k}{n}\right\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^1 \right), \delta_{\left\{\frac{k}{n}\right\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^2 \right), \dots, \delta_{\left\{\frac{k}{n}\right\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^N \right), \dots,$$

debido

a) a las igualdades

$$\delta_{(k)}(\bar{\delta}^j) = \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i \right) = \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i \right) ;$$

b) a que cada pareja $(\mathcal{E}_n, \{0, 1\}_{(k)})$ asociada a (\mathcal{E}_n, X_n) , correspondiente al pedazo P_k , perteneciente a \mathcal{P}_n de X_n , es un experimento de Bernoulli y

c) a que los colectivos

$$\begin{array}{ccccccc} \delta_{\{k\}}(\bar{\delta}^1), & \delta_{\{k\}}(\bar{\delta}^2), & \dots, & \delta_{\{k\}}(\bar{\delta}^N), & \dots \\ \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^1 \right), & \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^2 \right), & \dots, & \delta_{\{k\}} \left(\sum_{i=1}^n \delta_i^N \right), & \dots \\ \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^1 \right), & \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^2 \right), & \dots, & \delta_{\{\frac{k}{n}\}} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^N \right), & \dots \end{array}$$

son idénticos término a término, y siendo, entonces, la recta que *persiguen* sus gráficas

$$y = \binom{n}{k} p^k (1-p)^{n-k} x.$$

4.4. Comportamiento de la densidad binomial $b_n(k)$.

Dado que las distribuciones de los resultados de los colectivos de las parejas (\mathcal{E}_n, X_n) , $(\mathcal{E}_n, \overline{[0, n]})$ y $(\mathcal{E}_n, [0, 1]_n)$ se acumulan en los subconjuntos, de sus espacios de resultados, pertenecientes a las correspondientes particiones, $\mathcal{P}_n, \mathcal{P}_n^{(1)}, \mathcal{P}_n^{(2)}$ según lo rige o lo describe, *esencial* y cuantitativamente la densidad (de dominio discreto) binomial

$$b_n(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n,$$

lo cual no deja de ser paradójico, es conveniente saber más acerca de su comportamiento, concretamente de la forma de su gráfica, con lo cual, a su vez, se tendrá el conocimiento *esencial* y cuantitativo de aquellas acumulaciones; en términos generales la gráfica alcanza un valor máximo en cierto k_0 , y extraordinariamente en dos k_0 y $k_0 + 1$, de $\overline{[0, n]}$ pero siendo estrictamente creciente antes de k_0 y decreciente estrictamente después de k_0 , o de $k_0 + 1$; más detalladamente: $b_n(k)$ es

a) *creciente* $\Leftrightarrow \frac{b_n(k+1)}{b_n(k)} > 1 \Leftrightarrow k < np - (1-p)$.

y

b) *decreciente* $\Leftrightarrow \frac{b_n(k+1)}{b_n(k)} < 1 \Leftrightarrow k > np - (1 - p)$;

y si

i) $np - (1 - p)$ no es número natural, existe sólo un número natural k_0 tal que

$$np - (1 - p) < k_0 < np + p,$$

puesto que $np - (1 - p) + 1 = np + p$; así que

$$(1 - p)^n = b_n(0) < b_n(1) < \dots < b_n(k_0 - 1) < b_n(k_0),$$

debido a a). y debido a b)

$$b_n(k_0) > b_n(k_0 + 1) > \dots > b_n(n - 1) > b_n(n) = p^n$$

pues $k_0 + 1 > np + p$.

Pero si por casualidad

ii) $np - (1 - p)$ es número natural, entonces

$$k_0 = np - (1 - p), \quad k_0 + 1 = np + p,$$

que con las desigualdades de a) y b), se obtiene

$$b_n(0) < \dots < b_n(k_0 - 1) < b_n(k_0) = b_n(k_0 + 1) > b_n(k_0 + 2) > \dots > b_n(n)$$

A manera de resumen: para $b_n(k) = \binom{n}{k} p^k (1 - p)^{n-k}$, con $0 < p < 1$, existe k_0 , tal que, $np - (1 - p) \leq k_0 \leq np + 1$ y $(1 - p)^n = b_n(0) < \dots < b_n(k_0 - 1) < b_n(k_0) \geq b_n(k_0 + 1) > b_n(k_0 + 2) > \dots > b_n(n) = p^n$,

Por otro lado, aquí cabe dar una explicación del porque se afirma que esta densidad binomial describe cuantitativamente la repartición de los resultados de los colectivos de de (\mathcal{E}_n, X_n) por los subconjuntos P_0, P_1, \dots, P_n de X_n , *esencialmente*.

En diversos tratados introductorios de la teoría matemática de la probabilidad se presentan “demostraciones” de tal afirmación, en las cuales, por medio de argumentos *ad hoc*, se combinan las supuestas *probabilidades* (concepto que no se define) de los resultados de las ejecuciones independientes de \mathcal{E} . Pero cuando uno conoce ese fenómeno físico acerca de la igualdad estadística tan insistentemente tomado en cuenta aquí, tal tipo de “demostraciones” dejan un repugnante sabor de incredulidad.

Sin embargo, si uno se pone a experimentar con diversos experimentos de Bernoulli $(\mathcal{E}, X_1 = \{0, 1\})_p$ y diferentes experimentos de Bernoulli truncos (\mathcal{E}, X_n) , para diferentes rangos n , de aquel, efectuando un buen número de corridas de N ejecuciones sucesivas de \mathcal{E}_n , con N siempre lo más grande posible, registrando los correspondientes resultados, contando cuántos *caen* en cada pedazo perteneciente a la partición correspondiente, se encuentra se verifica que hay una aproximación sorprendente, tal y como lo describen las desigualdades

de Bernoulli, entre los números obtenidos por la experimentación y los números dados por la densidad binomial $(b_n(k))_{k=0}^n$ respectiva.

Entonces, es a partir de este trabajo experimental, ingente y laborioso, y por hacer un ejercicio de inducción, a la manera de los juicios sintéticos *a priori* de Kant, por lo que se afirma que tales acumulaciones mencionadas están descritas por una densidad binomial, aunque de ninguna manera, lo será debido a esas demostraciones espurias.

No obstante, es muy atractivo tener una fórmula, una ley numérica básica, la densidad binomial $(b_n(k))_{k=0}^n$ que rige la *repartición* de los resultados de los colectivos de (\mathcal{E}_n, X_n) , $(\mathcal{E}_n, [\overline{0, n}])$ y $(\mathcal{E}_n, [0, 1]_n)$ entre los pedazos pertenecientes a las particiones $\mathcal{P}_n, \mathcal{P}_n^{(1)}, \mathcal{P}_n^{(2)}$ correspondientes a sus respectivos espacios de resultados y que, por lo tanto, con su auxilio, puede saberse de antemano, predecirse, aproximadamente, cuántos resultados de un gran número de ejecuciones sucesivas de \mathcal{E}_n caerán en cada uno de esos pedazos de N , grande, ejecuciones (independientes) sucesivas de \mathcal{E}_n , del número N de resultados obtenidos, aproximadamente, en el sentido de las desigualdades de Bernoulli como se estableció en las tres últimas secciones,

$$\begin{aligned} Nm_n(P_k) &= Nm_n^{(1)}(\{k\}) = Nm_n^{(2)}\left(\left\{\frac{k}{n}\right\}\right) \\ &= N\binom{n}{k}p^k(1-p)^{n-k} \end{aligned}$$

caen en $P_k, \{k\}, \left\{\frac{k}{n}\right\}$ (o, equivalentemente, éste es el número de veces, aproximadamente, en el sentido anterior, que se *realiza* $P_k, \{k\}, \left\{\frac{k}{n}\right\}$ en esas N ejecuciones de \mathcal{E} aunque agrupándose mayoritariamente en $P_{k_0}, \{k_0\}, \left\{\frac{k_0}{n}\right\}$, para ir disminuyendo monótonamente a medida que k se aleja de k_0 , según lo dicta tal propiedad de $b_n(k)$; cuantitativamente:

$$Nm_n(P_0) < \dots < Nm_n(P_{k_0-1}) < Nm_n(P_{k_0}) \geq Nm_n(P_{k_0+1}) > \dots > Nm_n(P_n),$$

$$Nm_n^{(1)}(\{0\}) < \dots < Nm_n^{(1)}(\{k_0-1\}) < Nm_n^{(1)}(\{k_0\}) \geq Nm_n^{(1)}(\{k_0+1\}) > \dots > Nm_n^{(1)}(\{n\})$$

y

$$Nm_n^{(2)}(\{0\}) < \dots < Nm_n^{(2)}\left(\left\{\frac{k_0-1}{n}\right\}\right) < Nm_n^{(2)}\left(\left\{\frac{k_0}{n}\right\}\right) \geq Nm_n^{(2)}\left(\left\{\frac{k_0+1}{n}\right\}\right) > \dots > Nm_n^{(2)}(\{1\}).$$

Ejemplo.

En relación al ejemplo 1 del cap. I, introdúzcanse en el recipiente vacío 3 bolas blancas, o marcadas con el número 1, y una bola negra o marcada con el número 0. Interesa el experimento de Bernoulli $(\mathcal{E}, X_1 = \{0, 1\})_{\frac{3}{4}}$.

Para el experimento \mathcal{E}_6 , $n = 6$ repeticiones independientes de \mathcal{E} y como espacio de resultados $X_6 = \{(\delta_1, \dots, \delta_6) : \delta_i = 0 \text{ ó } 1, i = 1, 2, \dots, 6\}$

los resultados de cada colectivo de (\mathcal{E}_6, X_6) se reparten en los pedazos P_0, \dots, P_6 , constituyentes de \mathcal{P}_6 , según la distribución binomial

$$m_6 : \mathcal{P}_6 \rightarrow (0, 1) \ni$$

$$m_6(P_k) = \binom{6}{k} \left(\frac{3}{4}\right)^k \left(\frac{1}{4}\right)^{6-k}, \quad k = 0, 1, \dots, 6.$$

Haciendo los cálculos se obtiene $m_6(P_0) = 2.44 \dots \times 10^{-4}$, $m_6(P_1) = 4.39 \dots \times 10^{-3}$, $m_6(P_2) = 0.3295 \dots$, $m_6(P_3) = 0.1318 \dots$, $m_6(P_4) = 0.2966 \dots$, $m_6(P_5) = 0.3559 \dots$, $m_6(P_6) = 0.1779 \dots$, ejemplo del comportamiento general de la densidad binomial.

En cuanto el valor máximo $m_6(P_5) = 0.3559 \dots$, el máximo $k_0 = 5$ de esta binomial, cumple con las condiciones

$$\frac{3}{4} \times 6 - \frac{1}{4} < 5 < \frac{3}{4} \times 6 + \frac{3}{4}$$

o

$$\frac{3}{4} - \frac{1}{24} < \frac{5}{6} < \frac{3}{4} + \frac{1}{8} :$$

$k_0 = 5$ es el natural más cercano a $np = 6 \times \frac{3}{4}$ en la escala del 0 al 6 o, equivalentemente, $\frac{5}{6}$ el más proximo de los $\frac{k}{6}$, a $p = \frac{3}{4}$ en la escala de 0 a 1. El pedazo P_5 acapara, por decirlo así, más del 35 % del total 1 que se reparte entre los siete pedazos,

$$m_6(P_0) + m_6(P_1) + m_6(P_2) + m_6(P_3) + m_6(P_4) + m_6(P_5) + m_6(P_6) = 1.$$

Y P_5 junto con P_4 y P_6 se llevan

$$m_6(P_4) + m_6(P_5) + m_6(P_6) = 0.8305 \dots,$$

del total 1: acaparan más del 83 % del total del 1, o, equivalentemente, a los pedazos P_0, P_1, P_2, P_3 les toca

$$m_6(P_0) + m_6(P_1) + m_6(P_2) + m_6(P_3) = 0.1696 \dots,$$

menos del 17 %, del total 1.

Y en cuanto a las relaciones esenciales de esta medida básica binomial, y sus propiedades expuestas por medio de las propiedades algebraicas de la densidad binomial, con los resultados de los colectivos de (\mathcal{E}_6, X_6) , se tiene que para N ejecuciones de \mathcal{E}_6 , $N \gg 1$, los N resultados obtenidos se reparten entre los siete pedazos P_k o, lo que es lo mismo, el número de veces que se realiza cada P_k , están dados por los números

$$Nm_6(P_k), \quad k = 0, 1, \dots, 6$$

y siguiendo la *regla de repartición*

$$Nm_6(P_0) < Nm_6(P_1) < \dots < Nm_6(P_5) > Nm_6(P_6),$$

esencialmente, salvo accidentes aleatorios.

Y esto último en el sentido de que cada colectivo de (\mathcal{E}_n, X_n)

$$\begin{array}{l} \text{Ejecución de } \mathcal{E}_6 \quad 1, \quad 2, \quad \dots, \quad N, \dots \\ \text{Resultado} \quad \quad \quad \bar{\delta}^1, \quad \bar{\delta}^2, \quad \dots, \quad \bar{\delta}^N, \dots \end{array}$$

sigue la condiciones de los números grandes de Bernoulli:

$$\left| \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - Nm_6(P_k) \right| \ll N, \quad \forall N \gg 1$$

o, equivalentemente,

$$\left| \frac{1}{N} \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - m_6(P_k) \right| \ll 1, \quad \forall N \gg 1,$$

para cada $P_k \in \mathcal{P}_6$.

Pero, ahora, qué tan próximas son esas sumas, $\sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j)$, números aleatorios, a sus correspondientes $Nm_6(P_k)$, números éstos determinados de antemano, determinísticos, aunque, por lo general, no son números naturales, lo cual hace parecer absurdas esas proximidades.

Una respuesta puede darla la ley asta del logaritmo iterado:

$$\left| \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) - Nm_6(P_k) \right| < \sqrt{\frac{N}{2} \ln \ln N}, \quad \forall N \gg 1,$$

quien circunscribe los valores de las sumas a los números enteros, no negativos, que se encuentran en la vecindad de centro $Nm_6(P_k)$ y radio $\sqrt{\frac{N}{2} \ln \ln N}$,

$$\left(Nm_6(P_k) - \sqrt{\frac{N}{2} \ln \ln N}, Nm_6(P_k) + \sqrt{\frac{N}{2} \ln \ln N} \right);$$

así que

$$\left(Nm_6(P_k) - \sqrt{\frac{N}{2} \ln \ln N} < \sum_{j=1}^N \delta_{(k)}(\bar{\delta}^j) < Nm_6(P_k) + \sqrt{\frac{N}{2} \ln \ln N} \right),$$

para cada $N \gg 1$ y cada P_k de \mathcal{P}_6 .

Es lo mejorcito que se tiene, empíricamente hablando.

Por ejemplo, para $N = 1000$, que para estos menesteres es $\gg 1$, se obtiene que

$$\begin{aligned} 0 \leq \sum_{j=1}^{1000} \delta_{(0)}(\bar{\delta}^j) \leq 31; \quad 0 \leq \sum_{j=1}^{1000} \delta_{(1)}(\bar{\delta}^j) \leq 35; \quad 2 \leq \sum_{j=1}^{1000} \delta_{(2)}(\bar{\delta}^j) \leq 64; \quad 101 \leq \\ \sum_{j=1}^{1000} \delta_{(3)}(\bar{\delta}^j) \leq 162; \quad 266 \leq \sum_{j=1}^{1000} \delta_{(4)}(\bar{\delta}^j) \leq 327; \quad 325 \leq \sum_{j=1}^{1000} \delta_{(5)}(\bar{\delta}^j) \leq 387; \\ 143 \leq \sum_{j=1}^{1000} \delta_{(6)}(\bar{\delta}^j) \leq 209 : \end{aligned}$$

de 1000 ejecuciones sucesivas de \mathcal{E}_6 , de 0 a 31 de sus resultados *caen* en P_0, \dots y de 143 a 209 de sus resultados *caen* en P_6 ; o, lo que es equivalente: de 0 a 31 veces se *realiza* P_0 , de esas 1000 ejecuciones de \mathcal{E}_6 .

Y ese es el *pronóstico*, que como tal puede, también, verse lo anterior. Se invita al lector a corroborar lo establecido previamente con un experimento de Bernoulli que escoja y lo pueda efectuar, ya sea directamente o utilizando las tablas de dígitos al azar que se encuentran en la página electrónica.

4.5. Distribuciones asociadas a $(\mathcal{E}_n, X_n), (\mathcal{E}_n, [0, n]), (\mathcal{E}_n, [0, 1]_n)$.

Ahora bien, a partir de las *distribuciones básicas*

$$m_n : \mathcal{P}_n \rightarrow (0, 1), m_n^{(1)} : \mathcal{P}_n^{(1)} \rightarrow (0, 1), m_n^{(2)} : \mathcal{P}_n^{(2)} \rightarrow (0, 1),$$

despojadas de su origen e interpretación física, de las relaciones con las distribuciones de los colectivos $(\mathcal{E}_n, X_n), (\mathcal{E}_n, [0, n]), (\mathcal{E}_n, [0, 1]_n)$ entre los pedazos de las respectivas particiones de sus correspondientes espacios de resultados, resultan ser, ni más ni menos, que medidas normales básicas y, por lo tanto, usando las técnicas de extensión de medidas de la teoría de la medida de conjuntos, Cap. III, se obtienen sus correspondientes medidas normales

$$\mu_n : a(\mathcal{P}_n) \rightarrow [0, 1], \mu_n^{(1)} : a(\mathcal{P}_n^{(1)}) \rightarrow [0, 1], \mu_n^{(2)} : a(\mathcal{P}_n^{(2)}) \rightarrow [0, 1]$$

cuyos valores en cada

$$\begin{aligned} A &= \sum_{k \in N} P_k, & A &\in a(\mathcal{P}_n); \\ A^{(1)} &= \sum_{k \in N} \{k\}, & A^{(1)} &\in a(\mathcal{P}_n^{(1)}); \\ A^{(2)} &= \sum_{k \in N} \left\{ \frac{k}{n} \right\}, & A^{(2)} &\in a(\mathcal{P}_n^{(2)}), \end{aligned}$$

con $N \subset \{0, 1, \dots, n\}$, son

$$\mu(A) = \mu_n^{(1)}(A^{(1)}) = \mu_n^{(2)}(A^{(2)}) = \sum_{k \in N} \binom{n}{k} p^k (1-p)^{n-k}.$$

las cuales, interpretadas en el contexto físico del cual surgieron, o se extrajeron, sus medidas básicas, se convierten en *distribuciones* que describen (hasta cierto grado de poder de resolución, no más allá de los pedazos de las particiones) a las *distribuciones esenciales* respectivas.

Pero es imposible demostrar formalmente esta severación, ya que el fenómeno físico y la estructura matemática interrelacionadas (*Tesis de correspondencia*) no constituyen un sistema axiomático formal.

Podría declararse de manera radical, en última instancia y sin mayor conmiseración, que así funcionan las cosas y remitirse a la prueba de los hechos experimentales. Sin embargo, se expondrán algunos argumentos plausibles que sustentan aquella afirmación, los cuáles, por cierto, son muy aleccionadores.

Para esto, considérese N ejecuciones sucesivas de \mathcal{E}_n , entonces, los números

$$\begin{aligned} \sum_{j=1}^N \delta_A(\overline{\delta}^j) &= \sum_{j=1}^N \sum_{k \in N} \delta_{(k)}(\overline{\delta}^j) \\ &= \sum_{k \in N} \sum_{j=1}^N \delta_{(k)}(\overline{\delta}^j) \\ \sum_{j=1}^N \delta_{A^{(1)}}\left(\sum_{i=1}^n \delta_i^j\right) &= \sum_{k \in N} \sum_{j=1}^N \delta_{\{k\}}\left(\sum_{i=1}^n \delta_i^j\right) \\ \sum_{j=1}^N \delta_{A^{(2)}}\left(\frac{1}{n} \sum_{i=1}^n \delta_i^j\right) &= \sum_{k \in N} \sum_{j=1}^N \delta_{\left\{\frac{k}{n}\right\}}\left(\frac{1}{n} \sum_{i=1}^n \delta_i^j\right), \end{aligned}$$

(que en realidad son iguales) son el número de veces que en esas N ejecuciones se realizan los conjuntos A , $A^{(1)}$ y $A^{(2)}$ o, lo que es lo mismo, el número de resultados, del total N, que *caen* en A , en $A^{(1)}$ y en $A^{(2)}$.

Ahora, utilizando la condición o propiedad de los números grandes de Bernoulli, Secs 1, 2 y 3, cuando N es grande el número de resultados que *caen* en los pedazos en que se despedazan esos conjuntos, P_k , $\{k\}$, $\left\{\frac{k}{n}\right\}$, con $k \in N$, es del todo razonable inferir que

$$\begin{aligned} \sum_{j=1}^N \delta_A(\overline{\delta}^j) &\approx \sum_{k \in N} Nm_n(P_k) = N\mu_n(A), \\ \sum_{j=1}^N \delta_{A^{(1)}}\left(\sum_{i=1}^n \delta_i^j\right) &\approx \sum_{k \in N} Nm_n^{(1)}(\{k\}) = N\mu_n^{(1)}(A^{(1)}), \\ \sum_{j=1}^N \delta_{A^{(2)}}\left(\frac{1}{n} \sum_{i=1}^n \delta_i^j\right) &\approx \sum_{k \in N} Nm_n^{(2)}\left(\left\{\frac{k}{n}\right\}\right) = N\mu_n^{(2)}(A^{(2)}), \end{aligned}$$

o, equivalentemente, que

$$\begin{aligned} \frac{1}{N} \sum_{j=1}^N \delta_A(\overline{\delta}^j) &\approx \mu_n(A), \\ \frac{1}{N} \sum_{j=1}^N \delta_{A^{(1)}}\left(\sum_{i=1}^n \delta_i^j\right) &\approx \mu_n^{(1)}(A^{(1)}), \\ \frac{1}{N} \sum_{j=1}^N \delta_{A^{(2)}}\left(\frac{1}{n} \sum_{i=1}^n \delta_i^j\right) &\approx \mu_n^{(2)}(A^{(2)}). \end{aligned}$$

Aproximaciones que ya expresan cómo se distribuyen los resultados de cada colectivo de (\mathcal{E}_n, X_n) , $(\mathcal{E}_n, [\overline{0, n}])$, $(\mathcal{E}_n, [0, 1]_n)$ por sus respectivos espacios de resultados, aunque *respecto* a las correspondientes particiones seleccionadas \mathcal{P}_n , $\mathcal{P}_n^{(1)}$, $\mathcal{P}_n^{(2)}$; esto es, no se sabe, con estas particiones, cómo se distribuyen los resultados por los pedazos P_k , $\{k\}$ o $\left\{\frac{k}{n}\right\}$, sino sólo cuántos resultados de N ejecuciones sucesivas de \mathcal{E}_n *caen* en éstos, aproximadamente. (El límite del poder de resolución que se tiene es hasta dichos pedazos; por este camino ver más allá de esto, en subregiones de los pedazos, no es posible.) Es más, de tomarse

particiones *más finas* que las escogidas, lo que desde luego puede hacerse, haría muy difíciles las mediciones, se necesitarían una cantidad de ejecuciones de \mathcal{E}_n mucho mayores; se originaría una matemática muy engorrosa, que, por otro lado, desbordaría la intención de este libro y que, posiblemente, no resultara ser una descripción tan balanceada como lo son los espacios de medida normal binomial o los espacios con *distribución* binomial.)

O con otro enfoque: los experimentos $(\mathcal{E}_n, \{0, 1\}_A)$, $(\mathcal{E}_n, \{0, 1\}_{A^{(1)}})$ y $(\mathcal{E}_n, \{0, 1\}_{A^{(2)}})$ correspondientes a (\mathcal{E}_n, X_n) y $A \in a(\mathcal{P}_n)$; $(\mathcal{E}_n, [\overline{0, n}])$ y $A^{(1)} \in a(\mathcal{P}_n^{(1)})$; $(\mathcal{E}_n, [\overline{0, n}]_n)$ y $A^{(2)} \in a(\mathcal{P}_n^{(2)})$, son experimentos de Bernoulli; es más, es el mismo experimento de Bernoulli: por un lado los colectivos

$$\begin{array}{ccccccc} \delta_A(\bar{\delta}^1), & \delta_A(\bar{\delta}^2), & \dots, & \delta_A(\bar{\delta}^N), & \dots & & \\ \delta_{A^{(1)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^1\right), & \delta_{A^{(1)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^2\right), & \dots, & \delta_{A^{(1)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^N\right), & \dots & & \\ \delta_{A^{(2)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^1\right), & \delta_{A^{(2)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^2\right), & \dots, & \delta_{A^{(2)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^N\right), & \dots & & \end{array}$$

son idénticos, término a término. Por otro lado la recta que persigue su gráfica, es

$$\begin{aligned} y &= \mu_n(A)x \\ &= \mu_n^{(1)}(A^{(1)})x \\ &= \mu_n^{(2)}(A^{(2)})x \\ &= \sum_{k \in N} \binom{n}{k} p^k (1-p)^{n-k} x, \end{aligned}$$

manteniéndose a las distancias verticales dadas por las condiciones, propiedades o ley de los números grandes de Bernoulli,

$$\begin{aligned} \left| \sum_{j=1}^N \delta_{(A)}(\bar{\delta}^j) - N\mu_n(A) \right| &= \left| \sum_{j=1}^N \delta_{A^{(1)}}\left(\sum_{i=1}^n \delta_i^j\right) - N\mu_n^{(1)}(A^{(1)}) \right| \\ &= \left| \sum_{j=1}^N \delta_{A^{(2)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^j\right) - N\mu_n^{(2)}(A^{(2)}) \right| \ll N, \forall N \gg 1 \end{aligned}$$

o, equivalentemente,

$$\begin{aligned} \left| \frac{1}{N} \sum_{j=1}^N \delta_{(A)}(\bar{\delta}^j) - \mu_n(A) \right| &= \left| \frac{1}{N} \sum_{j=1}^N \delta_{A^{(1)}}\left(\sum_{i=1}^n \delta_i^j\right) - \mu_n^{(1)}(A^{(1)}) \right| \\ &= \left| \frac{1}{N} \sum_{j=1}^N \delta_{A^{(2)}}\left(\frac{1}{n}\sum_{i=1}^n \delta_i^j\right) - \mu_n^{(2)}(A^{(2)}) \right| \ll 1, \forall N \gg 1 \end{aligned}$$

siendo, entonces, $\mu(A)$, $\mu_n^{(1)}(A^{(1)})$, $\mu_n^{(2)}(A^{(2)})$ las constantes físicas asociadas a los conjuntos A , $A^{(1)}$, $A^{(2)}$ en los contextos de los colectivos de (\mathcal{E}_n, X_n) , $(\mathcal{E}_n, [\overline{0, n}])$, $(\mathcal{E}_n, [0, 1]_n)$, respectivamente, las cuales, interpretadas según aparecen en las desigualdades de Bernoulli, son las medidas de qué tan frecuentemente los resultados del correspondiente colectivo, *caen* en A , $A^{(1)}$, $A^{(2)}$ o,

equivalentemente, las medidas de qué tan frecuentemente se realizan estos subconjuntos, de sus correspondientes espacios de resultados, en cada sucesión de ejecuciones de \mathcal{E}_n .

Así pues, la medida normal

$$\mu_n : a(\mathcal{P}_n) \rightarrow [0, 1]$$

es una descripción de la *distribución esencial* de las distribuciones de los resultados de los colectivos de (\mathcal{E}_n, X_n) , *salvo accidentes aleatorios*, y a quien, en este contexto se le denominará *distribución binomial*. Y lo mismo puede afirmarse de las medidas normales, *distribuciones binomiales*,

$$\mu_n^{(1)} : a(\mathcal{P}_n^{(1)}) \rightarrow [0, 1] , \mu_n^{(2)} : a(\mathcal{P}_n^{(2)}) \rightarrow [0, 1]$$

en relación a sus respectivos colectivos $(\mathcal{E}_n, [\overline{0, n}])$ y $(\mathcal{E}_n, [0, 1]_n)$.

Tres estructuras matemáticas se han construido, tres *espacios con distribución*, de *medida normal* (o de *probabilidad*), en particular, *binomiales* (un aspecto de la *Tesis de correspondencia*),

$$(X_n, \mu_n : a(\mathcal{P}_n) \rightarrow [0, 1]) , ([\overline{0, 1}], \mu_n^{(1)} : a(\mathcal{P}_n^{(1)}) \rightarrow [0, 1]), ([0, 1]_n, \mu_n^{(2)} : a(\mathcal{P}_n^{(2)}) \rightarrow [0, 1])$$

a partir de un experimento de Bernoulli y un número natural n . Pero las cosas no quedan sólo en esto. En vista de la descripción teórica que estas medidas hacen acerca de las distribuciones de los resultados de los colectivos de (\mathcal{E}_n, X_n) , $(\mathcal{E}_n, [\overline{0, n}])$, $(\mathcal{E}_n, [0, 1]_n)$ por sus correspondientes espacios de resultados, se obtendrán formalmente, de estas estructuras, algunas consecuencias, teoremas, acerca del comportamiento teórico, muy peculiar, de las medidas cuando los valores de n son arbitrariamente grandes, es más cuando $n \rightarrow +\infty$, lo cual redundará en un conocimiento adicional en lo relativo a las distrsibuciones de resultados.

Ejemplo. Continuando con el Ejemplo del final de la Sec. 4., se obtiene la *medida binomial*

$$\mu_6 : a(\mathcal{P}_6) \rightarrow [0, 1] \ni$$

$$\mu_6(A) = \sum_{k \in N} m_6(P_k), \forall A = \sum_{k \in N} P_k \in a(\mathcal{P}_6);$$

en particular, para

$$\begin{aligned} A &= P_4 + P_5 + P_6, \\ \mu_6(A) &= \sum_{k=4}^6 m_6(P_k) \\ &= 0.8305 \dots \end{aligned}$$

y, por lo tanto,

$$\mu_6(A^c) = \sum_{k=0}^3 m_6(P_k) = 0.1696 \dots$$

De las relaciones de estos hechos del espacio de medida normal binomial, $(X, \mu_6 : a(\mathcal{P}_6) \rightarrow [0, 1])$, con los relativos a las distribuciones de los resultados de cada colectivo de la pareja (\mathcal{E}_6, X_6) mediante las condiciones (o ley) de los números grandes de Bernoulli, se tiene para aquellas $N = 1000$, por ejemplo, ejecuciones de \mathcal{E}_6 , que

$$\sum_{j=1}^{1000} \delta_A(\bar{\delta}^j) \approx 830.5 \dots$$

y

$$\sum_{j=1}^{1000} \delta_{A^c}(\bar{\delta}^j) \approx 169.6 \dots ,$$

aproximaciones según las desigualdades de Bernoulli; de los 1000 resultados obtenidos, aproximadamente 830.5... *caen* en A y, por consecuencia, alrededor de 169.6... en A^c ; o, equivalentemente, de $N = 1000$ ejecuciones sucesivas de \mathcal{E}_6 , alrededor de 830.5... de éstas se realiza A y aproximadamente 169.6... veces se realiza A^c .

Aunque, utilizando la ley basta del logaritmo iterado, esas aproximaciones pueden ser más específicas:

$$\left| \sum_{j=1}^{1000} \delta_A(\bar{\delta}^j) - 830.5 \dots \right| < 31.08 \dots$$

y

$$\left| \sum_{j=1}^{1000} \delta_{A^c}(\bar{\delta}^j) - 169.6 \dots \right| < 31.08 \dots$$

de las cuales se obtiene que

$$800 \leq \sum_{j=1}^{1000} \delta_A(\bar{\delta}^j) \leq 861$$

y

$$139 \leq \sum_{j=1}^{1000} \delta_{A^c}(\bar{\delta}^j) \leq 200,$$

o sea, que de 1000 ejecuciones sucesivas, entre 800 y 861 resultados *caen* en A y de 139 a 200 *caen* en A^c o, lo que es equivalente, de 800 a 861 veces se *realiza* A y de 139 a 200 se *realiza* A^c .

Un hecho que es importante resaltar acerca de esto es que el pedazo P_5 , con sus pedazos *contiguos*, P_4 y P_6 , funciona como un *atractor* de los resultados de \mathcal{E}_6 , en X_6 ; no hay que olvidar cómo son los índices de esos pedazos: son

próximos a $\frac{3}{4} \times 6$ o, lo que es lo mismo, $\frac{4}{6}, \frac{5}{6}, \frac{6}{6}$ son próximos a $\frac{3}{4}$, la pendiente de la recta que persiguen las gráficas del experimento de Bernoulli $(\mathcal{E}, \{0, 1\})_{\frac{3}{4}}$: la proporción de bolas marcadas son 1, o blancas, del total de bolas que se tienen en el recipiente.

Y esa función de *atractor* no se debe tanto al *tamaño* de P_5 o de $P_4 + P_5 + P_6$; *tamaño* en X_6 , en cuanto al número de elementos que los constituyen:

$\#(P_0) = 1$, $\#(P_1) = 6$, $\#(P_2) = 15$, $\#(P_3) = 20$, $\#(P_4) = 15$, $\#(P_5) = 6$, $\#(P_6) = 1$, , el cual es un buen referente con el cual comparar.

El *tamaño* de P_5 es significativamente menor que el de P_3 y, sin embargo, *caen* más de tres veces resultados, de un colectivo de (\mathcal{E}_6, X_6) , en P_5 que los que *caen* en P_3 .

Y en cuanto al *tamaño* de $A = P_4 + P_5 + P_6$, $\#(A) = 22$, comparado con el de $A^c = P_0 + P_1 + P_2 + P_3 + P_4$, $\#(A^c) = 42$, la mitad prácticamente; pero esto no obsta para que más del 83 % de esos resultados *caigan* en A mientras que en A^c menos de 17 %.

La tendencia, en un primer momento, sería pensar que por ser *más grande* una región de X_6 , aquellos resultados propenderían a caer en ésta; pero recordando el síndrome de Bernoulli de los colectivos de $(\mathcal{E}, \{0, 1\})_{\frac{3}{4}}$, se convence uno de lo razonable que es el comportamiento de la medida normal binomial, distribución binomial, del principio, de cuán fidedignamente ésta representa las propiedades relevantes de las distribuciones de los resultados de los colectivos de (\mathcal{E}_6, X_6) por X_6 . Más adelante se verá, teóricamente, la relevancia de los pedazos P_k , en \mathcal{P}_n , de X_n , cuyos índices sean próximos a np , en relación a los experimentos de Bernoulli truncos de rango n , (\mathcal{E}_n, X_n) .