Curso de Probabilidades

María Margarita Olivares

Octubre 2002

Índice General

PRÓLOGO

Estas notas es el resultado de la experiencia de la autora en el dictado del curso de probabilidades de la Licenciatura en Matemática de la Universidad Central de Venezuela. y de la consulta de las guías elaboradas para este curso por J. Ortega y M. Arriojas.

La misma corresponde al progarma vigente del curso e incluye ejercicios y problemas al final de cada capitulo.

El trabajo de mecanografía estuvo a cargo de la autora. Agradezco cualquier observación o comentario.

Capítulo 1

INTRODUCCIÓN Y CONCEPTOS BÁSICOS:

La acción de la naturaleza y de la obra del hombre, colocada en determinadas condiciones, produce en muchos casos, resultados, que hasta el hombre más prudente, puede predecir. Así, al relámpago sucede el trueno, a la noche el día, el agua en una marmita puesta al fuego se evapora, etc. Son hechos inmediatos al alcance del hombre más primitivo, ellos encendieron en él la fe en la existencia de leyes inmutables que gobiernan la naturaleza, en la ausencia del capricho del mundo de los fenómenos naturales. Esta fe ha sido la que ha incitado al hombre a aumentar su caudal de conocimientos, a investigar, que en definitiva es arrancar a la naturaleza, el secreto de las leyes que la gobiernan, de las regularidades de los fenómenos y así, la CAUSALIDAD figura en la primera página del libro de la historia del pensamiento científico.

Pero frente a este sentimiento de necesidad de las leyes naturales, coexistía la creencia en otros hechos imposibles de predecir, especialmente aquellos en los que interviene el libre albedrío humano, y así, al lado de la causalidad surgió la CASUALIDAD o el AZAR.

Parece a primera vista algo paradójico que el azar, lo imprevisible, pueda ser objeto de tratamiento matemático, esté sujeto a regularidades, sometido a leyes, y en definitiva, que pueda ser prevista la acción del azar.

La teoría de probabilidades tiene por objeto el análisis matemático de la noción intuitiva de SUERTE ó AZAR.

El cálculo de probabilidades se inició con el estudio de los juegos de azar, introduciéndose posteriormente en todas las ramas de la actividad científica, como por ejemplo en el análisis matemático (Teoría de Potencial), en la economía (Modelos Econométricos), en la genética (Leyes de Mendel), en la física corpuscular, en las ciencias sociales, en la informática y en otras disciplinas del conocimiento. Su amplio campo de aplicaciones se debe a que la teoría de probabilidades permite estudiar y ”medir” la incertidumbre que forma parte de casi todo lo que existe a nuestro alrededor.

El inicio oficial de la teoría de probabilidades, corresponde a mediados del siglo XVII ya que en la sociedad francesa, los juegos de azar eran una actividad generalizada entre las personas adineradas. Los juegos de dados, ruletas, cartas, etc., eran cada vez mas complicados, surgían preguntas sobre métodos para medir el riesgo sobre determinada apuesta.Uno de estos jugadores, De Méré, consultó en Paris a Pascal sobre algunas de estas preguntas. Así se originó un intercambio de correspondencias entre Pascal y matemáticos famosos de la época, entre ellos Fermat, surgiendo comentarios y soluciones a muchos problemas planteados, iniciándose de esa forma lo que hoy se conoce como la Teoría de Probabilidades.

El desarrollo de la Teoría de la Medida y de la Integración, en los años 1930, proporcionó a la Teoría de Probabilidades las bases matemáticas fundamentales que le permitieron desarrollarse de manera rigurosa.

EXPERIMENTO:

Un experimento es la producción de un fenómeno, con la finalidad de investigar sus propiedades o causas.

Hay dos tipos de experimentos: determinísticos y aleatorios.

EXPERIMENTOS DETERMINÍSTICOS :

Son aquellos que repetidos bajo las mismas condiciones, dan siempre un mismo resultado, el cual es predecible.

EJEMPLO:

Si dejamos caer una pelota en un tubo donde se ha hecho el vacío, la pelota llegará al suelo en el instante

$$t = \sqrt{\frac{2d}{g}}$$

donde d es la distancia desde donde se deja caer la pelota y g es la aceleración de gravedad.

EXPERIMENTOS ALEATORIOS:

Son aquellos cuyos resultados no pueden ser exactamente conocidos a priori, aunque se repitan bajo las mismas condiciones.

EJEMPLO:

Los juegos de azar proporcionan numerosos ejemplos de experiencias aleatorias, así, un juego de dados, lanzar una moneda y observar si sale cara o sale sello, constituyen experiencias aleatorias.

POBLACIÓN:

Es un conjunto de individuos u objetos acerca del cual se quiere saber algo.

EJEMPLO:

Alumnos de una escuela, piezas producidas por una máquina en cierto período.

MUESTRA:

Es una parte de la población que elegida adecuadamente, es representativa de ésta.

ESPACIO MUESTRAL:

Es el conjunto de posibles resultados que se obtienen al realizar un experimento aleatorio. Usualmente, denotamos por Ω el espacio muestral, por ω los puntos de Ω y escribimos

$$\omega \in \Omega$$

para indicar que ω pertenece a Ω.

Los elementos de un espacio muestral se denominan: **eventos elementales**.

Un espacio muestral se dice discreto si tiene un número finito o numerable de eventos elementales, en caso contrario se dice que es continuo.

EJEMPLOS:

1. En un proceso de fabricación extraemos un artículo elegido entre los artículos fabricados y observamos si es o no es defectuoso.

   Si denotamos por B cuando el artículo es bueno y por D cuando el artículo es defectuoso, podemos tomar

   $$\Omega = \{B, D\}$$

Si en lugar de un artículo elegimos n artículos, podemos considerar

$$\Omega = \{(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) : \varepsilon_i = 0 \text{ ó } 1;\ (i = 1, 2, \dots, n)\}$$

donde $\varepsilon_i = 1$ indica que el artículo es defectuoso y $\varepsilon_i = 0$ indica que el artículo es bueno.

2. En una planta de prueba, se elige un conjunto de lámparas eléctricas y se conectan cada una de ellas, hasta que se queme, observando en cada caso el tiempo de duración.

   Si se consideran n lámparas, podemos tomar:

$$\Omega = \{(t_1, t_2, \dots, t_n) : t_i \in \mathbb{R}, t_i \geq 0\}$$

3. Se lanzan dos dados y se observa la puntución de cada un de ellos.

$$\Omega = \{(i, j) : 1 \leq i \leq 6,\ 1 \leq j \leq 6,\ i, j \in \mathbb{N}\}$$

4. Se efectúan disparos sobre un disco de 1 metro de radio, situado a cien metros de distancia. No se toman en cuenta los impactos fuera del disco.

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

EVENTOS:

Estamos interesados en considerar familias de subconjuntos de Ω (familia de eventos). Por ejemplo, en el ejemplo 1, podemos interesarnos en el evento "entre los n artículos extraídos hay m defectuosos", es decir, en el subconjunto de Ω definido por

$$A = \{(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) : \varepsilon_i = 0 \text{ ó } 1;\ \sum_{i=1}^{n} \varepsilon_i = m\}$$

Análogamente, en el ejemplo 4, podríamos estar interesados en subconjuntos del tipo "la distancia al origen, del impacto, se encuentra comprendida entre $0,3$ metros y $0,5$ metros".

De modo que consideraremos una familia $\mathcal{F}$ de subconjuntos de Ω y diremos que el evento $F \in \mathcal{F}$ ocurre al realizar un experimento aleatorio cuyo resultado es el evento elemental ω, si $\omega \in F$.

La familia de eventos $\mathcal{F}$ debe satisfacer ciertas condiciones:

a) $\Omega \in \mathcal{F}$ es decir, "algo siempre ocurre". A Ω le llamaremos el evento cierto.

b) Si A es un evento, pediremos que "no ocurra A" también lo sea, es decir

$$A \in \mathcal{F} \implies A^c \in \mathcal{F}$$

c) Si se tiene una sucesión de eventos $A_1, A_2, \dots, A_n, \dots$ entonces "alguno de los A_n ocurre" también debe ser un evento, es decir

$$(A_k)_{k \geq 1} : A_k \in \mathcal{F}, k \geq 1, \implies \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$$

Toda familia de subconjuntos de Ω que satisfaga las tres condiciones anteriores se le llama una $\sigma-$álgebra de subconjuntos de Ω.

En lo que sigue supondremos que la familia de eventos asociada a un experimento aleatorio, de espacio muestral Ω es una $\sigma-$álgebra de subconjuntos de Ω.

PROPOSICIÓN:

Sea $\mathcal{F}$ una $\sigma-$álgebra de subconjuntos de Ω, entonces:

1. $\emptyset \in \mathcal{F}$. ($\emptyset$ es el conjunto vacío).
2. Si $A_1, A_2, \ldots, A_n \in \mathcal{F}$ entonces $\bigcup_{k=1}^{n} A_k \in \mathcal{F}$.
3. $(A_k)_{k\geq 1} : A_k \in \mathcal{F}, k \geq 1, \implies \bigcap_{k=1}^{\infty} A_k \in \mathcal{F}$.

(La demostración queda como ejercicio).

ESPACIO DE PROBABILIDAD.

FUNCIÓN DE PROBABILIDAD

Si Ω es un espacio muestral y $\mathcal{F}$ es una familia de eventos de Ω ($\mathcal{F}$ es una $\sigma-$álgebra de subconjuntos de Ω) queremos asignarle a cada evento $A \in \mathcal{F}$ un número real $\mathbb{P}(A)$ (la probabilidad de que ocurra A) de modo tal que se cumplan las siguientes condiciones:

1. $0 \leq \mathbb{P}(A) \leq 1$ para todo $A \in \mathcal{F}$.
2. $\mathbb{P}(\Omega) = 1$.
3. Si $A_n \in \mathcal{F}$, $n \geq 1$, es una suecesión de eventos disjuntos dos a dos ($A_n \cap A_m = \emptyset$ si n es distinto de m) entonces $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$

ESPACIO DE PROBABILIDAD

Es una terna $(\Omega, \mathcal{F}, \mathbb{P})$ formada por un espacio muestral Ω, una σ-álgebra $\mathcal{F}$ y una función de probabilidad $\mathbb{P}$.

PROPIEDADES.

1. $\mathbb{P}(\emptyset) = 0$.
2. Si $A_1, A_2, \ldots, A_n \in \mathcal{F}$, son disjuntos dos a dos, entonces $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k)$.
3. Si $A_1 \subseteq A_2$, $A_1, A_2 \in \mathcal{F}$ entonces $\mathbb{P}(A_1) \leq \mathbb{P}(A_2)$.
4. Si $A_1, A_2 \in \mathcal{F}$ entonces $A_1 - A_2 \in \mathcal{F}$ y $\mathbb{P}(A_1 - A_2) = \mathbb{P}(A_1) - \mathbb{P}(A_1 \cap A_2)$. En particular, si $A_2 \subseteq A_1$ $\mathbb{P}(A_1 - A_2) = \mathbb{P}(A_1) - \mathbb{P}(A_2)$.
5. Si $A_1, A_2 \in \mathcal{F}$ no son necesariamente disjuntos, entonces $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$.
6. Si $A_1 \subseteq A_2 \subseteq A_3 \subseteq A_4 \subseteq A_5 \subseteq A_6 \subseteq \ldots \subseteq A_k \subseteq A_{k+1} \subseteq \ldots$ es una sucesión creciente de eventos, entonces $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim\limits_{n\to\infty} \mathbb{P}(A_n)$.
7. Si $A_1 \supseteq A_2 \supseteq A_3 \supseteq A_4 \supseteq A_5 \supseteq A_6 \supseteq \ldots \supseteq A_k \supseteq A_{k+1} \supseteq \ldots$ es una sucesión decreciente de eventos, entonces $\mathbb{P}\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim\limits_{n\to\infty} \mathbb{P}(A_n)$.

DEMOSTRACIÓN:

1. Sea $A_1 = \Omega$ y $A_k = \emptyset$ para $k \geq 2$, esta sucesión de eventos es disjunta dos a dos, por definición de función de probabilidad,
$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k). = \mathbb{P}(\Omega) + \sum_{k=2}^{\infty} \mathbb{P}(A_k)$$
de aquí se deduce que $\mathbb{P}(A_k) = 0$ si $k \geq 2$, pero $\mathbb{P}(A_k) = \mathbb{P}(\emptyset)$.

2. Sea $B_k = A_k$ si $k = 1, 2, \ldots, n$ y $B_k = \emptyset$ para $k \geq n+1$. La sucesión de eventos $(B_k)_{k\geq 1}$ está formada por conjuntos disjuntos dos a dos, de modo que:

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(B_k) = \sum_{k=1}^{n} \mathbb{P}(A_k)$$

pues $\mathbb{P}(B_k) = 0$ si $k \geq n+1$,por otro lado $\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{n} A_k$, por lo tanto

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

3. Note que $A_2 - A_1 = A_2 \cap A_1^c \in \mathcal{F}$ si $A_1, A_2 \in \mathcal{F}$. Si $A_1 \subseteq A_2$ entonces $A_2 = A_1 \cup (A_2 - A_1)$ que es unión disjunta de dos eventos, por lo tanto $\mathbb{P}(A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2 - A_1)$, y como $\mathbb{P}(A_2 - A_1) \geq 0$ concluímos que $\mathbb{P}(A_2) \geq \mathbb{P}(A_1)$.

4. En general, $A_1 = (A_1 \cap A_2) \cup (A_1 - A_2)$ que es una unión disjunta de dos eventos, por lo tanto

$$\mathbb{P}(A_1) = \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 - A_2)$$

si $A_2 \subset A_1$ entonces $(A_1 \cap A_2) = A_2$, de aquí se deduce que

$$\mathbb{P}(A_1) - \mathbb{P}(A_2) = \mathbb{P}(A_1 - A_2)$$

5. Si $A_1, A_2 \in \mathcal{F}$ no son necesariamente disjuntos, podemos expresar

$$A_1 \cup A_2 = A_2 \cup (A_1 - A_2)$$

que es una unión disjunta de eventos, además

$$A_1 = (A_1 - A_2) \cup (A_1 \cap A_2)$$

de aquí se deduce que

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

6. A partir de la sucesión A_n de eventos crecientes, definimos una sucesión disjunta de eventos de la siguiente manera:

$$B_0 = A_0 = \emptyset,\ B_n = A_n - A_{n-1},\ n \geq 1$$

puesto que $\bigcup_{k=1}^{\infty} A_n = \bigcup_{k=1}^{\infty} B_n$ se cumple que

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(B_k) = \\ \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(A_k - A_{k-1}) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(A_k) - \mathbb{P}(A_{k-1}) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

7. A partir de la sucesión A_n de eventos decrecientes, obtenemos una sucesión creciente de eventos, definida por

$$B_n = A_n^c$$

puesto que

$$\bigcap_{k=1}^{\infty} A_{k=} \left(\bigcup_{k=1}^{\infty} B_k\right)^c$$

por las propiedades anteriores resulta

$$\mathbb{P}\left(\bigcap_{k=1}^{\infty} A_{k=}\right) = \mathbb{P}\left[\left(\bigcup_{k=1}^{\infty} B_k\right)^c\right] = 1 - \mathbb{P}\left[\left(\bigcup_{k=1}^{\infty} B_k\right)\right] = \\ 1 - \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

EJERCICIOS

Indique cuáles de las siguientes igualdades entre conjuntos son correctas y cuáles no lo son:

1. $(A \cup B) - C = A \cup (B - C)$
2. $A \cap B \cap C = [(A \cap B) \cap (C \cup B)] - B^c$
3. $A \cup B \cup C = A \cup [B - (A \cap B)] \cup (C - (A \cap C))$
4. $A \cup B = A \cup (B - (A \cap B))$
5. $(A \cup B) - A = B$
6. $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$
7. $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$

El problema de cómo definir la función $\mathbb{P}$, es decir, de cómo asignar una probabiliadad a cada evento, debe ser resuelto de acuerdo a las condiciones concretas de cada experimento aleatorio considerado.

Si Ω es un espacio muestral de N puntos, la única forma de asignar una probabilidad sobre este espacio, de manera que todos los eventos elementales tengan la misma probabilidad es definiendo

$$\mathbb{P}(A) = \frac{\#A}{N}$$

donde $\#A$ es el cardinal de A ó números de elementos de A.

En efecto, si $\Omega = \{\omega_1, \omega_2, \omega_3, \cdots, \omega_N\}$ y $p = \mathbb{P}(\{\omega_i\})$ $i = 1, 2, \cdots, N$ entonces

$$\mathbb{P}(\Omega) = \sum_{i=1}^{n} \mathbb{P}(\{\omega_i\}) = \sum_{i=1}^{n} p = Np$$

y puesto que $\mathbb{P}(\Omega)$ =1, obtenemos que $p = \frac{1}{N}$. Por lo tanto si A es un evento de k elementos $(k \leq N)$ entonces

$$\mathbb{P}(A) = kp = \frac{k}{N} = \frac{\#A}{N}.$$

En aquellas experiencias aleatorias relacionadas con juegos de azar, las cuales dan lugar a espacios muestrales finitos, es natural desear que todos los sucesos elementales tengan la misma probabilidad. Esta probabilidad, llamada probabilidad canónica, puede ser natural para ciertas experiencias aleatorias mientras que para otras puede resultar totalmente inadecuada.

En la práctica, la composición exacta del espacio muestral Ω, se conoce muy pocas veces, por lo que los valores exactos de las probabilidades de los eventos son desconocidas.

En estos casos se recurre a la noción de frecuencia relativa. Se procede de la siguiente manera: si A es un evento $(A \in \mathcal{F})$, repetimos N veces la experiencia y observamos si A ocurre. Si en las N repeticiones A ocurre k_N veces, llamaremos frecuencia relativa a

$$f_N = \frac{k_N}{N}.$$

Al definir $\mathbb{P}(A) = f_N$ nos tropezamos con dos dificultades principales:

a) Tendríamos que precisar las condiciones de repetición del experimento que consiste en observar A las N veces que lo realizamos.

b) Nada asegura que k_N sea siempre el mismo cada vez que se realizan las N observaciones, por lo tanto $\mathbb{P}(A)$ no estaría bien definido de esta manera.

Veremos más adelante (leyes de grandes números) cuáles son las relaciones entre frecuencia relativa y probabilidad, y cómo, en un sentido que precisaremos más adelante, la frecuencia relativa tiende a la probabilidad del evento considerado, cuando el número N de observaciones aumenta indefinidamente.

Describir el espacio muestral Ω explícitamente puede ser bastante tedioso cuando consta de un número muy grande de puntos y se debe estar seguro de no omitir ningun evento elemental. Puesto que en muchos problemas debemos calcular $\#A, \#\Omega$, es conveniente conocer algunos métodos de enumeración que describiremos en el próximo capítulo.

EJERCICIOS PARA LA PRÁCTICA

1. Considere la experiencia aleatoria consisitente en seleccionar al azar un matrimonio de esta ciudad y registrar las edades de ambos cónyugues.

   a) Cuál es el Espacio Muestral?. Represéntelo en términos de conjuntos y gráficamente.
   b) Represente gráficamente sobre el Espacio Muestral los siguientes sucesos:
      i) Ambos cónyugues son mayores de 21 años.
      ii) El marido es mayor que la mujer.
      iii) La suma de las edades es mayor que 50 y menor que 80 años.

2. Se lanza una moneda tres veces consecutivas, describa el Espacio Muestral en cada uno de los siguientes casos:

   a) Si interesa el resultado de cada lanzamiento.
   b) Si interesa el número total de caras.

3. Se arrojan 5 dados y se suman los puntos.

   a) Describa el Espacio Muestral.
   b) Cite dos sucesos
   c) El conjunto $\{3, 4\}$ será un suceso de este espacio muestral?. Explique.

4. Se arrojan 2 dados y se observan los puntos de cada dado.

   a) Describa el Espacio Muestral.
   b) Describa los siguientes sucesos:
      i) Sale al menos un tres.
      ii) Sale a lo sumo un tres.
      iii) Sale exactamente un tres.
      iv) La suma de los puntos es siete.

5. Sabiendo que $\mathbb{P}(A_1) = a$, $\mathbb{P}(A_2) = b$, $\mathbb{P}(A_1 \cap A_2) = c$. Hallar:

   a) $\mathbb{P}(A_1 \cap A_2^c)$
   b) $\mathbb{P}(A_1 \cup A_2^c)$
   c) $\mathbb{P}\left((A_1 \cap A_2)^c\right)$
   d) $\mathbb{P}\left((A_1 \cap A_2^c)^c\right)$
   e) $\mathbb{P}(A_1^c \cup A_2^c)$
   f) $\mathbb{P}\left(A_1^c \cap (A_1 \cup A_2)\right)$

6. De los cinco dígitos $\{1, 2, 3, 4, 5\}$ se elige uno al azar y luego el otro se escoge entre los cuatro restantes. Suponga que los veinte resultados posibles tienen la misma probabilidad. Encontrar la probabilidad de que un dígito impar sea seleccionado

   a) la primera vez.
   b) la segunda vez
   c) ambas veces.

7. Se lanza una moneda tres veces, cuál es la probabilidad de que salga cara

   a) al menos una vez

b) dos veces
c) a lo sumo una vez
d) una vez.

8. En una habitación se encuentra el siguiente grupo de personas: 5 hombres mayores de 21 años, 4 hombres menores de 21 años, 9 mujeres (6 mayores y 3 menores de 21 años). Se elige una persona al azar y se definen los siguientes sucesos:

$$\begin{aligned} A &= \{\text{la persona es mayor de 21 años}\} \\ B &= \{\text{la persona es menor de 21 años}\} \\ C &= \{\text{la persona es hombre}\} \\ D &= \{\text{la persona es mujer}\} \end{aligned}$$

Evaluar

a) $\mathbb{P}(B \cup D)$
b) $\mathbb{P}(A \cup C^c)$
c) $\mathbb{P}(A - D)$

9. En una habitación 10 personas tienen insignias numeradas del 1 al 10. Se eligen tres personas al azar y se les pide que se retiren de la habitación y se anotan los números de las insignias.

a) Cuál es la probabilidad de que el número menor de las insignias sea el 5?
b) Cuál es la probabilidad de que el número mayor de las insignias sea el 5?

10. Con conjuntos $\{a_1, a_2, \cdots, a_m\}$ y $\{b_1, b_2, \cdots, b_n\}$ de m y n elementos respectivamente, cuántos pares se pueden formar?. Haga una representación gráfica de la situación.

11. Sea Ω un Espacio Muestral finito (no vacío), si $A \subseteq \Omega$, definimos

$$\mathbb{P}(A) = \frac{\#A}{\#\Omega}$$

demuestre que $\mathbb{P}$ satisface

a) $\mathbb{P}(A) \geq 0$
b) $\mathbb{P}(A) \leq 1$ para todo $A \subseteq \Omega$
c) $\mathbb{P}(\Omega) = 1$
d) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ si $A \cap B = \emptyset$

12. Un experimento aleatorio tiene dos resultados posibles. El primero de ellos tiene probabilidad p y el otro tiene probabilidad p^2. Cuánto vale p?.

13. Si $\mathbb{P}$ es una probabilidad en Ω, sabemos que para todo par de eventos A y B

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Generalice este resultado a tres eventos A, B, C.

14. Cuántos términos habrá en el desrrollo del producto

$$(a+b+c)(d+e+f)(x+y+u+v+w)\,?$$

15. Considere el conjunto de dos dígitos que se puede formar con los enteros

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

    Cuántos de ellos son pares?

    Si consideramos el conjunto de 4 dígitos que podemos formar, en cuántos de ellos, el 1 y el 2 preceden al 9?

16. El carro y las cabras (The New York Times, 1991). Suponga que está en el final de un concurso de TV. El animador del concurso le coloca delante de tres puertas, dos de ellas contienen una cabra y la otra un automóvil, Ud. escoge una puerta y se gana lo que está detrás, al hacerlo, había una cabra. El animador le da la oportunidad de volver a elegir. Asumiendo que Ud.quiere ganarse el carro, es ventajoso para Ud. hacer un cambio de puerta?

17. Variación del problema del carro y las cabras. Suponga que hay 4 puertas, detrás de ellas hay 3 cabras y un auto. Ud.escoge una puerta al azar, el animador escoge una puerta que contiene una cabra. Ud.cambia a una de las otras puertas escogiendo al azar (incluyendo la ya elegida por Ud. y que supuestamente no abrió), cuál es la probabilidad de ganar el carro?.Cuál es la probabilidad de ganar el carro si no cambia de puerta?

18. Suponga que la alarma de su reloj suena en algún momento entre las 6 am y las 7 am. Describa el espacio muestral correspondiente a cada uno de los instantes de tiempo en que la alarma puede sonar.Este espacio será continuo o discreto?

    Suponga ahora que la alarma sólo suena en intervalos de 5 minutos, es decir, a las 6 am, 6.05 am, 6.10 am, etc. Describa el espacio muestral en este caso. Este espacio será continuo o discreto?

19. Un dado es lanzado y luego una moneda dos veces.

    (a) Describa el espacio muestral.

    (b) Asumiendo todos los eventos elementales equiprobables, halle las probabilidades de los siguientes eventos:

        i. Se obtiene un seis y al menos una cara.
        ii. Se obtiene un número par y una cara en el segundo lanzamiento.
        iii. Se obtiene un número menor que 5 y al menos una cara.

Capítulo 2

ANÁLISIS COMBINATORIO

En el estudio de los juegos sencillos de azar, de los procedimientos de muestreo, de los problemas de ocupación y de orden, etc., nos encontramos en general con espacios muestrales finitos, en los cuales se atribuye la misma probabilidad a todos los puntos. Si queremos calcular la probabilidad de un evento A, tenemos que dividir el número de elementos del evento A (casos favorables) entre el número de elementos del espacio muestral (casos posibles). Esto resulta más sencillo, si recurrimos al análisis combinatorio, que nos permitirá, mediante unas técnicas estándar ahorrar trabajo y darle sencillez a la materia.

ARREGLOS:

ARREGLOS POR PAREJA:

Con m elemntos $a_1, a_2, \cdots, a_m$ y n elementos $b_1, b_2, \cdots, b_n$ podemos formar $m \cdot n$ parejas (a_j, b_k), $j = 1, 2, \cdots, m$; $k = 1, 2, \cdots, n$ que contengan un elemento de cada grupo.

Para convencernos de este resultado, es suficiente hacer un gráfico de la situación mencionada:

$$
\begin{array}{cccccccc}
b_n & \bullet & \bullet & \cdots & \bullet & \bullet & \cdots & \bullet \\
b_{n-1} & \bullet & \bullet & \cdots & \bullet & \bullet & \cdots & \bullet \\
b_{n-2} & \bullet & \bullet & \cdots & \bullet & \bullet & \cdots & \bullet \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
b_k & \bullet & \bullet & \cdots & \bullet & \bullet & \cdots & \bullet \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \bullet \\
b_2 & \bullet & \bullet & \cdots & \bullet & \bullet & \cdots & \bullet \\
b_1 & \bullet & \bullet & \cdots & \bullet & \bullet & \cdots & \bullet \\
 & a_1 & a_2 & \cdots & a_j & a_{j+1} & \cdots & a_m
\end{array}
$$

El número de puntos representa el número de parejas que podemos formar, obviamente obtendremos $m \cdot n$ parejas.

ARREGLOS MÚLTIPLES:

Se tienen r grupos

$$\{a_{11}, a_{12}, \cdots, a_{1n_1}\}, \{a_{21}, a_{22}, \cdots, a_{2n_2}\}, \cdots, \{a_{r1}, a_{r2}, \cdots, a_{rn_r}\}.$$

El número de $r-$uplas distintas de la forma $(a_{1j_1}, a_{2j_2}, \cdots, a_{rj_r})$ con

$$1 \leq j_k \leq n_k, 1 \leq k \leq r$$

que podemos formar es

$$\sqcap_{k=1}^{r} n_k = n_1 \cdot n_2 \cdot \ldots \cdot n_r,$$

donde n_i con $i = 1, 2, \cdots, r$ representa el número de elementos del $i-$ésimo grupo.

Para probar este resultado basta observar que cada terna de la forma $(a_{1j_1}, a_{2j_2}, a_{3j_3})$ es por sí misma una pareja constituída por (a_{1j_1}, a_{2j_2}) y el elemento a_{3j_3}. Puesto que el número de elementos de la forma (a_{1j_1}, a_{2j_2}) es $n_1 \cdot n_2$, obviamente, el número de parejas de la forma $((a_{1j_1}, a_{2j_2}), a_{3j_3})$ que podemos formar es $(n_1 \cdot n_2) \cdot n_3 = n_1 \cdot n_2 \cdot n_3$. Para demostrar la afirmación para cualquier r, se generaliza el procedimiento por inducción.

EJEMPLOS:

CLASIFICACIONES MÚLTIPLES: Supongamos que se clasifica a la gente según su sexo, su estado civil y su profesión. Las diversas categorías desempeñan el papel de elementos. Si se consideran 17 profesiones y tomamos en cuenta los cuatro estados civiles (casado, soltero, divorciado y viudo) tendremos $2 \times 4 \times 17 = 136$ clases en total.

TRATAMIENTOS AGRÍCOLAS: En un tratamiento agrícola se prueban tres tratamientos diferentes (aplicación de un fertilizante, empleo de un rociador y temperatura). Si esos tratamientos se aplican en r_1, r_2, r_3 niveles o concentraciones, respectivamente, entonces existen $r_1 \cdot r_2 \cdot r_3$ combinaciones o modos de tratamiento.

CARTAS DE PÓKER: Consideremos las cuatro pintas como un grupo y los trece números como el otro. Se obtienen por lo tanto $4 \times 13 = 52$ combinaciones de cartas.

POBLACIONES Y MUESTRAS:

Consideremos el conjunto ó población, $S = \{a_1, a_2, \cdots, a_n\}$ de n elementos, llamamos muestra de esta población, de tamaño r a toda $r-$upla

$$(a_{j_1}, a_{j_2}, \cdots, a_{j_r})$$

formada por elementos de esta pobalción.

Se pueden distinguir cuatro tipos de muestra:

1. Muestras ordenadas con reemplazo.
2. Muestras ordenadas sin reemplazo.
3. Muestras no ordenadas con reemplazo.
4. Muestras no ordenadas sin reemplazo o subpoblaciones.

Nuestro problema es calcular el número de muestras de tamaño r que pueden ser seleccionadas a partir de una población de tamaño n (ó de n elementos). Obviamente este número dependerá del tipo de muestra que estemos considerando.

Simultáneamente al análisis de muestreo, veremos que este problema es equivalente al de calcular el número de maneras posibles de colocar r bolas en n cajas numeradas. Las diferentes situaciones se obtendrán según se consideren las r bolas distinguibles o indistinguibles combinado con el hecho de permitir o no ocupación múltiple.

MUESTRA ORDENADA CON REEMPLAZO:

Dada una población $S = \{a_1, a_2, \cdots, a_n\}$ de n elementos, cuando extraemos r objetos de esta población, seleccionados uno a uno, tomando en cuenta el orden en que son elegidos, con la posibilidad de que un objeto sea elegido varias veces, ya que al hacer una elección, lo devolvemos o reemplazamos enseguida, antes de elegir el siguiente, diremos que tenemos una muestra ordenada con reemplazo de tamaño r.

En este tipo de muestreo, cada uno de los r elementos se puede elegir de n maneras diferentes, por lo tanto el número de r-uplas de este tipo que podemos obtener es:

$$n \cdot n \cdot n \overset{r}{\cdots} \cdot n = n^r$$

puesto que este muestreo corresponde a un arreglo múltiple, de r grupos cada uno de n elementos.

PROBLEMA EQUIVALENTE: Se tienen n cajas numeradas y r bolas distinguibles (distintas). Supongamos que las r bolas están numeradas: De cuántas maneras se pueden colocar las r bolas numeradas en las n cajas numeradas, permitiendo ocupación múltiple?

Para demostrar la equivalencia de los dos planteamientos, definiremos una biyección entre los conjuntos de resultados posibles de cada uno de ellos, obteniendo además la respuesta del último problema planteado.

Sea $(a_{j_1}, a_{j_2}, \cdots, a_{j_r})$ una muestra genérica, ordenada, de tamaño r, extraída de la población S, $a_{j_k} \in S, 1 \leq j_k \leq n$, $1 \leq k \leq r$. Asociémosle la siguiente configuración de bolas:

la bola # 1 se coloca en la caja # j_1
la bola # 2 se coloca en la caja # j_2
$\vdots$
la bola # $r-1$ se coloca en la caja # j_{r-1}
la bola # r se coloca en la caja # j_r

Como la muestra es con reemplazo, pueden haber números ó subíndices j_i repetidos, es decir, varias bolas pueden ser asignadas en la caja j_i, si este subíndice aparece varias veces en la muestra.

Ilustremos esto con un ejemplo: sea $S = \{a_1, a_2, a_3, a_4\}$. Tomemos una muestra de tamaño 2 tomando en cuenta el orden y con reemplazo. Esta operación se puede realizar de $4^2 = 16$ maneras diferentes. Alternativamente podemos considerar dos bolas numeradas y cuatro cajas numeradas.

A la muestra $\{a_3, a_3\}$ le corresponde la siguiente configuración de bolas: las bolas 1 y 2 están en la caja número 3.

A la muestra $\{a_1, a_4\}$ le corresponde la siguiente configuración de bolas: la bola 1 está en la caja #1 y la bola 2 está en la caja número 4.

EJEMPLOS:

1. Lanzamos una moneda 100 veces y observamos el resultado de cada lanzamiento. En este caso la población S tiene 2 elementos
$$S = \{c, s\}$$
c representa cara y s representa sello. Al lanzar la moneda 100 veces obtenemos una muestra de tamaño $r = 100$, ordenada, con reemplazo, extraída de la población S. El número de resultados posibles es 2^{100}.

2. Se dispone de 3 libros diferentes para ser rifados entre 5 personas. Una misma persona puede ganarse más de un libro.

   Si las personas representan las cinco cajas numeradas y los libros las tres bolas numeradas, el problema corresponde a configuraciones con ocupación múltiple, por lo tanto el número de resultados posibles es $n^r = 53 = 125$.

   En este caso resultó mas sencillo resolver el problema en términos de cajas y bolas, cómo se plantearía en términos de muestreo?

MUESTRAS ORDENADAS SIN REEMPLAZO

Si elegimos r objetos de una población $S = \{a_1, a_2, \cdots, a_n\}$ de n elementos, tomando en cuenta el orden en que son seleccionados, de tal manera que cada objeto pueda ser elegido a lo sumo una vez, diremos que tenemos una muestra ordenada, sin reemplazo, de tamaño r.

Para calcular el número de casos posibles, basta observar que nos encontramos nuevamente delante un problema de arreglos múltiple.

En efecto:

el primer elemento puede elegirse de n maneras diferentes
el segundo elemento puede elegirse de $n-1$ maneras diferentes
$\vdots$
el $r-1$-ésimo elemento puede elegirse de $n-r+2$ maneras diferentes
el r-ésimo elemento puede elegirse de $n-r+1$ maneras diferentes

Obsérvese que

$$n(n-1)(n-2)\cdot\ldots\cdot(n-r+1)=\frac{n!}{(n-r)!}$$

Esta expresión se conoce con el nombre de Variaciones de n elementos tomadas de r en r y la denotaremos por:

$$V_{n,r}=\frac{n!}{(n-r)!}$$

En el muestreo ordenado sin reemplazo, cuando $n=r$, la muestra incluirá toda la población y representará una permutación o reordenamiento de sus elementos. El número total de permutaciones es:

$$P_n=V_{n,n}=\frac{n!}{0!}=n!$$

ya que $0!=1$.

PROBLEMA EQUIVALENTE El muestreo ordenado sin reemplazo, es equivalente al problema de colocar r bolas distinguibles (ó numeradas) en n cajas numeradas, sin permitir ocupación múltiple.

Veamos que existe una biyección entre el número de muestras ordenadas de tamaño r extraída de una población de n elementos, sin reemplazo y el conjunto de configuraciones de bolas de este tipo.

En efecto, una muestra típica tendrá la forma $(a_{j_1},a_{j_2},\cdots,a_{j_r})$ con $1\leq j_k\leq n$, $1\leq k\leq r$. Los j_k son todos diferentes ya que el muestreo se hace sin reemplazo del elemento elegido. Si numeramos las cajas desde 1 hasta n, a esta muestra típica, se le puede asociar la siguiente configuración de bolas:

la bola # 1 se coloca en la caja # j_1
la bola # 2 se coloca en la caja # j_2
$\vdots$
la bola # $r-1$ se coloca en la caja # j_{r-1}
la bola # r se coloca en la caja # j_r

Como la muestra es sin reemplazo, ningún índice j_i estará repetido en la muestra, es decir, a lo sumo una bola puede ser asignada en la caja j_i.

EJEMPLOS:

1. Para calcular la cantidad de números de tres cifras que se pueden formar con los dígitos $\{1,2,3,4,5\}$,si los dígitos deben ser diferentes entre sí, procedemos de la siguiente manera: la población $S=\{1,2,3,4,5\}$, queremos extraer muestras ordenadas de tamaño $r=3$ sin reemplazo. La respuesta a la pregunta es

$$V_{5,3}=\frac{5!}{2!}=60$$

2. **CUMPLEAÑOS**

Los cumpleaños de r personas forman una muestra de tamaño r de todos los días del año. No todos los años tienen una misma longitud y sabemos que las tasas de natalidad no son constantes en distintas épocas del año pero como primera aproximación (para simplificar el problema) podemos considerar el año de 365 días y tomar una selección aleatoria de personas como equivalente a una selección aleatoria de cumpleaños. Con este convenio, el problema se reduce a calcular, el número de configuraciones que se pueden obtener, al colocar r bolas numeradas en $n=365$ cajas numeradas, permitiendo ocupación múltiple, este número es

$$n^r=(365)^r$$

Si queremos considerar solamente el número de cumpleaños diferentes, ya no se permite ocupación múltiple, en este caso, las respuesta es:

$$V_{365,r}=\frac{(365)!}{(365-r)!}$$

De aquí se deduce que, la probabilidad de que todos los cumpleaños sean diferentes, si seleccionamos r personas al azar, es:

$$p = \frac{V_{365,r}}{(365)^r}$$

OBSERVACIÓN: El cociente $\frac{V_{n,r}}{n^r}$ tiende a 1 si $n \to \infty$ y r permanece fijo. Esto significa que para poblaciones de tamaño n muy grande y r "pequeño", una muestra ordenada con ó sin reemplazo, es casi lo mismo. En efecto:

$$\begin{aligned} &\frac{V_{n,r}}{n^r} = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \cdots \cdot \frac{n-r+1}{n} = \\ &1 \cdot (1 - \frac{1}{n}) \cdot (1 - \frac{2}{n}) \cdot \cdots \cdot (1 - \frac{r-1}{n}) \\ &\lim_{n\to\infty} \frac{V_{n,r}}{n^r} = 1. \end{aligned}$$

MUESTRAS NO ORDENADAS SIN REEMPLAZO O SUBPOBLACIONES.

Sea $S = \{a_1, a_2, \cdots, a_n\}$ una población de n elementos, llamamos subpoblación ó muestra no ordenada, sin reemplazo de tamaño r, a todo subconjunto de r elementos, elegidos a partir de S, uno a uno, sin reemplazo del objeto seleccionado.

En este tipo de muestreo, el orden en que se eligen los elementos, no interviene, luego, cada subpoblación de tamaño r corresponde a $r!$ muestras de tamaño r, ordenadas y sin reemplazo, es decir, $V_{n,r}$ es $r!$ veces mayor que el número buscado. Dicho de otra forma, el valor buscado es

$$\frac{V_{n,r}}{r!} = \frac{n!}{(n-r)!r!} = C_{n,r} = \binom{n}{r}$$

que es el coeficiente binomial que aparece en la fórmula del Binomio de Newton

$$(a+b)^n = \sum_{r=0}^{n} C_{n,r} a^r b^{n-r}$$

y representa el número de muestras no ordenadas, de tamaño r elegidas de una población de n elementos, sin reemplazo.

PROBLEMA EQUIVALENTE El muestreo no ordenado, sin reemplazo, puede también ser enunciado en término de cajas y bolas. En efecto, este problema es idéntico al siguiente : de cuántas maneras se pueden colocar r bolas indistinguibles (no numeradas) en n cajas numeradas, sin permitir ocupación múltiple? La siguiente biyección entre el conjunto de subpoblaciones y el de configuraciones admisibles, prueba la equivalencia entre ambos problemas:

A cada subpoblación típica $(a_{j_1}, a_{j_2}, \cdots, a_{j_r})$ con $1 \leq j_k \leq n$, $1 \leq k \leq r$, donde los j_k son todos diferentes ya que el muestreo se hace sin reemplazo del elemento elegido, le asociamos la siguiente configuración de bolas, si numeramos las cajas desde 1 hasta n :

hay una bola en la caja # j_1
hay una bola en la caja # j_2
$\vdots$
hay una bola en la caja # j_{r-1}
hay una bola en la caja # j_r

Como la muestra es sin reemplazo, ningún índice j_i estará repetido en la muestra, es decir, a lo sumo una bola puede ser asignada en la caja j_i. La presencia del elemento j_k en la muestra significa que hay una bola en la caja $\#j_k$.

EJEMPLOS:

1. Se cuenta con 8 personas en el Departamento de Matemática, de la Escuela de Matemática de la Facultad de Ciencias, para elegir entre ellos un comité de 3 personas que constituirán la Comisón Departamental. Esta selección puede ser hecha de varias maneras, para calcular este número, planteamos este problema como un muestreo no ordenado, sin reemplazo, de tamaño $r = 3$ elegido a partir de una población de tamaño $n = 8$. Luego, el número posible de Comités que se pueden formar, es:

$$C_{8,3} = \binom{8}{3} = \frac{8!}{5!3!} = 56.$$

2. Se dispone de 5 afiches idénticos para ser repartidos entre 20 estudiantes, cada estudiante recibirá a lo sumo un afiche. Por supuesto, al distribuirlos, quedarán 15 estudiantes sin afiche. Para calcular el número de resultados posibles, pensaremos los estudiantes como las $n = 20$ cajas numeradas y los afiches idénticos como las $r = 5$ bolas indistinguibles. La respuesta al problema es:

$$C_{20,5} = \binom{20}{5} = \frac{20!}{15!5!} = 15504.$$

ORDENACIONES EN QUE INTERVIENEN DOS CLASES DE ELEMENTOS. Supongamos que tenemos un conjunto constituido por p elementos del tipo α y q elementos del tipo β. Queremos calcular el número de sucesiones distinguibles de tamaño $p + q$ que podemos formar con estos elementos. Para ello nos valemos del siguiente artificio: numeramos todos lo elementos α desde 1 hasta p y los elementos β desde $p + 1$ hasta $p + q$, obteniendo así una sucesión de $p + q$ elementos diferentes. El número de sucesiones de este tipo que podemos formar es $(p + q)!$. A cada sucesión de elementos α y β sin numerar le corresponden $p!q!$ sucesiones numeradas, es decir, el número de sucesiones numeradas es $p!q!$ veces mayor que el número de sucesiones no numeradas, por lo tanto el número de sucesiones no numeradas es

$$\frac{(p+q)!}{p!q!} = \binom{p+q}{q} = C_{p+q,q} = \binom{p+q}{p} = C_{p+q,p}$$

MUESTRAS NO ORDENADAS CON REEMPLAZO:

Sea $S = \{a_1, a_2, \cdots, a_n\}$ una población de n elementos, llamamos muestra no ordenada, con reemplazo de tamaño r, a todo conjunto de r elementos, elegidos a partir de S, uno a uno, reemplazando cada objeto seleccionado inmediatamente después de su elección, sin importar el orden en que los objetos quedan elegidos. (Los objetos pueden aparecer repetidos en la muestra).

Nos interesa el cardinal ó número de elementos del espacio muestral asociado a esta experiencia. Como en los casos anteriores, este problema es equivalente al de distribuir r bolas en n cajas, pero en este caso, resulta menos complicado resolver este último problema para calcular el cardinal del espacio muestral correspondiente a la experiencia de muestreo enunciada.

PROBLEMA EQUIVALENTE Se tienen n cajas numeradas y r bolas no numeradas o indistinguibles. Queremos distribuir estas bolas en las cajas permitiendo ocupación múltiple.

Para convencernos de la equivalencia entre ambos problemas vamos a poner en evidencia la biyección entre el espacio muestral correspondiente al muestreo y el conjunto de configuraciones de bolas.

A cada subpoblación típica $(a_{j_1}, a_{j_2}, \cdots, a_{j_r})$ con $1 \leq j_k \leq n$, $1 \leq k \leq r$, donde los j_k pueden repetirse ya que el muestreo se hace con reemplazo del elemento elegido, le asociamos la siguiente configuración de bolas, al numerar las cajas desde 1 hasta n :

hay una bola en la caja # j_1
hay una bola en la caja # j_2
$\vdots$
hay una bola en la caja # j_{r-1}
hay una bola en la caja # j_r

Como la muestra es con reemplazo, algunos de los índices j_i estarán repetidos en la muestra, es decir, puede haber más de una bola en la caja j_i, en los casos en que el índice esté repetido en la muestra, habrá una bola cada vez que el índice aparece.

La muestra estará constituida por r_i objetos a_i, $i = 1, 2, \cdots, n$; con $\sum_{i=1}^{n} r_i = r$. Si $r < n$, entonces al menos $(n - r)$ de estos r_i deben ser nulos. En este tipo de muestreo r puede ser

$$r < n \text{ ó } r \geq n$$

Decir que en la muestra hay r_i elementos a_i significa que en la caja $\# i$ hay r_i bolas.

Resolvamos el problema del número de configuraciones posibles en este caso: representemos las bolas por * y las cajas por los n espacios comprendidos entre $(n + 1)$ barras. Una configuración posible de * y | sería:

$$| * * | * | * * * * | \cdots\cdots | * | \cdots | * * | * |$$

Toda configuración comienza y termina con una barra, y el resto estará formado por sucesiones de $(n-1)$ barras (|)y r estrellas ($*$), combinadas en un orden arbitario.

Este problema es idéntico al de ordenaciones en que intervienen dos clases de elementos, podemos pensar que las $(n - 1)$ barras representan los $p = n - 1$ elementos del tipo α y las r estrellas representan los $q = r$ elementos del tipo β. El número de sucesiones posibles que podemos hacer con ellos, es:

$$A_{n,r} = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

este número combinatorio responde la pregunta formulada.

Se concluye así que, el número de muestras no ordenadas, de tamaño r, con reemplazo, que se pueden elegir a partir de una población de tamaño n, es

$$A_{n,r} = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}.$$

EJEMPLOS:

1-. Se quieren rifar 5 lápices idénticos entre 8 niños, permitiendo que cualquiera de ellos se pueda ganar más de un lápiz.

Para calcular el número de resultados posibles, consideramos lo $r = 5$ lápices como las bolas idénticas y los $n = 8$ niños como las cajas numeradas. El número que queremos calcular es

$$A_{8,5} = \binom{12}{5} = \frac{12!}{5!7!} = 792.$$

2-. Se lanzan 8 dados indistinguibles, queremos calcular el número de resultados diferentes en esta experiencia. La población es $S = \{1, 2, 3, 4, 5, 6\}$, queremos extraer un muestra no ordenada con reemplazo de tamaño $r = 8$. La respuesta a la pregunta es:

$$A_{6,8} = \binom{13}{8} = \frac{13!}{8!5!} = 1287.$$

PARTICIONES

Sea $S = \{a_1, a_2, \cdots, a_n\}$ un conjunto de n elementos, $A_k \subseteq S$, $k = 1, 2, \cdots, m$. Diremos que estos subconjuntos de S, forman una partición de S, si $A_i \cap Aj = \emptyset$ si i es diferente de j y

$$S = \bigcup_{k=1}^{m} A_k.$$

PARTICIONES EN DOS SUBCONJUNTOS

Sabemos que si $S = \{a_1, a_2, \cdots, a_n\}$ es una población de n objetos, el número de subpoblaciones ó muestras no ordenadas de tamaño r, sin reemplazo, es

$$C_{n,r} = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Esta cantidad representa el número de particiones del conjunto S, en dos subconjuntos complementarios, de r y $(n-r)$ elementos respectivamente. Por lo tanto el número de particiones posibles, en dos subconjuntos de cardinales r y $(n-r)$ respectivamente es

$$C_{n,r} = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

PARTICIONES EN K SUBCONJUNTOS:

El número de particiones diferentes de una población $S = \{a_1, a_2, \cdots, a_n\}$ de n elementos, en k subconjuntos de cardinales $r_1, r_2, \cdots, r_k$ respectivamente, es

$$\frac{n!}{r_1!r_2!\cdots r_k!}$$

con $r_1 + r_2 + \cdots + r_k = n$, $r_i \geq 0, i = 1, 2, \cdots, k$.

Este coeficiente se llama coeficiente multinomial y aparece en la siguiente fórmula algebráica:

$$(u_1 + u_2 + \cdots + u_k)^n = \sum_{r_1+r_2+\cdots+r_k=n} \frac{n!}{r_1!r_2!\cdots r_k!} u_1^{r_1} u_2^{r_2} \cdots u_k^{r_k}.$$

Para demostrar el resultado enunciado, seleccionamos de la población r_1 objetos, esta operación se puede hacer de $\binom{n}{r_1}$ maneras diferentes. El segundo conjunto de la partición se elige entre los $n - r_1$ elementos restantes de la población y como el cardinal es r_2, esta última operación se hará de $\binom{n-r_1}{r_2}$ maneras diferentes. Así sucesivamente, el $(k-1)$ ésimo grupo de r_{k-1} elementos se seleccionará entre los

$$n - r_1 - r_2 - \cdots - r_{k-2}$$

elementos que quedan en la población, de

$$\binom{n - r_1 - r_2 - \cdots - r_{k-2}}{r_{k-1}}$$

maneras posibles. El k ésimo grupo de r_k elementos sólo podrá ser seleccionado de una forma, ya que sólo quedan por seleccionar

$$n - r_1 - r_2 - \cdots - r_{k-2} - r_{k-1} = r_k$$

elementos en S. Así, la respuesta al problema planteado es:

$$\binom{n}{r_1}\binom{n-r_1}{r_2}\binom{n-r_1-r_2}{r_3}\cdots\binom{n-r_1-r_2-\cdots-r_{k-2}}{r_{k-1}}$$

simplificando esta expresión obtenemos el coeficiente multinomial:

$$\frac{n!}{r_1!r_2!\cdots r_k!}$$

con $r_1 + r_2 + \cdots + r_k = n$, $r_i \geq 0, i = 1, 2, \cdots, k.$.

DISTRIBUCIÓN HIPERGEOMÉTRICA

Muchos problemas combinatorios se reducen a la siguiente forma: en una población de n elementos, n_1 son rojos y $n_2 = n - n_1$ son negros; se elige aleatoriamente un grupo de r elementos, si pedimos que esta muestra contenga exactamente k elementos rojos, con $0 \leq k \leq n_1$, $k \leq r$, queremos saber de cuántas maneras posibles podemos seleccionarla. Para hallar esta cantidad, observemos que la muestra que elegiremos, contendrá exactamente k elementos rojos y $r - k$ elementos negros. Los rojos se pueden elegir de $\binom{n_1}{k}$ maneras diferentes y los negros de $\binom{n_2}{r-k}$ maneras diferentes, por lo tanto el número que nos interesa es

$$\binom{n_1}{k} \cdot \binom{n_2}{r-k}$$

con $n_2 = n - n_1$, $0 \leq k \leq n_1 \wedge r$, donde

$$n_1 \wedge r = \begin{cases} n_1 \text{ si } r \geq n_1 \\ r \text{ si } r \leq n_1 \end{cases}$$

Si queremos calcular la probabilidad de elegir una muestra no ordenada sin reemplazo, que contenga exactamente k elementos rojos, con

$$0 \leq k \leq n_1,\ k \leq r,$$

debemos dividir entre el cardinal del espacio muestral asociado a esta experiencia, a saber

$$p_k = \frac{\binom{n_1}{k} \cdot \binom{n_2}{r-k}}{\binom{n}{r}}$$

EJEMPLO:

Se cuenta con tres bolas distintas para ser colocadas al azar en cuatro cajas numeradas, permitiendo ocupación múltiple.

Se quiere calcular la probabilidad de que quede una bola en la primera caja y dos bolas en la tercera caja. El cardinal del espacio muestral es

$$\#\Omega = 4^3 = 64$$

El eventos que nos interesa está representado en el espacio muestral, por varios eventos elementales, ya que las bolas están numeradas. (Si las bolas no estuviesen numeradas, cuál es el cardinal de este evento?). El cardinal del evento que nos interesa se puede calcular pensando el conjunto de bolas como una población

$$S = \{1, 2, 3\}$$

de tres elementos (las bolas numeradas) y lo que se quiere calcular es el número de particiones en dos subconjuntos de cardinales 1 y 2 respectivamente. Sabemos que la respuesta es $C_{3,1} = \binom{3}{1} = 3$ ($n = 3$, $r = 1$, $n - r = 2$).

La probabilidad del eventos es:

$$p = \frac{3}{64}.$$

GENERALIZACIÓN:

Se tienen r bolas numeradas y n cajas numeradas, cual es la probabilidad de que

$$\begin{array}{l} \text{haya } r_1 \text{ bolas en la caja } \#\ j_1 \\ \text{haya } r_2 \text{ bolas en la caja } \#\ j_2 \\ \text{haya } r_3 \text{ bolas en la caja } \#\ j_3 \\ \vdots \\ \text{haya } r_k \text{ bolas en la caja } \#\ j_k \end{array}$$

con $r_1 + r_2 + \cdots + r_k = r$, $r_i \geq 0$, $1 \leq j_i \leq n$, $i = 1, 2, \cdots, k$.

El número de casos posibles o cardinal del espacio muestral asociado al experimento, es n^r.

Para calcular el número de casos favorables, identificamos las bolas numeradas con la población

$$S = \{1, 2, 3, \cdots, r\}$$

de r elementos, las k cajas las identificamos con k subconjuntos de la población S de $r_1, r_2, \cdots, r_k$ elementos respectivamente. El número de casos favorables es

$$\frac{r!}{r_1!r_2!\cdots r_k!}$$

y la probabilidad de este evento es

$$\frac{r!}{n^r \cdot r_1!r_2!\cdots r_k!}$$

CONTROL DE CALIDAD. De una población de N artículos, entre los que hay n defectuosos, se extraen sucesivamente r artículos sin reemplazo. El espacio muestral Ω correspondiente tiene

$$\#\Omega = \binom{N}{r}$$

La probabilidad de que en la muestra se encuentren exactamente k artículos defectuosos, con $0 \leq k \leq n \wedge r$, es

$$p_k = \frac{\binom{n}{k} \cdot \binom{N-n}{r-k}}{\binom{N}{r}}$$

ESTIMACIÓN DE N Como aplicación de la distribución hipergeométrica, consideremos el siguiente ejemplo. Supongamos que queremos estimar el número total N de peces en un lago. Este es un problema característico de estimación estadística, describir los métodos que un estadístico moderno usaría nos alejaría de los objetivos de este curso, nos limitaremos a enseñar que la distribución hipergeométrica nos da un indicio de la solución de este problema.

Los investigadores proceden del siguiente modo: extraen una muestra de tamaño n y marcan los peces antes de reintegrarlos al agua. Posteriormente se extrae una muestra de tamaño $r = n$ y se cuenta el número k de peces marcados, $0 \leq k \leq n$.

Para simplificar, se supone que el número total N no ha variado y que todos los peces tienen la misma probabilidad de ser seleccionados.

Bajo estas suposiciones, la probabilidad de seleccionar k peces marcados, es

$$p_k = L_N(k) = \frac{\binom{n}{k} \cdot \binom{N-n}{n-k}}{\binom{N}{n}}$$

Vamos a determinar N de manera que $L_N(k)$ sea máximo. Con una simplificación elemental, obtenemos que

$$\frac{L_N(k)}{L_{N-1}(k)} = \frac{(N-n)^2}{N(N-2n+k)}.$$

Esta fracción

$$\frac{L_N(k)}{L_{N-1}(k)} \begin{cases} < 1 \text{ si } n^2 < Nk \\ > 1 \text{ si } n^2 > Nk \end{cases}$$

pues, las parábolas (como función de N) $(N-n)^2$ y $N(N-2n+k)$, se intersectan en el punto $N = \frac{n^2}{k}$, por lo tanto, $L_N(k)$ como función de N crece primeramente y luego decrece, alcanzando el máximo cuando N es el mayor entero tal que $N < \dfrac{n^2}{k}$.

En particular, si $n = 1000$ (número de peces marcados) y $k = 100$ (número de peces pintados cuando se selecciona la muestra después de haber marcado n peces en el lago), resulta que el valor estimado $\hat{N}$ de N es

$$\hat{N} = \frac{10^6}{10^2} = 10^4 = 10000$$

ORDENACIONES CON K CLASES DE ELEMENTOS. Supongamos que tenemos un conjunto S constituido por n_1 elementos a_1, n_2 elementos a_2, n_3 elementos $a_3, \cdots$, n_k elementos a_k, con $n_1+n_2+\cdots+n_k = n$; queremos calcular el número de sucesiones que podemos construir con todos ellos. Si numeramos todos los elementos, tendríamos n elementos diferentes y $n!$ sucesiones diferentes que podríamos construir con todos ellos. Por cada sucesión no numerada se tienen $n_1! \cdot n_2! \cdot \cdots \cdot n_k!$ sucesiones numeradas, así el número de sucesiones del tipo planteado que podemos construir es

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

Observe que este número es idéntico al número de particiones de un conjunto de n elementos, en k subconjuntos de $n_1, n_2, \cdots, n_k$ elementos respectivamente, donde $n_1 + n_2 + \cdots + n_k = n$.

El modelo anterior puede expresarse en términos de particiones, si identificamos las posiciones de los diferentes elementos en la sucesión, con los números $1, 2, 3, \cdots, n$.

Si

$$S = \{1, 2, 3, \cdots, n\}$$

$\#A_1 = n_1, \#A_2 = n_2, \cdots, \#A_k = n_k$, $S = A_1 \cup A_2 \cup \cdots \cup A_k$, $A_i \cap A_j = \emptyset$ si i es diferente de j, podemos identificar A_i con la posición de los elementos de tipo a_i en la sucesión.

EJEMPLO: Se tiene una población constituida por las letras que forman la palabra PROBABILIDAD, a saber

$$S = \{\text{PROBBIIAALDD}\},$$

se quiere calcular la probabilidad de obtener la palabra PROBABILIDAD al ordenar aleatoriamente las letras que aparecen en el conjunto S.

Esta población está contiene 8 tipos de elementos:

uno del tipo P
uno del tipo R
uno del tipo O
uno del tipo L
dos del tipo B
dos del tipo I
dos del tipo A
dos del tipo D

En total tenemos 12 elementos, el número de sucesiones distinguibles es

$$\frac{12!}{2!2!2!2!}$$

y sólo una de ellas es favorable, por lo tanto la probabilidad de obtener la palabra probabilidad es

$$\frac{2!2!2!2!}{12!} = 3,340281 \cdot 10^{-8}$$

EJERCICIOS PARA LA PRÁCTICA

1. En un juego de barajas de 52 cartas, se extraen al azar, 4 de ellas, sin reposición. Determine la probabilidad de los siguientes eventos:

   a) Por lo menos una carta es un As.

   b) Las cuatro cartas tienen el mismo color.

2. Se reparten al azar 5 bolas distinguibles en 3 casillas. Determine la probabilidad de que por lo menos halla una bola en cada casilla.

3. Cuántos conjuntos diferentes de iniciales pueden formarse, si cada persona tiene un apellido y
    a) exactamente 2 nombres
    b) no más de 2 nombres
    c) no más de 3 nombres
    (Suponga que el abecedario tiene 26 letras).
4. De cuántas maneras pueden colocarse n bolas idénticas en n cajas numeradas de tal modo que exactamente 2 cajas queden vacías? y que quede exactamente una caja vacía?.
5. Los números $1, 2, 3, \cdots, n$ están colocados en un orden aleatorio, encuentre la probabilidad de que los dígitos
    a) 1 y 2 estén en ese orden y juntos.
    b) $1, 2$ y 3 estén en ese orden y juntos.
6. Encontrar la probabilidad que en el Triple de la Lotería del Zulia salgan exactamente
    a) tres números diferentes
    b) dos números diferentes
    c) los tres números iguales
7. Si en lugar del Triple, el número premiado es de cuatro cifras, como ocurre en otras loterías, cuál es la probabilidad que
    a) todos sean diferentes?
    b) sólo tres sean diferentes?
    c) sólo dos sean diferentes?
    d) todos sean iguales?
8. Si se colocan aleatoriamente n bolas distinguibles en n cajas numeradas, encuentre la probabilidad que exactamente una caja quede vacía.
9. Encontrar la probabilidad p_r que en una muestra de r dígitos aleatorios, no hayan dos dígitos iguales. Estime el valor p_{10} mediante la fórmula de Stirling (para una demostración de esta fórmula, ver W.Feller, cap.1, Introducción a la teoría de probabilidades y sus aplicaciones)

$$n! \simeq \sqrt{2\pi} n^{n+1/2} e^{-n}$$

10. Se tienen $2k$ cajas numeradas (desde 1 hasta $2k$) y k bolas numeradas (desde 1 hasta k). Se distribuyen las k bolas en las $2k$ cajas de modo que cada caja contenga a lo sumo una bola.
    a) calcular el número de configuraciones tales que la caja #1 resulte ocupada
    b) calcular el número de configuraciones tal que la caja #1 resulte desocupada
    c) calcular el número de configuraciones tales que la caja #1 resulte desocupada y la caja #2 contenga la bola #2.
11. Un ascensor con 7 pasajeros se detiene en 10 pisos. Cuál es la probabilidad que 2 pasajeros no se bajen en el mismo piso?
12. Si se colocan n bolas numeradas en n cajas numeradas, cuál es la probabilidad que cada caja contenga una bola?
13. Cuál es la probabilidad que entre n personas haya al menos dos que cumplan año el mismo día?

14. Se disponen en fila 3 bolas blancas y 6 bolas negras de manera que no haya 2 bolas blancas consecutivas. Cuántas maneras hay de hacerlo?

15. Se lanza una moneda equilibrada 10 veces consecutivas,

    a) cuál es la probabilidad que se obtengan 10 caras?
    b) cuál es la probabilidad que en un décimo primer lanzamiento se obtenga un sello?
    c) cuál es la probabilidad que los tres primeros lanzamientos sean cara y los otros siete sean sello?
    d) cuál es la probabilidad de obtener 3 caras y 7 sellos?

16. Se lanzan 6 dados distinguibles

    a) de cuántas maneras pueden salir sus caras superiores?
    b) de cuántas maneras puede ocurrir que sus caras superiores sean todas distintas?
    c) cuál es la probabilidad que todas sus caras superiores sean iguales?
    d) cuál es la probabilidad que al menos dos dados, muestren sus caras superiores iguales?

17. Qué pasaría en el problema anterior si los dados no son distinguibles?

18. Cuántos enteros hay entre 10^6 (un millón) y 10^7 (diez millones), es decir en el intervalo $\left(10^6, 10^7\right)$.

    a) que no tengan dos dígitos consecutivos iguales
    b) que tengan todos sus dígitos diferentes.

19. En un paquete de barajas de Póker, bien barajado, cuál es la probabilidad que los 4 ases queden situados consecutivamente, al formar una sucesión aleatoria con todas ellas?

20. Consideremos el conjunto de todas las poligonales (n, S_n) con $n = 0, 1, 2, \cdots$, que parten del origen, es decir, $S_0 = 0$ y que en cada paso saltan una unidad hacia arriba o una unidad hacia abajo, explícitamente,

    $$S_n = X_1 + X_2 + \cdots + X_n,$$

    donde $X_k = 1$ ó -1 ($k = 1, 2, \cdots, n$). S_n así definido, es un paseo al azar ó caminata aleatoria.

    a) cuántas poligonales podemos construir en el intervalo de tiempo $[0, n]$?
    b) cuántas poligonales son tales que $S_n = 0$?

21. Dos cajas contienen $2n$ bolas cada una, numeradas desde 1 hasta $2n$. Selecionamos n bolas de cada caja. Calcular la probabilidad que haya a lo sumo una bola con el mismo número, en ambos conjuntos.

22. En un cesto de 550 manzanas el 20% está en mal estado. Calcule la probabilidad que una muestra de 25 manzanas, tomadas al azar, contengan dos manzanas en mal estado.

23. Se tiene un lote de N productos de los cuales n defectuosos, hallar la probabilidad que en una muestra aleatoria de tamaño k se encuentren

    a) j piezas defectuosas, $j \leq n$ y $j \leq k$.
    b) al menos j piezas defectuosas, $j \leq n$ y $j \leq k$.
    c) a lo sumo j piezas defectuosas, $j \leq n$ y $j \leq k$.

24. Se tienen 6 palos de igual longitud y cada uno de ellos se parte en uno largo y uno corto, con los 12 palos así obtenidos, se vuelven a formar 6, reagrupándolos al azar, en parejas.

    a) calcular la probabilidad de juntar los trozos en la forma original
    b) calcular la probabilidad que cada trozo largo sea juntado con uno corto

RESPUESTAS DE LOS PROBLEMAS.

1. $CARD(\Omega) = \binom{52}{4}$

   a) $1 - \dfrac{\binom{48}{4}}{\binom{52}{4}}$

   b) $4 \cdot \dfrac{\binom{13}{4}}{\binom{52}{4}}$

2. $150/3^5 = 3.\frac{5!}{2!2!1!}\frac{1}{3^5} + 3.\frac{5!}{3!1!1!}\frac{1}{3^5}$
   Este problema se ha resuelto considerando la población $\{1,2,3,4,5\}$ y hallando el número de particiones en tres subconjuntos $A_1, A_2, A_3, card(A_1) = 2, card(A_2) = 2, card(A_3) = 1$, por otro lado, el número de particiones en tres subconjuntos $A_1, A_2, A_3, card(A_1) = 3, card(A_2) = 1, card(A_3) = 1$, de cada uno de ellos hay 3.

3. Respuestas:

   a) 26^3

   b) $26^2 + 26^3$

   c) $26^2 + 26^3 + 26^4$

4. Para dos cajas vacías: $\binom{n}{2}\binom{n-1}{2}$, para una caja vacía $\binom{n}{1}\binom{n-1}{1}$

5. $CARD(\Omega) = n!$

   a) Para que el 1 y 2 estén en ese orden y juntos hay

   $$(n-2)!(n-1) = (n-1)!$$

   posibilidades, por lo tanto, la probabilidad del evento es:

   $$\frac{(n-1)!}{n!} = \frac{1}{n}$$

   b) Para que los dígitos 123, estén en ese orden y juntos, se tienen

   $$(n-3)!(n-2) = (n-2)!$$

   posibilidades, por lo tanto, la probabilidad del evento es:

   $$\frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

6. $CARD(\Omega) = 10^3$

   a) $CARD\,\{\text{los tres números son diferentes}\} = \#A = 10 \cdot 9 \cdot 8$

   $$\mathbb{P}(A) = 0,72$$

   b) $CARD\,\{\text{dos números son diferentes}\} = \#B = 10 \cdot 9 \cdot 3$

   $$\mathbb{P}(B) = 0,27$$

   Se puede resolver como colocación de 3 bolas numeradas en 10 cajas numeradas, permitiendo ocupación múltiple. El evento de interés es

   $$\left\{ \begin{array}{c} \text{dos cajas tienen una bola cada una y la tercera bola} \\ \text{está en cualquiera de ellas dos} \end{array} \right\}$$

El cardinal de este evento es

$$\binom{10}{2}\frac{3!}{2!1!}\cdot 2 = 10\cdot 9\cdot 3$$

ya que las cajas se eligen de $\binom{10}{2}$ maneras, luego se construyen particiones de $\{1,2,3\}$ en dos subconjuntos A_1 y A_2 tales que $\#A_1 = 1, \#A_2 = 2$ y viceversa.

c) $CARD\,\{\text{los tres números son iguales}\} = \#C = 10$

$$\mathbb{P}(C) = 0,01$$

Observemos que $A\cup B\cup C = \Omega$ con A, B, C dos a dos disjuntos.

7. $CARD(\Omega) = 10^4$

a) $CARD\,\{\text{los cuatro números son diferentes}\} = \#A = 10\cdot 9\cdot 8\cdot 7$

$$\mathbb{P}(A) = 0,504$$

b) $CARD\,\{\text{tres números son diferentes}\} = \#B$

$$\#B = 10\cdot 9\cdot 8\cdot 6$$

$$\mathbb{P}(B) = \frac{10\cdot 9\cdot 8\cdot 6}{10^4} = 0,432$$

En efecto, podemos pensar en distribuir 4 bolas numeradas en 10 cajas numeradas permitiendo ocupación múltiple. El evento que nos interesa es

$$\left\{\begin{array}{c}\text{Tres de las cajas tienen una bola cada una y}\\ \text{la } 4^a \text{ bola está en cualquiera de ellas}\end{array}\right\}$$

El cardinal de este evento es

$$\binom{10}{3}\binom{3}{1}\frac{4!}{1!1!2!} = 10\cdot 9\cdot 8\cdot 6$$

ya que las cajas se eligen de $\binom{10}{3}$ maneras, luego se hacen particiones de $\{1,2,3,4\}$ en 3 subconjuntos de cardinales 1,1 y 2 respectivamente, permutando los cardinales de $\binom{3}{1}$ maneras diferentes.

c) $CARD\,\{\text{dos números son diferentes}\} = \#C$

$$\#C = 10\cdot 9\cdot 7$$

de dónde

$$\mathbb{P}(C) = 0,063$$

Se puede resolver expresando el evento que nos interesa como

$$\left\{\begin{array}{c}\text{dos bolas en dos cajas diferentes y las otras dos}\\ \text{en cualquiera de esas dos}\end{array}\right\}$$

el cardinal de este conjunto es

$$\binom{10}{2}\left(2\cdot\frac{4!}{1!3!} + \frac{4!}{2!2!}\right) = 10\cdot 9\cdot 7$$

ya que las cajas se eligen de $\binom{10}{2}$ maneras, luego se hacen particiones de $\{1,2,3,4\}$ en 2 subconjuntos de cardinales 1 y 3; 3 y 1; 2 y 2 respectivamente.

d) $CARD\{\text{los cuatro números son iguales}\} = \#D = 10$

$$\mathbb{P}(D) = 0,001$$

Observemos que

$$\Omega = A \cup B \cup C \cup D$$

con A, B, C, D dos a dos disjuntos.

8. $A = \{\text{exactamente hay una caja vacía}\}$

$$\#A = n\binom{n-1}{1}\frac{n!}{2!\cdot 1!\cdot 1!\cdots 1!}$$

(hay un sólo 2 y $n-2$ unos). Puesto que $\#\Omega = n^n$

$$\mathbb{P}(A) = \frac{n!}{n^n}\binom{n}{2}$$

9. $CARD(\Omega) = 10^r$

$$A = \{\text{que no hayan dos dígitos iguales}\} = \{\text{los } r \text{ dígitos son distintos}\}$$

$$\#A = V_{10,r} = \frac{10!}{(10-r)!}$$
$$\mathbb{P}(A) = \frac{10!}{(10-r)!10^r} = p_r$$
$$p_{10} = \frac{10!}{10^{10}} \simeq \sqrt{20\pi}e^{-10} \simeq 0,0003598.$$

10. Respuestas:

a) $k \cdot V_{2k-1,k-1} = \dfrac{(2k-1)!}{(k-1)!}$

b) $V_{2k-1,k} = \dfrac{(2k-1)!}{(k-1)!}$

c) $V_{2k-2,k-1} = \dfrac{(2k-2)!}{(k-1)!}$

11. Los pisos son las $n = 10$ cajas numeradas y los pasajeros son las $r = 7$ bolas numeradas. Se permite ocupación múltiple:

$$A = \{\text{dos pasajeros no se bajan en un mismo piso}\}$$
$$\#A = V_{10,7}, \#\Omega = 10^7$$
$$\mathbb{P}(A) = \frac{10!}{7!10^7} = 720.10^{-7}$$

12. $\mathbb{P}(A) = \dfrac{n!}{n^n}$

13. $\mathbb{P}(A) = 1 - \dfrac{V_{365,n}}{(365)^n}$

14. $\binom{7}{3} = 35$

15. $CARD(\Omega) = 2^{10}$

a) $\dfrac{1}{2^{10}}$

b) $\dfrac{2^{10}}{2^{11}} = \dfrac{1}{2}$

c) $\dfrac{1}{2^{10}}$

d) $\dfrac{\binom{10}{3}}{2^{10}} = \dfrac{120}{2^{10}}$

16. $CARD(\Omega) = 6^6$

a) 6^6

b) $6!$

c) $\dfrac{6}{6^6} = \dfrac{1}{6^5}$

d) $1 - \dfrac{6!}{6^6}$

17. $S = \{1, 2, 3, 4, 5, 6\}, r = 6$. Corresponde a un muestreo no ordenado con reposición.

a) $\binom{11}{6} = 462$

b) 1

c) $\frac{6}{462} \simeq 0.013$

d) $1 - \dfrac{1}{462} = 0.9928$

18. Un número con estas condiciones solamente puede tener 7 cifras, pues es estrictamente menor que 10^7 y es mayor que 10^6.

a) 9^7

b) $9 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = 544320$.

19. $\dfrac{4!49!}{52!}$

20. Respuestas:

a) $S_n = \sum_{i=1}^{n} X_i$ con $X_i = \pm 1$. El Cardinal del espacio muestral es el mismo que corresponde a elegir una muestra de tamaño $r = n$, ordenda con reemplazo a partir de una población $S = \{-1, 1\}$ de dos elementos, así

$$Card(\Omega) = 2^n.$$

b) $CARD\,\{\text{poligonales tales que } S_n = 0\} = 0$ si n es impar. Si n es par $n = 2k$ y habrán k variables $X_i = 1$ y k variables $X_i = -1$. Por lo tanto la respuesta, para n par, es

$$\binom{n}{n/2}$$

21. $A = A_0 \cup A_1$

$$A_0 = \{\text{ninguna bola igual}\}$$
$$A_1 = \{\text{Una bola igual}\}$$

$$\mathbb{P}(A) = \frac{1 + n^2}{\binom{2n}{n}}$$

22. La población S está constituida por 550 manzanas de las cuales el 20% está en mal estado. Si A es el evento que nos interesa

$$\mathbb{P}(A) = \frac{\binom{110}{2}\binom{440}{23}}{\binom{550}{25}}$$

23. $CARD(\Omega) = \binom{N}{k}$

  a) $A = \{j \text{ piezas defectuosas}\}$

$$\mathbb{P}(A) = \frac{\binom{n}{j}\binom{N-n}{k-j}}{\binom{N}{k}}$$

  b) $B = \{\text{al menos } j \text{ piezas defectuosas}\}$

$$\mathbb{P}(B) = \sum_{i=j}^{k\wedge n} \frac{\binom{n}{i}\binom{N-n}{k-i}}{\binom{N}{k}}$$

  c) $C = \{\text{a lo sumo } j \text{ piezas defectuosas}\}$

$$\mathbb{P}(C) = \frac{\sum_{i=0}^{j}\binom{n}{i}\binom{N-n}{k-i}}{\binom{N}{k}}$$

24. Numeramos los doce palitos de tal forma que los largos sean impares y los cortos pares, la unión original corresponderá a 1 con $2, 3$ con $4, 5$ con 6,etc.

    Considere el conjunto $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ y las particiones posibles en seis subconjuntos de cardinal dos para cada uno de ellos. Este resultado es conocido

$$\frac{12!}{2^6}$$

    pero en nuestro caso las 6! maneras

$$\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}, \{9, 10\}, \{11, 12\}$$

    son indistinguibles, así

$$\#\Omega = \frac{12!}{2^6 6!}, a)\frac{2^6 6!}{12!}, b)\frac{2^6 6! 6!}{12!}$$

Capítulo 3

PROBABILIDAD CONDICIONAL.INDEPENDENCIA.

INTRODUCCIÓN:

En este capítulo estudiaremos dos conceptos importantes de la teoría de Probabilidad, como son, la probabilidad condicional y la independencia entre eventos.

Cuando se observan los resultados de un experimento aleatorio, con frecuencia nos interesa saber de qué manera el resultado de cierto evento A es influenciado por el de otro evento B.

Para caracterizar la relación entre A y B introducimos el concepto de probabilidad condicional de A, dado que el evento B ha ocurrido. Esta probabilidad puede ser distinta que la probabilidad de A y refleja de qué manera la información adicional que tenemos (el evento B ha ocurrido) altera la probabilidad de que suceda A. Cuando ambas probabilidades son iguales, diremos que los eventos A y B son independientes.

A continuación, definiremos ambos conceptos con precisión, pero antes vamos a considerar un ejemplo para ilustrar la situación:

Consideremos una población de 10 artículos entre los cuales hay 7 en buen estado, que llamaremos del tipo A y 3 en mal estado que llamaremos del tipo B. De esta población se extraen dos artículos al azar, sin reemplazo y se consideran los siguientes eventos:

$$\begin{aligned} E_1 &= \text{ el primer artículo extraído está en buen estado} \\ E_2 &= \text{ el segundo artículo extraído está en buen estado} \end{aligned}$$

El espacio muestral

$$\Omega = \{(a_i, a_j), (a_i, b_k), (b_k, a_i), (b_m, b_n)\}$$

donde a_i significa que el ariículo extraído es el i-ésimo artículo en buen estado (ó de tipo A), de la población y b_j significa que el artículo extraído es el j-ésimo artículo en mal estado (ó de tipo B), de la población. Puesto que el muestreo es sin reemplazo, i es diferente de j y m es diferente de n, con

$$\begin{aligned} 1 \le i, j \le 7 \\ 1 \le m, n, k \le 3 \end{aligned}$$

El cardinal de Ω es $V_{10,2} = \dfrac{10!}{8!} = 90$.

La siguiente tabla de doble entrada, indica el número de eventos elementales correspondientes a la partición de Ω según los eventos E_1, E_2 y sus complementos:

$\cap$	E_2	E_2^c	Ω
E_1	7×6	7×3	7×9
E_1^c	3×7	3×2	3×9
Ω	7×9	3×9	10×9

PROBLEMA: si suponemos que en la primera extracción, el artículo está en buen estado, cuál es la probabilidad de que en la segunda extracción, el artículo esté en buen estado?.

Dicho de otra forma, si restringimos nuestra atención al subconjunto de Ω cuyos eventos cumplen la condición: "el primer artículo extraído está en buen estado" o equivalentemente, consideramos los eventos de la forma $E_1 \cap A$ siendo A un evento cualquiera de Ω, lo que nos preguntamos es: cuál es la probabilidad de que ocurra $E_1 \cap E_2$ siendo el nuevo espacio muestral $E_1 \cap \Omega = E_1$, la respuesta a la pregunta es:

$$\frac{\#(E_1 \cap E_2)}{\#(E_1)} = \frac{7 \times 6}{7 \times 9} = \frac{7 \times 6/10 \times 9}{7 \times 9/10 \times 9} = \frac{\mathbb{P}(E_1 \cap E_2)}{\mathbb{P}(E_1)}$$

Por lo tanto, la probabilidad de obtener un artículo en buen estado, en la segunda extracción, sabiendo que en la primera se obtuvo un artículo en buen estado, es:

$$\frac{\mathbb{P}(E_1 \cap E_2)}{\mathbb{P}(E_1)} = \frac{6}{9} = \frac{2}{3}$$

La probabilidad que hemos calculado, se llama Probabilidad Condicional de E_2 dado E_1 y se denota por

$$\mathbb{P}(E_2/E_1).$$

Observemos que $\mathbb{P}(E_2) = \dfrac{7 \times 9}{10 \times 9} = \dfrac{7}{10}$ y que $\mathbb{P}(E_2/E_1) = \dfrac{2}{3}$ no coinciden. Lo que sucede es que al saber que ya ocurrió E_1 disponemos de cierta información adicional: la nueva población correspondiente a la segunda extracción no coincide con la original, ya que sólo hay 6 artículos en buen estado en un total de 9.

Notemos además, que si las extracciones se realizan con reemplazo, esto no ocurre, ya que el resultado de la primera extracción no nos da ninguna información sobre la segunda. En este último caso se tiene

$$\begin{aligned}
&\mathbb{P}(E_1) = \frac{7 \times 10}{10^2}, \quad \mathbb{P}(E_1 \cap E_2) = \frac{7 \times 7}{10^2}, \ \mathbb{P}(E_2) = \frac{7 \times 10}{10^2} \\
&\mathbb{P}(E_2/E_1) = \frac{\mathbb{P}(E_1 \cap E_2)}{\mathbb{P}(E_1)} = \frac{7}{10} = \mathbb{P}(E_2) \\
&\#\Omega = 10 \times 10 = 100
\end{aligned}$$

PROBABILIDAD CONDICIONAL.

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $B \in \mathcal{F}$ un evento tal que

$$\mathbb{P}(B) > 0.$$

Definimos una función

$$\mathbb{P}(\cdot/B) : \mathcal{F} \to \mathbb{R}$$

mediante la fórmula

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

para todo $A \in \mathcal{F}$.

Esta función así definida, es una función de probabilidad, en efecto:

1. $\mathbb{P}(A/B) \geq 0$ para todo $A \in \mathcal{F}$.
2. $\mathbb{P}(\Omega/B) = 1$
3. $\mathbb{P}(\bigcup_{k=1}^{\infty} A_k /B) = \dfrac{\mathbb{P}\left(\bigcup_{k=1}^{\infty}(A_k \cap B)\right)}{\mathbb{P}(B)} = \sum_{k=1}^{\infty} \mathbb{P}(A_k/B)$ para toda sucesión de eventos $A_k \in \mathcal{F}$, $k \geq 1$, disjuntos dos a dos.

La función $\mathbb{P}(\cdot/B)$ se llama probabilidad condicional dado el evento B.

PROPIEDADES:

1. Si $A \cap B = \emptyset$ entonces $\mathbb{P}(A/B) = 0$
2. Si $B \subset A$ entonces $\mathbb{P}(A/B) = 1$.
3. Si todos los puntos de un espacio muestral finito son equiprobables, entonces

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

   si B es un evento tal que $\mathbb{P}(B) > 0$, (en este caso $\#B > 0$).

Observe que la probabilidad condicional respecto a un suceso particular B , equivale a considerar B como un nuevo espacio muestral, cuyos eventos serán los eventos del espacio original intersectados con B, es decir, si $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidades,

$$\mathcal{F}_B = \{A \cap B : \ A \in \mathcal{F}\}$$

es la familia de eventos correspondiente al espacio muestral B, y definimos sobre ellos la misma función de probabilidad del espacio original, multiplicada por un factor de normalización $\dfrac{1}{\mathbb{P}(B)}$. Denotamos esta función de probabilidad como

$$\mathbb{P}(\cdot/B)$$

ó equivalentemente

$$\mathbb{P}_B(\cdot)$$

Cuando empleamos el símbolo $\mathbb{P}(A)$ para la probabilidad de A, aludimos en realidad a la probabilidad de A dado algún espacio muestral Ω, así que $\mathbb{P}(A)$ es en realidad $\mathbb{P}(A/\Omega)$, pero para simplificar utlizamos simplemente $\mathbb{P}(A)$.

EJERCICIO: Demuestre que si $\mathbb{P}(\cdot/B)$ es una probabilidad condicional correspondiente al espacio de probabilidades $(\Omega, \mathcal{F}, \mathbb{P})$ y $B \in \mathcal{F}$ con $\mathbb{P}(B) > 0$ entonces:

1. $\mathbb{P}(A^c/B) = 1 - \mathbb{P}(A/B)$ para todo $A \in \mathcal{F}$.
2. Si $A, C \in \mathcal{F}$ entonces

$$\mathbb{P}\left((A - C)/B\right) = \mathbb{P}(A/B) - \mathbb{P}\left((A \cap C)/B\right)$$

   en particular si $C \subset A$ entonces

$$\mathbb{P}\left((A - C)/B\right) = \mathbb{P}(A/B) - \mathbb{P}(C/B).$$

INTERSECCIÓN DE DOS Ó MAS EVENTOS:

De la definición de probabilidad condicional se deducen las siguientes expresiones para la probabilidad de la intersección de dos ó mas eventos:

1. Si A y B son dos eventos tales que $\mathbb{P}(B) > 0$, entonces

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

2. Si $A_1, A_2, \cdots, A_n$ son eventos tales que $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_k) > 0$ para $1 \leq k \leq n$, entonces

$$\begin{aligned}&\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \\ &\mathbb{P}(A_n/A_1 \cap A_2 \cap \cdots \cap A_{n-1})\mathbb{P}(A_{n-1}/A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \cdots \\ &\cdots \mathbb{P}(A_3/A_1 \cap A_2)\mathbb{P}(A_2/A_1)\mathbb{P}(A_1).\end{aligned}$$

   esta última fórmula se puede demostrar por inducción.

EJEMPLOS:

1. Consideremos las combinaciones posibles de los hijos de una familia de dos niños siendo todas equiprobables. Definamos los siguientes eventos:

$$\begin{aligned} A &= \{\text{la familia tiene un varón}\} = \{(v,v),(v,h),(h,v)\} \\ B &= \{\text{el mayor es un varón}\} = \{(v,v),(h,v)\} \\ C &= \{\text{los dos niños son varones}\} = \{(v,v)\} \end{aligned}$$

   Queremos calcular

   (a) la probabilidad que los dos niños sean varones sabiendo que la familia tiene un hijo varón

   (b) la probabilidad que los dos niños sean varones sabiendo que el hijo mayor es varón.
   A primera vista, pareciera que estas dos probabilidades condicionales son iguales, hagamos los cálculos para aclarar la situación:
   El espacio muestral es

$$\Omega = \{(v,v),(v,h),(h,v),(h,h)\}$$

   queremos calcular $\mathbb{P}(C/A)$ y $\mathbb{P}(C/B)$:

$$\begin{aligned} \mathbb{P}(C/A) &= \frac{\mathbb{P}(A\cap C)}{\mathbb{P}(A)} = \frac{1/4}{3/4} = \frac{1}{3} \\ \mathbb{P}(C/B) &= \frac{\mathbb{P}(B\cap C)}{\mathbb{P}(B)} = \frac{1/4}{1/2} = \frac{1}{2} \end{aligned}$$

2. Se lanzan dos dados distinguibles, se quiere hallar la probabilidad que la suma de los números sea 7 y al mismo tiempo la diferencia entre el mayor y el menor sea 1.

$$\begin{aligned} A &= \{\text{la suma es 7}\} \\ B &= \{\text{la diferencia entre el mayor y el menor es 1}\} \end{aligned}$$

   El cardinal del espacio muestral es $6^2 = 36$, $Card(A) = 6$, $Card(B) = 10$, por lo tanto:

$$\begin{aligned} \mathbb{P}(A) &= \frac{6}{36} = \frac{1}{6} \\ \mathbb{P}(B/A) &= \frac{2}{6} = \frac{1}{3} \end{aligned}$$

   de aquí podemos obtener

$$\mathbb{P}(A\cap B) = \mathbb{P}(B/A)\mathbb{P}(A) = \frac{1}{6}\cdot\frac{1}{3} = \frac{1}{18}$$

   Observe que $\mathbb{P}(B) = \dfrac{10}{36} = \dfrac{5}{18}$, por lo tanto, en esta caso

$$\mathbb{P}(A\cap B) = \frac{1}{18},\ \mathbb{P}(A)\mathbb{P}(B) = \frac{5}{108}$$

   Cuando $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$, diremos que los dos eventos A y B son independientes.

PRINCIPIO DE EXPANSIÓN. TEOREMA DE BAYES .

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidades, A un evento del espacio muestral y B_i con $1 \leq i \leq n$, una sucesión finita de eventos disjuntos dos a dos. Si la realización del suceso A va acompañada de la realización de alguno de los sucesos B_i, es decir;

$$A = \bigcup_{i=1}^{n}(A\cap B_i)$$

entonces, por definición de probabilidad:

$$\mathbb{P}(A) = \sum_{i=1}^{n}\mathbb{P}(A\cap B_i)$$

PRINCIPIO DE EXPANSIÓN: Si $A = \bigcup_{i=1}^{n}(A \cap B_i)$ y $\mathbb{P}(B_i) > 0$ para $1 \leq i \leq n$, entonces, la probabilidad del evento A puede expresarse mediante la fórmula

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A/B_i)\mathbb{P}(B_i)$$

con B_i , $1 \leq i \leq n$, eventos disjuntos dos a dos. Esta expresión es consecuencia de la definición de probabilidad condicional y se conoce con el nombre de Principio de Expansión.

TEOREMA DE BAYES: Sea $A = \bigcup_{i=1}^{n}(A \cap B_i)$ y $\mathbb{P}(B_i) > 0$,con B_i, $1 \leq i \leq n$, eventos disjuntos dos a dos. Supongamos que conocemos $\mathbb{P}(A/B_i)$, para $1 \leq i \leq n$; entonces podemos calcular $\mathbb{P}(B_i/A)$ mediante la siguiente fórmula:

$$\mathbb{P}(B_i/A) = \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A/B_i)\mathbb{P}(B_i)}{\sum_{k=1}^{n} \mathbb{P}(A/B_k)\mathbb{P}(B_k)}$$

Este resultado recibe el nombre de Teorema de Bayes.

EJEMPLOS:

1. En un día de lluvia, la probabilidad que Bertha llegue tarde a clase es de $0,8$; mientras que en un día de sol es solo de $0,1$.

   Un 70% de los días son lluviosos y el resto son de sol.

   (a) Cuál es la probabilidad que Bertha llegue tarde a clase?

   (b) Un día miércoles de este año, Bertha llegó tarde a clase, cuál es la probabilidad que ese día haya sido lluvioso?

   Consideremos los siguientes sucesos:

$$\begin{aligned} T &= \{\text{Bertha llega tarde}\} \\ L &= \{\text{el día es lluvioso}\} \\ S &= \{\text{el día es de sol}\} \end{aligned}$$

   Sabemos que

$$\begin{aligned} &\mathbb{P}(T/L) = 0,8\,;\, \mathbb{P}(T/S) = 0,1 \\ &\mathbb{P}(L) = 0,7;\, \mathbb{P}(S) = 0,3. \end{aligned}$$

   Puesto que

$$T = (T \cap L) \cup (T \cap S)$$

   aplicando el principio de expansión, obtenemos:

$$\mathbb{P}(T) = \mathbb{P}(T/L)\mathbb{P}(L) + \mathbb{P}(T/S)\mathbb{P}(S) = 0,59$$

   que representa la probabilidad que Bertha llegue tarde.

   Para calcular la probabilidad, que ese miércoles, halla sido un día lluvioso, puesto que Bertha llegó tarde ese día, nos valemos del Teorema de Bayes

$$\mathbb{P}(L/T) = \frac{\mathbb{P}(T/L)\mathbb{P}(L)}{\mathbb{P}(T)} = \frac{0,56}{0,59} = 0,95$$

2. Se tienen tres cajas con el siguiente contenido: la primera contiene tres bolas blancas y una bola negra, la segunda dos bolas blancas y dos bolas negras, la tercera contiene 3 bolas blancas.

   Se elige una caja al azar, si resulta seleccionada la primera ó la segunda caja, se extraen dos bolas con reemplazo y si resulta elegida la tercera, se extraen dos bolas sin reemplazo.

   Al realizar la experiencia, se obtienen dos bolas blancas, cuál es la probabilidad que provenga de la caja $\#i, i = 1, 2, 3$.

Consideremos los siguientes eventos:

$$A_i = \{\text{elegimos la caja número } i\}$$
$$B = \{\text{las dos bolas seleccionadas son blancas}\}$$

$$\mathbb{P}(B/A_1) = \frac{\binom{4}{2}}{\binom{5}{2}} = \frac{3}{5}$$
$$\mathbb{P}(B/A_2) = \frac{\binom{3}{2}}{\binom{5}{2}} = \frac{3}{10}$$
$$\mathbb{P}(B/A_3) = \frac{\binom{3}{2}}{\binom{3}{2}} = 1$$

Por el Teorema de Bayes

$$\mathbb{P}(A_i/B) = \frac{\mathbb{P}(B/A_i)\mathbb{P}(A_i)}{\mathbb{P}(B)},\ i = 1, 2, 3.$$

Por el principio de Expansión

$$\mathbb{P}(B) = \sum_{i=1}^{3} \mathbb{P}(B/A_i)\mathbb{P}(A_i) = \frac{3}{5} \cdot \frac{1}{3} + \frac{3}{10} \cdot \frac{1}{3} + \frac{1}{3} = \frac{19}{30} \simeq 0,63$$

por lo tanto se tienen los siguientes resultados:

$$\mathbb{P}(A_1/B) = \frac{6}{19},\ \mathbb{P}(A_2/B) = \frac{3}{19},\ \mathbb{P}(A_3/B) = \frac{10}{19}$$

Observe que $\sum_{i=1}^{3} \mathbb{P}(A_i/B) = 1$, podría justificar este hecho?

PROBABILIDAD DE LA UNIÓN FINITA DE EVENTOS.

Sea $\{A_i\}, 1 \leq i \leq n$, una sucesión finita de eventos, no necesariamente disjuntos, entonces

$$\mathbb{P}(\bigcup_{i=1}^{n} A_i) = S_1 - S_2 + S_3 - S_4 + \cdots + (-1)^{k-1} S_k + \cdots + (-1)^{n-1} S_n$$

donde

$$S_1 = \textstyle\sum_{i=1}^{n} \mathbb{P}(A_i)$$
$$S_k = \sum_{1 \leq j_1 < j_2 < \cdots j_k \leq n} \mathbb{P}(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}),\ 2 \leq k \leq n.$$

DEMOSTRACIÓN:

Este resultado se demuestra por inducción sobre n. La fórmula ya ha sido demostrada para $n = 2$ ya que

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

Supongamos que la fórmula es cierta para todo entero $1 \leq k \leq n-1$. Aplicando el resultado conocido para $k = 2$ y la hipótesis de inducción, se tiene que:

$$\begin{aligned}
&\mathbb{P}\left(\textstyle\bigcup_{k=1}^{n} A_k\right) = \mathbb{P}\left((\textstyle\bigcup_{k=1}^{n-1} A_k) \cup A_n\right) = \\
&\mathbb{P}\left(\textstyle\bigcup_{k=1}^{n-1} A_k\right) + \mathbb{P}(A_n) - \mathbb{P}\left((\textstyle\bigcup_{k=1}^{n-1} A_k) \cap A_n\right) = \\
&\textstyle\sum_{i=1}^{n} \mathbb{P}(A_i) - \displaystyle\sum_{1 \leq j_1 < j_2 \leq n-1} \mathbb{P}(A_{j_1} \cap A_{j_2}) + \\
&\sum_{1 \leq j_1 < j_2 < j_3 \leq n-1} \mathbb{P}(A_{j_1} \cap A_{j_2} \cap A_{j_3}) - \cdots - \\
&+(-1)^{n-2}\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) - \\
&[\textstyle\sum_{i=1}^{n-1} \mathbb{P}(A_i \cap A_n) - \displaystyle\sum_{1 \leq j_1 < j_2 \leq n-1} \mathbb{P}(A_{j_1} \cap A_{j_2} \cap A_n) + \cdots \\
&+(-1)^{n-2}\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n)] = \\
&S_1 - S_2 + \cdots + (-1)^{n-1} S_n.
\end{aligned}$$

PROBLEMA DE COINCIDENCIAS. Se tienen n objetos y se permutan aleatoriamente. Diremos que hay una coincidencia en el lugar i, si el objeto número i queda en la $i-$ésima posición al realizar la permutación. Queremos calcular la probabilidad que no haya coincidencia.

Denotemos por

$$A_i = \{\text{hay coincidencia en la } i \text{ -ésima posición}\}$$

El conjunto $\cup_{i=1}^n A_i$ representa el evento "hay al menos una coincidencia". Por lo tanto el complementario de este conjunto representa el evento "no hay coincidencias" cuya probabilidad queremos calcular.

Hallaremos la probabilidad de la unión usando la fórmula anteriormente demostrada:

$$\mathbb{P}(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

ya que el objeto i estará fijo en su posición y el resto permutará en las otras $(n-1)$ posiciones. De aquí obtenemos:

$$S_1 = \sum_{i=1}^{n} \mathbb{P}(A_i) = 1$$

La intersección de cualquier par de estos eventos, es:

$$\mathbb{P}(A_i \cap A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

El cardinal del conjunto

$$\{(i,j) : 1 \leq i < j \leq n\}$$

es idéntico al número de muestras no ordenadas, de tamaño $r = 2$, sin reemplazo, que se pueden formar a partir de la población

$$S = \{1, 2, 3, \cdots, n\}$$

obteniéndose así

$$S_2 = \binom{n}{2} \cdot \frac{1}{n(n-1)} = \frac{1}{2!}$$

En general,

$$\mathbb{P}(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}) = \frac{(n-k)!}{n!}$$

$$S_k = \tbinom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}$$

De donde, la probabilidad que hay al menos una coincidencia, es:

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1} \frac{1}{n!}$$

así, la probabilidad que no haya coincidencias, es:

$$\mathbb{P}\left((\bigcup_{k=1}^{n} A_k)^c\right) = 1 - \mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \frac{1}{2!} - \frac{1}{3!} - \cdots + (-1)^n \frac{1}{n!}.$$

Observe que el n-ésimo polinomio de Taylor alrededor de $x = 0$ de $f(x) = e^{-x}$ es

$$P_n(x) = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots + (-1)^n \frac{x^n}{n!},$$

haciendo $x = 1$ obtenemos que

$$\mathbb{P}\left((\bigcup_{k=1}^{n} A_k)^c\right) \simeq e^{-1}$$

EJEMPLO: Se escriben n cartas a n personas de esta ciudad. Se escribe cada dirección en un sobre, se colocan aleatoriamente las n cartas en los n sobres. Cuál es la probabilidad que ninguna carta llegue a su destino?

Este problema es una aplicación directa del problema de las coincidencias. Denotamos por

$$A_i = \{\text{la } i\text{-ésima carta está colocada en el sobre correcto}\}$$

Lo que se quiere calcular es

$$\mathbb{P}\left((\bigcup_{k=1}^{n} A_k)^c\right) = \frac{1}{2!} - \frac{1}{3!} - \cdots + (-1)^n \frac{1}{n!}.$$

EJERCICIOS RESUELTOS:

1. **PARADOJA DE BERTRAND.**

   Tres cajas $C_i, 1 \leq i \leq 3$, contienen dos monedas cada una. En la primera, ambas son de oro, en la segunda, ambas son de plata y en la tercera una es de oro y la otra de plata. Se escoje una caja al azar y luego una moneda al azar. Si la moneda elegida es de oro, cuál es la probabilidad que provenga de la caja que contiene dos monedas de oro?

   Si consideramos el problema sin mucho cuidado, podríamos razonar de la siguiente manera: al sacar una moneda de oro, sólo podría provenir de la primera o de la tercera caja y como las cajas se eligen al azar, la probabilidad que provenga de C_1 es $1/2$. En este razonamiento irresponsable no estamos tomando en cuenta que es más probable sacar una moneda de oro de la primera caja que de la segunda. Resolveremos el problema con mas cuidado:

   $$G = \{\text{la moneda elegida es de oro}\}$$

   queremos calcular

   $$\mathbb{P}(C_1/G) = \frac{\mathbb{P}(G/C_1)\mathbb{P}(C_1)}{\mathbb{P}(G)}$$

   por el principio de expansión

   $$\mathbb{P}(G) = \sum_{i=1}^{3} \mathbb{P}(G/C_i)\mathbb{P}(C_i) = \frac{3}{6} = \frac{1}{2}$$

   por lo tanto

   $$\mathbb{P}(C_1/G) = \frac{2}{3}.$$

2. Se lanzan 6 dados distinguibles, balanceados. Cuál es la probabilidad de obtener seis números distintos?

   Este ejercicio es fácil de resolver directamente ya que $\#\Omega = 6^6$ y el cardinal del evento

   $$A = \{\text{se obtienen seis números distintos}\}$$

   es $6.5.4.3.2.1 = 6!$, pero queremos ilustrar la expresión de $\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6)$, en función de las probabilidades condicionales.

   Consideremos los siguientes eventos:

   $$\begin{aligned}
   A_1 &= \{\text{se obtiene cualquier número en el dado \#1}\} \\
   A_2 &= \{\text{el resultado del dado\#2 es distinto al del \#1}\} \\
   A_3 &= \{\text{el tercer resultado es distinto a los dos primeros}\} \\
   A_4 &= \{\text{el resultado \#4 es distinto a los tres primeros}\} \\
   A_5 &= \{\text{el resultado \#5 es distinto a los cuatro primeros}\} \\
   A_6 &= \{\text{el resultado \#6 es distinto a los cinco primeros}\}
   \end{aligned}$$

Queremos calcular la probabilidad de la intersección de estos seis eventos, para ello nos valemos de

$$
\begin{aligned}
&\mathbb{P}(A_1) = 1\\
&\mathbb{P}(A_2/A_1) = \frac{5}{6}\\
&\mathbb{P}(A_3/A_1 \cap A_2) = \frac{4}{6}\\
&\mathbb{P}(A_4/A_1 \cap A_2 \cap A_3) = \frac{3}{6}\\
&\mathbb{P}(A_5/A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{2}{6}\\
&\mathbb{P}(A_6/A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = \frac{1}{6}
\end{aligned}
$$

de aquí obteneos el resultado

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6) = \frac{5!}{6^5} = \frac{6!}{6^6}$$

3. Un estudiante responde una pregunta de un examen de múltiple escogencia, que tiene cuatro respuestas posibles. Supongamos que la probabilidad que el estudiante conozca la respuesta a la pregunta es $0,8$ y la probabilidad que adivine es $0,2$.

   Si el estudiante adivina, la probabilidad que acierte es $0,25$. Si el estudiante responde acertadamente la pregunta, cuál es la probabilidad que el estudiante realmente sepa la respuesta?

   Definamos los siguientes eventos:

$$
\begin{aligned}
&E = \{\text{el estudiante conoce la respuesta}\}\\
&D = \{\text{el estudiante adivina}\}\\
&A = \{\text{el estudiante responde acertadamente}\}
\end{aligned}
$$

   Queremos calcular

$$\mathbb{P}(E/A) = \frac{\mathbb{P}(A/E)\mathbb{P}(E)}{\mathbb{P}(A)}$$

   por el principio de expansión

$$\mathbb{P}(A) = \mathbb{P}(A/E)\mathbb{P}(E) + \mathbb{P}(A/D)\mathbb{P}(D) = 1 \cdot 0,8 + 0,25 \cdot 0,2 = 0,85$$

   por lo tanto

$$\mathbb{P}(E/A) = \frac{\mathbb{P}(A/E)\mathbb{P}(E)}{\mathbb{P}(A)} = \frac{0,8}{0,85} = 0,941$$

EVENTOS INDEPENDIENTES

Al introducir el concepto de probabilidad condicional, hemos visto que en ciertos casos, cuando la ocurrencia del evento B no nos aporta información respecto a la ocurrencia o no del suceso A, se tiene que

$$\mathbb{P}(A/B) = \mathbb{P}(A)$$

En este caso decimos que A es independiente de B. Ahora bien, como

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

si $\mathbb{P}(A/B) = \mathbb{P}(A)$ podemos expresar esta igualdad como:

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

Esta expresión es válida, aun cuando $\mathbb{P}(A) = 0$ ó $\mathbb{P}(B) = 0$. Usamos pues, esta igualdad como definición.

DEFINICIÓN: Dos eventos A y B son independientes, si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

PROPOSICIÓN:

1. Ω y $\emptyset$ son independientes de cualquier suceso.
2. Si A y B son independientes, también lo son:
   (a) A y B^c
   (b) A^c y B^c

Demostración:

1. $\mathbb{P}(\Omega \cap A) = \mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(\Omega)$
   $\mathbb{P}(\emptyset \cap A) = \mathbb{P}(\emptyset) = 0 = \mathbb{P}(\emptyset)\mathbb{P}(A)$.
   (a) $\mathbb{P}(A \cap B^c) = \mathbb{P}(A - B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B^c)$
   (b) $\mathbb{P}(A^c \cap B^c) = 1 - \mathbb{P}(A \cup B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B) = \mathbb{P}(A^c)\mathbb{P}(B^c)$.

GENERALIZACIÓN A UNA FAMILIA CUALQUIERA. Sea $\mathcal{C} = \{A_i, i \in I\}$ una familia cualquiera de eventos. Diremos que los eventos $A_i, i \in I$,

son independientes, si para cualquier subcolección finita de eventos

$$A_{i_1}, A_{i_2}, A_{i_3}, \ldots A_{i_n} \in \mathcal{C}$$

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} \mathbb{P}\left(A_{i_k}\right)$$

en este caso diremos que la familia $\mathcal{C}$ es una familia de eventos independientes.

OBSERVACIÓN. En esta definición sólo intervienen colecciones finitas de eventos de la familia $\mathcal{C}$, pero intervienen todas las colecciones finitas de dicha familia.

Por ejemplo: se lanzan dos dados distinguibles, consideremos la siguiente familia de eventos

$$\mathcal{C} = \{A_1, A_2, A_3\}$$

donde

$$\begin{aligned} A_1 &= \{\text{se obtiene seis en el primer dado}\} \\ A_2 &= \{\text{se obtiene uno en el segundo dado}\} \\ A_3 &= \{\text{la suma de las caras superiores es siete}\} \end{aligned}$$

claramente

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \frac{1}{6}$$

y la probabilidad de las intersecciones

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = \frac{1}{6^2}$$

por lo tanto

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

si i es diferente de j. Pero, la familia no es de eventos independientes, pues

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \frac{1}{6^2}$$

y este valor es diferente al producto de las probabilidades de los tres eventos.

Así, no es suficiente verificar la independencia solamente, para las parejas de eventos si queremos verificar la independencia de toda la familia.

EJEMPLO Un lote de 10 objetos, contiene 4 defectuosos y 6 en buen estado. Se extraen dos objetos sucesivamente y sin reemplazo. Definimos los siguientes eventos

$$\begin{aligned} D_1 &= \{\text{el primer objeto es defectuoso}\} \\ D_2 &= \{\text{el segundo objeto es defectuoso}\} \end{aligned}$$

Usando el principio de expansión, obtenemos que

$$\mathbb{P}(D_2) = \mathbb{P}(D_2/D_1)\mathbb{P}(D_1) + \mathbb{P}(D_2/D_1^c)\mathbb{P}(D_1^c) = \frac{3}{9} \cdot \frac{4}{10} + \frac{4}{9} \cdot \frac{6}{10} = \frac{2}{5}.$$

Puesto que

$$\mathbb{P}(D_2/D_1) = \frac{3}{9} = \frac{1}{3}$$

concluímos que D_1 y D_2 no son independientes.

Si extraemos los objetos con reemplazo, resultarán independientes, pues

$$\begin{aligned} \mathbb{P}(D_1) = \mathbb{P}(D_2) &= \frac{4}{10} = \frac{2}{5} \\ \mathbb{P}(D_1 \cap D_2) &= \frac{4^2}{10^2} = \mathbb{P}(D_1) \cdot \mathbb{P}(D_2) \end{aligned}$$

EJERCICIOS PARA LA PRÁCTICA

1. Se quiere construir una autopista, pero dicha obra puede demorarse en virtud de una huelga. Si la probabilidad que ocurra una huelga es 0.6, la pobabilidad que el trabajo se termine a tiempo si no hay huelga es 0.85 y la probabilidad que el trabajo se termine a tiempo si hay huelga es 0.35. Cuál es la probabilidad que la obra se concluya a tiempo? Si el trabajo se terminó a tiempo, cuál es la probabilidad que, pese a ello, hubiese estallado una huelga?

   RESPUESTAS: $0,55; 0.3818$.

2. Los miembros de una firma de consultoría renta automóviles en tres agencias: el 60% de la agencia A, el 30% de la agencia B y el 10% de la agencia C. Si el 9% de los vehículos que provienen de la agencia A necesitan entonación, el 20% de las unidades de la agencia B necesitan entonación y el 6% de los que provienen de la agencia C necesitan entonación, cuál es la probabilidad que un auto rentado a la firma necesite entonación? Si un automóvil rentado a la firma necesitó entonación, cuál es la probabilidad que provenga de la agencia B ?

   RESPUESTAS: $0.12, 0.5$.

3. En la ciudad de Caracas, el 25% de los automóviles emiten excesiva cantidad de contaminantes. Si la probabilidad que un automóvil que emite excesiva cantidad de contaminantes no apruebe la prueba de emisión vehicular es de 0.99 y la probabilidad que un automóvil que no emite excesiva cantidad de contaminantes no apruebe es 0.17, cuál es la probabilidad que un automóvil que no apruebe la prueba emita excesiva cantidad de contaminantes? RESPUESTA: 0.66

4. En una planta electrónica, se sabe que la probabilidad que un nuevo trabajador que ha asistido al programa de capacitación de la compañía cumpla con la cuota de producción es de 0.84 y que la probabilidad correspondiente es 0.49 de un nuevo trabajador que no ha asistido al programa de capacitación. Si el 70% de todos los trabajadores de nuevo ingreso asisten al programa de capacitación, cuál es la probabilidad que un nuevo trabajador cumplirá con su cuota de producción? , cuál es la probabilidad que un trabajador que cumple su cuota de producción asistió al programa de capacitación de la compañía? RESPUESTAS: 0.735; 0.8.

5. Una compañía de ventas por correo tiene tres empleados de almacén: U, V y W, quienes toman productos de la bodega y los ensamblan para la subsiguiente verificación y empaquetado. Se cometen errores en un pedido ya sea tomando un producto equivocado o seleccionando la cantidad equivocada. La probabilidad de equivocarse de cada uno de los empleados es una vez en cien, cinco veces en cien y tres veces en cien,

respectivamente. Si U, V y W cubren , el 30, , 40 y 30% de todos los pedidos, cuál es la probabilidad que se cometa un error? Se encuentra un error en un pedido, cuál es la probabilidad que lo haya cometido el empleado V? RESPUESTAS: 0.032; 0.625.

6. El 5% de las unidades producidas en una fábrica se encuentran defectuosas cuando el proceso de fabricación se encuentra bajo control. Si el proceso se encuentra fuera de control, se produce un 30% de unidades defectuosas. La probabilidad marginal que el proceso se encuentre bajo control es de 0.92. Si se escoje aleatoriamente una unidad y se encuentra que es defectuosa, cuál es la probabilidad que el proceso se encuentre bajo control? RESPUESTA: $0,6571$.

7. Un inversionista está pensando en comprar un número muy grande en acciones de una compañía. La cotización de las acciones en la bolsa, durante los seis meses anteriores, es de interés para el inversionista. Con base en esta información, se observa que la cotización se relaciona con el producto nacional bruto. Si el PNB aumenta, la probabilidad que las acciones aumenten es de 0.8. Si el PNB es el mismo, la probabilidad que las acciones aumenten su valor es de 0.2. Si el PNB disminuye, la probabilidad es de sólo 0.1. Si para los siguientes seis meses se asignan las probabilidades $0.4, 0.3$ y 0.3 a los eventos, el PNB aumenta, es el mismo y disminuye, respectivamente, determine la probabilidad que las acciones aumenten su valor en los próximos seis meses. RESPUESTA: 0.41.

8. Mañana tendrá lugar el encuentro de fútbol entre los equipos de Alemania y Brasil. Nuestro comentarista deportista ve la situación así:: Estoy seguro que se abrirá el marcador y cualquiera de los dos equipos tiene igual probabilidad de hacerlo; si Alemania anota el primer gol su moral se levantará y la probabilidad que el próximo gol también sea suyo es de 2/3 contra 1/3 que sea de Brasil; en cambio si es Brasil quien anota primero la reacción no se hará esperar y habrá un segundo gol que puede ser con igual probabilidad para cualquier bando. Si el marcador llega a ponerse $2-0$ a favor de cualquiera de los equipos, por desmoralización de uno y apatía del otro no habrá más goles; en cambio, si llega a ponerse $1-1$ pueden ocurrir tres cosas con iguales probabilidades: que Alemania anote y gane $2-1$, que Brasil anote y gane $1-2$ ó que no haya más goles". Cuál es, de acuerdo a nuestro comentarista, la probabilidad que:

   (a) Brasil gane? RESPUESTA: 7/18.
   (b) Brasil gane, dado que abrió el marcador? RESPUESTA: 2/3.
   (c) Alemania haya abierto el marcador, dado que Brasil ganó?
   RESPUESTA: 1/7.
   (Usar diagrama de árbol).

9. A y B son sucesos disjuntos. Dar una condición necesaria y suficiente para que A, B sean independientes.

10. Sean A, B dos eventos disjuntos y de probabilidad positiva. Pueden ser independientes estos dos eventos?

11. Si $A_1, A_2, \cdots, A_n$ son eventos independientes, muestre que:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1-(1-\mathbb{P}(A_1))(1-\mathbb{P}(A_2))\cdots(1-\mathbb{P}(A_n)).$$

12. Una muestra de tamaño 4 se extrae con reposición de una bolsa que contiene 6 bolas, de las cuales 4 son blancas. Sea A el evento, "exactamente una de las cuatro bolas extraídas es blanca" y sea B el evento "la cuarta bola es blanca". Son A y B independientes?. Qué sucede si el muestreo se realiza sin reposición? Si definimos un tercer evento C como el evento "exactamente las dos primeras bolas extraídas son blancas", son A, B, C independientes?

13. Se extrae una bola de una caja que contiene cuatro blancas y dos negras. Si la bola es blanca, se le deja fuera de la bolsa, mientras que si es negra, se la vuelve a colocar dentro. Extraemos luego otra bola. Sea A el evento "la primera bola es blanca"y B = "la segunda bola es blanca". Diga si las siguientes proposiciones son ciertas o falsas, en caso de ser falsas dé la respuesta correcta:

   (a) $\mathbb{P}(A) = 2/3$

(b) $\mathbb{P}(B) = 3/5$

(c) $\mathbb{P}(B/A) = 3/5$

(d) $\mathbb{P}(A/B) = 9/14$

(e) Los eventos A y B son disjuntos.

14. Cuál es el menor valor de n para el cual la probabilidad de obtener al menos un 6 en una serie de n lanzamientos de un dado es mayor que 3/4?

    RESPUESTA: $n = 8$.

15. Si $A_1, A_2, \cdots, A_n$ son eventos independientes y $\mathbb{P}(A_i) = p$, $i = 1, 2, \cdots, n$. Halle el menor valor de n para el cual

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \geq p_0$$

    donde p_0 es un número fijo. $(0 < p_0 < 1)$.

Capítulo 4

VARIABLES ALEATORIAS

INTRODUCCIÓN:

En los capítulos anteriores, estudiamos los conceptos básicos de probabilidad con respecto a eventos de un espacio muestral. Los experimentos se conciben de manera que los elementos del espacio muestral o eventos elementales son cualitativos o cuantitativos.

Como ejemplos de resultados cualitativos se tiene:

a) los resultados posibles al lanzar una moneda son "cara" o "sello".

b) un producto manufacturado en una fábrica puede ser "defectuoso" o "no defectuoso".

c) una persona en particular puede preferir el perfume A o el perfume B.

La cuantificación de los resultados cualitativos de un experimento, puede ser útil y, mediante el empleo de medidas numéricas, se estudia su comportamiento aleatorio.

Es decir, con frecuencia nos interesa asociar a cada elemento del espacio muestral (ya sea numérico ó no) valores numéricos, según sea sugerido por el problema que estemos investigando. Por ejemplo, si lanzamos un dado dos veces, el espacio muestral es:

$$\Omega = \{(m, n) : 1 \leq m, n \leq 6; m, n \in \mathbb{N}\}$$

si nos interesamos en el puntaje total, podemos asociar a cada evento elemental (m, n) el número $m + n$, luego, lo que hacemos es definir una función numérica sobre el espacio muestral asociado a esta experiencia aleatoria, esta función que debe cumplir ciertas propiedades, las cuales precisaremos cuando demos la definición rigurosa, se llamará variable aleatoria. Este término resulta un poco confuso, sería más apropiado llamarla función aleatoria, pues la variable independiente es un punto del espacio muestral, es decir, constituye el resultado de un experimento.

Antes de dar la definición formal de Variable Aleatoria, vamos a introducir algunos detalles preliminares:

Sea A un subconjunto de $\mathbb{R}$, X una función definida sobre un espacio muestral Ω con valores reales. Denotaremos por

$$[X \in A] = \{\omega \in \Omega : X(\omega) \in A\}.$$

Observe que este subconjunto del espacio muestral Ω no es más que $X^{-1}(A)$. Si Ω, es un espacio muestral discreto, el conjunto $[X \in A]$ (que puede ser eventualmente vacío) es siempre un evento de Ω si la familia de eventos $\mathcal{F} = \mathcal{P}(\Omega)$. En caso contrario, $[X \in A]$ no tiene por qué pertenecer a la familia de eventos del espacio muestral. Como ejemplo consideremos:

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \mathcal{F} = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$$

la función $X(\omega) = \omega$, no cumple que $[X \in A] \in \mathcal{F}$ para todo $A \subset \mathbb{R}$, pues si $A = [0, 3/2]$

$$X^{-1}(A) = \{1\}$$

no pertenece a la familia $\mathcal{F}$ de eventos.

Estamos interesados en pedir que para ciertos subconjuntos A de números reales, se cumpla que $[X \in A] \in \mathcal{F}$, pues la probabilidad $\mathbb{P}$ asociada a $(\Omega, \mathcal{F})$ solamente está definida sobre los eventos del espacio muestral. Pediremos que esta condición se cumpla para los intervalos de la recta real. Si se cumple para los intervalos de la recta real, se cumplirá para una familia más extensa, de subconjuntos de números reales, llamados Borelianos de $\mathbb{R}$, de hecho, todo intervalo de números reales, es un conjunto de Borel o Boreliano. Esta familia, que denotaremos por $\mathcal{B}(\mathbb{R})$, es muy extensa y constituye una $\sigma-$álgebra de subconjuntos de $\mathbb{R}$, de hecho es la menor $\sigma-$álgebra que contiene los conjuntos abiertos de $\mathbb{R}$, (tan extensa es, que a simple vista parece incluir todos los subconjuntos de la recta real, la construcción de un conjunto que no sea Boreliano, es bastante complicada y está fuera de los objetivos de este curso).

DEFINICIÓN:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, una función

$$X : \Omega \to \mathbb{R}$$

es una variable aleatoria real ó simplemente una variable aleatoria si para cualquier intervalo $I \subseteq \mathbb{R}$, el conjunto

$$[X \in I] = X^{-1}(I) \in \mathcal{F}$$

EJEMPLOS:

1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $A \in \mathcal{F}$. Definamos X por

$$X(\omega) = \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \in A^c \end{cases}$$

Observemos que el evento A ocurre si y solo si $X(\omega) = 1$. Si I es un intervalo cualquiera

$$[X \in I] = \begin{cases} \emptyset \text{ si } 0 \in I^c, 1 \in I^c \\ A \text{ si } 0 \in I^c, 1 \in I \\ A^c \text{ si } 0 \in I, 1 \in I^c \\ \Omega \text{ si } 0 \in I, 1 \in I \end{cases}$$

por lo tanto $[X \in I]$ siempre pertenece a la familia de eventos $\mathcal{F}$ y concluímos que la función X es una variable aleatoria. Esta variable aleatoria se conoce con el nombre de función indicatriz de A porque X nos indica si A ocurrió ó no ocurrió. (En otros contextos se denomina, función característica del conjunto A, pero en esta rama de las matemáticas, función característica representa otro concepto que veremos más a delante). Las notaciones usuales son:

$$\mathcal{X}_A \text{ y } 1_A$$

Si X es una variable aleatoria sobre un espacio de probabilidades, que toma sólo los valores 1 y 0, entonces X es la función indicatriz del evento

$$A = \{\omega : X(\omega) = 1\}.$$

Este conjunto A siempre es un evento, si X es una variable aleatoria (por qué?)

2. Sea c un número real, la función X definida por $X(\omega) = c$ es un variable aleatoria, pues para cualquier intervalo I,

$$[X \in I] = \begin{cases} \Omega \text{ si } c \in I \\ \emptyset \text{ si } c \in I^c \end{cases}$$

y tanto Ω como $\emptyset$ son siempre eventos. Esta es una variable aleatoria constante.

3. Sea $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ un espacio de probabilidades, donde $\mathcal{B}(\mathbb{R})$ son los conjuntos de Borel y $\mathbb{P}$ es alguna probabilidad definida sobre $\mathcal{B}(\mathbb{R})$. Sea c un número real y definamos la función $X : \mathbb{R} \to \mathbb{R}$ por

$$X(\omega) = \omega + c.$$

Si I es algún intervalo en $\mathbb{R}$,

$$[X \in I] = \{\omega : X(\omega) \in I\} = \{\omega : \omega + c \in I\} = \{\omega : \omega \in I - c\}$$

donde, para un conjunto cualquiera A y c una constante, definimos:

$$A + c = \{x : x = a + c, a \in A\}\,.$$

Por lo tanto el conjunto $I - c = [X \in I]$ es un intervalo, luego un Boreliano. Así, X es una variable aleatoria.

DEFINICIÓN: Rango.

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $X : \Omega \to \mathbb{R}$ una variable aleatoria. Llamamos rango de X al conjunto de valores que ella toma, es decir,

$$R(X) = \{x \in \mathbb{R} : \text{ existe } \omega \in \Omega : X(\omega) = x\}\,.$$

Si el rango de X es finito ó numerable, diremos que la variable aleatoria X es discreta.

EJEMPLOS:

1. En la experiencia que consiste en lanzar una moneda, el espacio muestral es:

$$\Omega = \{Cara, Sello\}$$

La función X definida por:

$$X(Cara) = 1, X(Sello) = 0$$

es una variable aleatoria discreta de rango

$$R(X) = \{0, 1\}$$

2. Se lanza una moneda N veces, a cada uno de los 2^N eventos elementales del espacio muestral le asociamos el número de caras obtenidas en los N lanzamientos. El rango de esta variable aleatoria es el conjunto

$$\{0, 1, 2, 3, \cdots, N\}\,.$$

3. Se lanza una moneda hasta que salga cara por primera vez. El espacio muestral asociado a esta experiencia es numerable. Definimos la variable aleatoria

$$X(\omega) = \{\text{ n}^o \text{ de lanazamientos necesarios para obtener cara por } 1^a \text{ vez}\}$$

El rango de esta variable aleatoria es el conjunto

$$\{1, 2, 3, \cdots, n, n+1, \cdots\}\,.$$

4. Se efectúan disparos sobre un disco de 1 metro de radio, no se toman en cuenta los disparos fueras del disco. El espacio muestral Ω es el círculo unitario. Para cada ω definimos $X(\omega) =$ distancia desde ω hasta el origen. El rango de esta variable aleatoria es:

$$R(X) = [0, 1]\,.$$

DISTRIBUCIÓN O LEY DE UNA VARIABLE ALEATORIA.

Antes de definir lo que se entiende por distribución de una variable aleatoria, vamos a ilustrar el concepto con un ejemplo.

Un jugador lanza una moneda tres veces, gana Bs.3 por cada cara y pierde Bs.3 por cada sello.

El espacio muestral asociado tiene 2^3 elementos que pueden representarse por:

$$\Omega = \{(ccc), (ccs), (css), (scs), (ssc), (csc), (scc), (sss)\}$$

Más que el punto muestral en sí, al jugador le interesa la ganancia ó pérdida que ésto representa.

Definimos $X(\omega) =$ "ganancia del jugador", siendo el rango de X

$$R(X) = \{-9, -3, 3, 9\}$$

Si I es un intervalo del conjunto de números reales, $X^{-1}(I) = \emptyset$ si $I \cap R(X) = \emptyset$, luego, sólo nos interesan los conjuntos de la forma

$$[X = e_i]$$

con $e_i \in R(X)$. Así:

$$\begin{array}{l} [X = -9] = \{(sss)\} \\ [X = -3] = \{(ssc), (css), (scs)\} \\ [X = 3] = \{(ccs), (csc), (scc)\} \\ [X = 9] = \{(ccc)\} \end{array}$$

Podemos calcular la probabilidad de cada uno de los eventos

$$\begin{array}{l} \mathbb{P}([X = -9]) = \frac{1}{8} \\ \mathbb{P}([X = -3]) = \frac{3}{8} \\ \mathbb{P}([X = 3]) = \frac{3}{8} \\ \mathbb{P}([X = 9]) = \frac{1}{8} \end{array}$$

A cada punto $e_i \in R(X)$ le hemos asociado un valor

$$\mathbb{P}_X(\{e_i\}) = \mathbb{P}([X = e_i]).$$

Es fácil ver que $\mathbb{P}_X$ es una probabilidad sobre $R(X)$, considerando como familia de eventos todos los subconjuntos del rango de X.

DEFINICIÓN:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $X : \Omega \to \mathbb{R}$ una variable aleatoria. La distribución de la variable aleatoria X es una probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ definida por

$$\mathbb{P}_X(I) = \mathbb{P}([X \in I])$$

para todo intervalo I de $\mathbb{R}$. Note que si $I \cap R(X) = \emptyset$, $\mathbb{P}_X(I) = 0$.

OBSERVACIÓN:

Para conocer enteramente la distribución $\mathbb{P}_X$ de una variable aleatoria X, es suficiente poder calcular $\mathbb{P}_X(I)$ cualquiera sea el intervalo I de $\mathbb{R}$.

En el ejemplo anterior, $R(X) = \{-9, -3, 3, 9\}$, por lo tanto para conocer $\mathbb{P}_X$ fué suficiente evaluar $\mathbb{P}_X(\{e_i\})$ con $e_i \in R(X)$, ya que si I es un intervalo, se obtiene que

$$I \cap R(X) = \emptyset$$

ó

$$e_i \in I$$

para algún $i = 1, 2, 3, 4$. En este último caso, para calcular $\mathbb{P}_X(I)$, basta calcular la probabilidad de cada elemento del rango.

Cuando X no es discreta, resulta impráctico valerse de la probabilidad $\mathbb{P}_X$ para describir la distribución de X.

Vamos a introducir otras funciones que nos permitirán describir X de manera más sencilla.

FUNCIÓN DE DISTRIBUCIÓN.

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $X : \Omega \rightarrow \mathbb{R}$ una variable aleatoria. Si $x \in \mathbb{R}$ (no necesariamente en el rango de X), el conjunto

$$[X \leq x] = \{\omega \in \Omega : X(\omega) \leq x\} = X^{-1}(-\infty, x]$$

es siempre un evento del espacio muestral.

La probabilidad de este evento depende del punto $x \in \mathbb{R}$.

Para cada $x \in \mathbb{R}$, definimos la función

$$F(x) = \mathbb{P}([X \leq x])$$

la cual se llama "función de distribución" de la variable aleatoria X.

En algunas ocasiones, para resaltar que F es la función de distribución de X escribimos F_X en lugar de F.

EJEMPLO:

Sea X una variable aleatoria de rango el conjunto

$$R(X) = \{1, 2, 3\}$$

con $\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \mathbb{P}(X = 3) = \dfrac{1}{3}$. Para calcular su función de distribución debemos calcular $\mathbb{P}([X \leq x])$ para cada $x \in \mathbb{R}$.

Si $-\infty < x < 1$, $F(x) = \mathbb{P}([X \leq x]) = \mathbb{P}(\emptyset) = 0$
Si $1 \leq x < 2$, $\quad F(x) = \mathbb{P}([X \leq x]) = \mathbb{P}([X = 1]) = \frac{1}{3}$
Si $2 \leq x < 3$, $\quad F(x) = \mathbb{P}([X \leq x]) = \mathbb{P}([X = 1 \text{ ó } X = 2]) = \frac{2}{3}$
Si $3 \leq x < \infty$, $\quad F(x) = \mathbb{P}([X \leq x]) = \mathbb{P}([X = 1 \text{ ó } X = 2 \text{ ó } X = 3]) = 1$.

$F(x)$ es una función en escalera con los saltos en los puntos del rango de X, sus discontinuidades son de primer grado, es decir, en los puntos de discontinuidad, existen los límites laterales, pero son diferentes, es continua por la derecha, con salto igual a $\frac{1}{3}$ en cada discontinuidad.

Por ser F continua por la derecha

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

La diferencia

$$\lim_{x \to a^+} F(x) - \lim_{x \to a^-} F(x)$$

se denomina salto de F en el punto "a".

PROPOSICIÓN:

Sea X una variable aleatoria y F su función de distribución, entonces:

1. F es monótona creciente.
2. F es continua por la derecha.
3. $\lim\limits_{x \to -\infty} F(x) = 0, \ \lim\limits_{x \to \infty} F(x) = 1$

DEMOSTRACIÓN:

1. Si $x_1 < x_2$ entonces.

$$F(x_2) - F(x_1) = \mathbb{P}([X \leq x_2]) - \mathbb{P}([X \leq x_1]) = \mathbb{P}([x_1 < X \leq x_2]) \geq 0$$

2. Sea $a \in \mathbb{R}$, para demostrar que F es continua por la derecha en el punto a, es suficiente verificar que si $(x_n)_{n\geq 1}$ es una sucesión decreciente de números reales que tiende al punto "a", entonces

$$\lim_{n\to\infty} F(x_n) = F(a)$$

Observe que

$$[X \leq a] = \bigcap_{n=1}^{\infty} [X \leq x_n]$$

y por ser $A_n = [X \leq x_n]$ una sucesión decreciente de eventos, resulta que:

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mathbb{P}([X \leq x_n]) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} [X \leq x_n]\right) = \mathbb{P}([X \leq a]) = F(a).$$

3. Sea (x_n) una sucesión decreciente de números reales tales que

$$\lim_{n\to\infty} x_n = -\infty.$$

Entonces la sucesión de eventos $[X \leq x_n]$ es decreciente y

$$\bigcap_{n=1}^{\infty} [X \leq x_n] = \emptyset$$

por lo tanto,

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mathbb{P}([X \leq x_n]) = \mathbb{P}(\emptyset) = 0,$$

esto prueba que

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

Análogamente, si (x_n) es una sucesión creciente de números reales tales que

$$\lim_{n\to\infty} x_n = \infty,$$

la sucesión de eventos $[X \leq x_n]$ es creciente y

$$\bigcup_{n=1}^{\infty} [X \leq x_n] = \Omega.$$

Por lo tanto,

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mathbb{P}([X \leq x_n]) = \mathbb{P}(\Omega) = 1,$$

esto prueba que

$$\lim_{x\to\infty} F(x) = 1.$$

■

VARIABLES ALEATORIAS DISCRETAS.

DEFINICIÓN:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $X : \Omega \to \mathbb{R}$ una variable aleatoria. Diremos que X es discreta si su rango es finito ó numerable. Esto equivale a pedir que exista un conjunto finito ó numerables de valores $\{x_n\}$ tal que

$$\sum_{n} \mathbb{P}([X = x_n]) = 1.$$

OBSERVACIÓN:

Sea

$$R(X) = \{x_1, x_2, \cdots, x_n, x_{n+1}, \cdots\}$$

el Rango de X, el cual es finito ó numerable, si x es un número real que no pertence a $R(X)$, entonces $\mathbb{P}([X = x]) = 0$, ya que el evento $[X = x] = \emptyset$.

Si F es la función de distribución de X, x un número real cualquiera, $\mathbb{P}([X = x]) =$ salto de F en x, puesto que

$$[X = x] = \bigcap_{n=1}^{\infty} [a_n < X \leq b_n]$$

donde (b_n) es una sucesión que decrece con límite igual a x y (a_n) es un sucesión que crece con límite igual a x. Por otra parte

$$\begin{aligned} \mathbb{P}([a_n < X \leq b_n]) &= \mathbb{P}([X \leq b_n]) - \mathbb{P}([X \leq a_n]), \\ \mathbb{P}([X = x]) &= \lim_{n\to\infty} \mathbb{P}([X \leq b_n]) - \lim_{n\to\infty} \mathbb{P}([X \leq a_n]) = \text{salto de } F \text{ en } x \end{aligned}$$

luego, si x no está en el rango de X, el salto de F en x es cero, es decir, F es continua en los puntos que no están en el rango de X.

Si $x = x_i \in R(X)$, $\mathbb{P}([X = x_i]) =$ salto de F en x_i es diferente de cero, por lo tanto el conjunto de puntos de discontinuidad de la función de distribución F de una variable aleatoria X, discreta, coincide con su rango.

Notemos además que, F es constante en los intervalos de la forma $[x_i, x_{i+1})$ ya que si $x \in [x_i, x_{i+1})$

$$F(x) = \mathbb{P}([X \leq x]) = \mathbb{P}([X \leq x_i]) = F(x_i),$$

y así, F es una función en escalera, con saltos en los puntos del rango de X.

Sea $p_i = \mathbb{P}([X = x_i])$, lo cual representa la función de probabilidad de la variable aleatoria X, discreta. Entonces:

$$F(x) = \mathbb{P}([X \leq x]) = \sum_{x_i \leq x} p_i$$

donde la suma se extiende a aquellos valores de i para los cuales $x_i \leq x$.

DEFINICIÓN:

Llamamos función de probabilidad de una variable aleatoria X, discreta, a la función g_x (ó simplemente g), definida por

$$g_x(x) = \mathbb{P}([X = x]) = \text{ salto de } F \text{ en } x,$$

donde F es la función de distribución de X, $x \in \mathbb{R}$.

PROPIEDADES:

Si $g(x) = \mathbb{P}([X = x])$ es la función de probabilidad de una variable aleatoria X, discreta, entonces:

1. $g(x) = \mathbb{P}([X = x]) = 0$, si x no está en el rango de X.
2. $0 \leq g(x) = \mathbb{P}([X = x]) \leq 1$, si $x \in R(X)$.
3. $\displaystyle\sum_{x\in\mathbb{R}} g(x) = \sum_{x\in R(X)} g(x) = 1$

   pues,

$$\sum_{x\in R(X)} \mathbb{P}([X = x]) = \mathbb{P}\left(\bigcup_{x\in R(X)} [X = x]\right) = \mathbb{P}(\Omega) = 1$$

4. Si A es un Boreliano de $\mathbb{R}$, (en particular, un intervalo),

$$\mathbb{P}([X \in A]) = \sum_{x \in A} \mathbb{P}([X = x]) = \sum_{x \in A \cap R(X)} \mathbb{P}([X = x])$$

su función de distribución, es

$$F(x) = \mathbb{P}([X \leq x]) = \sum_{x_i \leq x} \mathbb{P}([X = x_i]) = \sum_{x_i \leq x} g(x_i).$$

EJEMPLOS:

1. Lanzamos dos dados distinguibles y llamamos X la suma de los resultados. Esta suma puede tomar once valores distintos y si suponemos que los dados están balanceados, podemos calcular la probabilidad de cada uno de los resultados.
$$R(X) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$
y su función de probabilidad viene dada por:
$$\begin{array}{l} p_2 = p_{12} = \frac{1}{36},\ p_3 = p_{11} = \frac{2}{36},\ p_4 = p_{10} = \frac{3}{36} \\ p_5 = p_9 = \frac{4}{36},\ p_6 = p_8 = \frac{5}{36},\ p_7 = \frac{6}{36} \\ \text{donde} \qquad p_i = \mathbb{P}([X = i]) = g_x(i), i = 2, 3, 4, \cdots, 11, 12 \end{array}$$
(Haga el gráfico de $g(x)$ y de $F(x)$)

2. Elegimos al azar un número en el intervalo $[0, 1)$ y definimos $X(\omega)$, como la tercera cifra después de la coma, en el desarrollo decimal de ω. Los posibles valores de $X(\omega)$ ó su rango, es
$$R(X) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$
y cada uno tiene probabilidad $\frac{1}{10}$, es decir:
$$g(x_i) = p_i = \mathbb{P}([X = x_i]) = \frac{1}{10},\ i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.$$
(Haga el gráfico de $g(x)$ y de $F(x)$).

DISTRIBUCIONES DISCRETAS FAMOSAS.

Distribución Binaria ó de Bernoulli.

Consideremos una experiencia aleatoria en la cual hay sólo dos resultados de interés, que llamaremos: “Exito” y “Fracaso”.

El ejemplo inmediato es el lanzmiento de una moneda, en donde podemos llamar Exito (E) la aparición de cara y Fracaso (F) la aparición de sello, (o viceversa).

Una experiencia aleatoria de esta naturaleza, se llama, prueba de Bernoulli; su espacio muestral consta únicamente de dos puntos.

Es frecuente definir sobre este espacio muestral una variable aleatoria X que tome el valor 1 si el resultado es un éxito y el valor 0 si el resultado es un fracaso.

Si $p \in (0, 1)$ es la probabilidad asignada a un éxito, la función de probabilidad de la variable aleatoria X, será:
$$g_b(1; p) = \mathbb{P}([X = 1]) = p,\ g_b(0; p) = \mathbb{P}([X = 0]) = 1 - p.$$

DISTRIBUCIÓN BINOMIAL:

Consideremos la experiencia aleatoria que consiste en realizar N pruebas independientes de Bernoulli; por ejemplo, lanzar una moneda N veces.

Si $p \in (0,1)$ es la probabilidad de obtener éxito en cada prueba, podemos preguntarnos, cuál es la distribución del número total de éxitos en las N pruebas?

Sea X el número total de éxitos en la N pruebas, el rango de X es

$$R(X) = \{0, 1, 2, 3, \cdots, N\}.$$

El espacio muestral Ω asociado a esta experiencia está constituído por $N-$uplas cuyas componentes son "éxito" ó "fracaso", luego, su cardinal es 2^N.

Observemos que si $p = \frac{1}{2}$, todos los eventos elementales de Ω, serán equiprobables y se tendría que

$$\mathbb{P}([X = n]) = \frac{Card([X = n])}{2^N} = \frac{\binom{N}{n}}{2^N},\ n = 0, 1, 2, \cdots, N.$$

Pero si p no es $\frac{1}{2}$, los eventos elementales no serán todos igualmente probables, su probabilidad dependerá del mayor ó menor número de éxitos presentes en la N -upla. Si en un resultado particular hay n éxitos y N-n fracasos, como cada prueba es independiente, la probabilidad asociada a este resultado es

$$p^n(1-p)^{N-n},$$

entre los 2^N puntos del espacio muestral Ω, existen $\binom{N}{n}$ que contienen n éxitos y $N-n$ fracasos, luego, la probabilidad del evento $[X = n]$ será la suma de las probabilidades de estos $\binom{N}{n}$ eventos elementales, cada uno de los cuales tiene probabilidad $p^n(1-p)^{N-n}$, es decir,

$$\mathbb{P}([X = n]) = \binom{N}{n} p^n(1-p)^{N-n},\ n = 0, 1, 2, \cdots, N.$$

La distribución de esta variable aleatoria X recibe el nombre de "Distribución Binomial" de parámetros p y N.

Su función de probabilidad es

$$g_B(n; N, p) = \mathbb{P}([X = n]) = \binom{N}{n} p^n(1-p)^{N-n},\ n = 0, 1, 2, \cdots, N.$$

N es el número de pruebas independientes de Bernoulli y p es la probabilidad de éxito

EJEMPLO: Se extraen con reposición cinco cartas de un juego de barajas. Sea X el número de diamantes en la muestra. Cuál es la probabilidad que haya exactamente 2 diamantes entre las 5 cartas?. Cuál es la probabilidad que haya a lo sumo 2 diamantes?

Observemos que si el muestreo es con reposición, las $N = 5$ extracciones representan $N = 5$ pruebas independientes de Bernoulli, con probabilidad de éxito $p = \frac{1}{4}$, así la variable aleatoria X tiene distribución Binomial de parámetros $N = 5, p = \frac{1}{4}$.

Para responder la primera pregunta, calculamos

$$\mathbb{P}([X = 2]) = \binom{5}{2} \left(\frac{1}{4}\right)^2 (1 - \frac{1}{4})^3 = 0,264.$$

Para responder la segunda pregunta tenemos que:

$$\mathbb{P}([X \leq 2]) = \sum_{n=0}^{2} \mathbb{P}([X = n]) = \sum_{n=0}^{2} \binom{5}{n} \left(\frac{1}{4}\right)^n \left(\frac{3}{4}\right)^{5-n} = 0,897.$$

DISTRIBUCIÓN GEOMÉTRICA.

Si llevamos a cabo una serie de pruebas independientes de Bernoulli, nos podemos preguntar, cuál es la distribución del número de puebas efectuadas antes de la obtención del primer éxito?

Si

$$X = \text{"número de fracasos consecutivos seguidos de un éxito"},$$

nuestro problema es calcular la probabilidad del evento $[X = n]$ donde $n = 0, 1, 2, 3, \cdots$.

Note que el rango de X es numerable, así como el espacio muestral Ω asociado a esta experiencia, ya que sus eventos elementales están constituídos por sucesiones de la forma

$$FFFF \overset{n}{\smile}{\cdots} FE,\ n = 0, 1, 2, \cdots;$$

si llamamos p la probabilidad de éxito, la probabilidad de obtener n fracasos consecutivos, seguidos de un éxito, en virtud de la independencia de las pruebas, es

$$(1 - p)^n p.$$

Esta distribución recibe el nombre de "Distribución Geométrica" de parámetro p, su función de probabilidad la denotamos por

$$g_G(n; p) = \mathbb{P}([X = n]) = (1 - p)^n p;\ n = 0, 1, 2, \cdots,$$

donde p representa la probabilidad de éxito en cada una de las pruebas independientes de Bernoulli y n es el número de fracasos consecutivos antes de obtener el primer éxito.

OBSERVACIÓN: Si en lugar de investigar el número de fracasos antes del primer éxito, nos interesamos en el número de pruebas necesarias para obtener éxito por primera vez, definamos

$$Y = \text{"el n}^o \text{ de pruebas necesarias para obtener éxito por } 1^a \text{ vez"}$$

esta nueva variable se relaciona con la anterior por

$$Y = X + 1$$

así, $R(Y) = \{1, 2, 3, \cdots, \}$, la distribución de Y es

$$\mathbb{P}([Y = k]) = \mathbb{P}([X + 1 = k]) = \mathbb{P}([X = k - 1]) = (1 - p)^{k-1} p$$

con $k = 1, 2, 3, \cdots$; representa la probabilidad de efectuar k pruebas para obtener el primer éxito.

EJEMPLO: Supongamos que un organismo viviente tiene probabilidad $p = 10^{-4}$ que se produzca una mutación en su descendencia. Si llamamos X el número de generaciones consecutivas que preceden la primera mutación genética, X tendrá distribución geométrica de parámetro $p = 10^{-4}$, siendo

$$\mathbb{P}([X = n]) = (0,9999)^n 10^{-4};\ n = 0, 1, 2, 3, \cdots.$$

DISTRIBUCIÓN HIPERGEOMÉTRICA:

Si realizamos un muestreo al azar,de tamaño k, sin reemplazo, elegidos en una población que contiene n objetos, de los cuales r son del tipo I (por ejemplo, defectuosos) y $n - r$ son del tipo II (por ejemplo, en buen estado), la variable aleatoria $X =$"n^o de objetos de tipo I en la muestra", tiene como función de probabilidad

$$\mathbb{P}([X = j]) = \frac{\binom{r}{j}\binom{n-r}{k-j}}{\binom{n}{k}},\ j = 0, 1, 2, \cdots, k \wedge r,$$

usando propiedades de los números combinatorios, es posible reescribir la fórmula anterior, como:

$$\mathbb{P}([X = j]) = \frac{\binom{k}{j}\binom{n-k}{r-j}}{\binom{n}{r}},\ j = 0, 1, 2, \cdots, k \wedge r,$$

EJEMPLO: Consideremos una población de 100 personas, de las cuales 10 tienen miopía. La probabilidad que haya a lo sumo dos personas miopes en un grupo de diez escogidos al azar, sin reemplazo, es

$$\mathbb{P}([X \leq 2]) = \sum_{j=0}^{2} \frac{\binom{10}{j}\binom{90}{10-j}}{\binom{100}{10}} = 0,94.$$

DISTRIBUCIÓN DE POISSON.

Un volumen de agua (ó de aire) contiene partículas puntuales en suspensión, distribuídas al azar, pero uniformemente en todo el volumen. Si aislamos un volumen V de fluído, el número total de partículas contenidas en dicho volumen, es una variable aleatoria.

La tela que se produce en un telar presenta defectos puntuales distribuídos al azar. Si cortamos una longitud L, el número de defectos incluídos en ella será una variable aleatoria.

Las llamadas telefónicas que llegan a una central, ocurren al azar, pero distribuídas uniformemente en el tiempo. Si consideramos un tiempo T, el número de llamadas que ocurren en ese tiempo, es una variable aleatoria.

Para resolver este tipo de problemas, precisaremos algunos términos:

1. Observemos que en todos los casos citados tenemos un "continuo" que puede ser, volumen, longitud, tiempo, etc., dentro del cual hay puntos aislados al azar pero con cierta "uniformidad" de distribución, lo que queremos hallar es la distribución del número total de tales puntos en un "segmento" (de longitud, volumen, tiempo,etc.,) finito del continuo.

2. Decir que los puntos están distribuídos al azar, pero uniformemente, significa que si tomamos un segmento de longitud, volumen, tiempo,etc., tan pequeño, que la probabilidad que dicho segmento presente dos ó más puntos pueda despreciarse, la probabilidad que presente solamente un punto, sea proporcional a su tamaño y no dependa de la ubicación en el continuo. Esto significa, que si la longitud ΔS de uno de esos segmentos es suficientemente pequeña, la probabilidad que presente dos ó más puntos es despreciable y la probabilidad que presente exactamente un punto es

$$\lambda \cdot \Delta S,$$

siendo λ la constante de proporcionalidad.

3. Se supone que la ocurrencia ó no de un punto en un segmento S_1, es probabilísticamente independiente de la ocurrencia ó no de un punto en otro segmento S_2 sin puntos comunes con S_1.

En base a estas observaciones, si consideramos un segmento de longitud S y lo dividimos en un número N suficientemente grande de subsegmentos, de igual longitud $\frac{S}{N}$, cada subsegmento contendrá uno ó ningún punto, siendo

$$\lambda\frac{S}{N} \text{ y } (1-\lambda\frac{S}{N})$$

las probabilidades que contenga uno ó ningún punto, respectivamente.

Si consideramos uno a uno los segmentos de extremos $k\frac{S}{N}$ y $(k+1)\frac{S}{N}$, con $k = 0,1,2,3,\cdots,N-1$; los examinamos y observamos si contienen ó no un punto, este examen puede ser considerado como una prueba de Bernoulli, en la cual un éxito corresponde a encontrar un punto. Nos interesa la distribución del número total de éxitos, es decir la distribución del número total de puntos en el segmento de longitud S.

Si $X =$ " número total de puntos en el segmento S ", esta variable tendrá distribución Binomial de parámetros $N,\ p = \lambda\frac{S}{N}$, luego,

$$\mathbb{P}([X = n]) = \binom{N}{n}\left(\lambda\frac{S}{N}\right)^n (1-\lambda\frac{S}{N})^{N-n},\ n = 0,1,2,\cdots,N.$$

En todo este proceso hemos supuesto que N es tan grande que cada subsegmento tiene a lo sumo un punto, hagamos entonces tender N a infinito:

$$\mathbb{P}([X = n]) = \frac{N}{N}\cdot\frac{N-1}{N}\cdot\frac{N-2}{N}\cdot\cdots\cdot\frac{N-n+1}{N}\cdot\frac{(\lambda S)^n}{n!}(1-\lambda\frac{S}{N})^{N-n}$$

Si $N \to \infty$, obtendremos que la distribución de la variable aleatoria X tiende a:

$$\mathbb{P}([X = n]) = \frac{(\lambda S)^n}{n!} e^{-\lambda S}; n = 0, 1, 2, \cdots,$$

lo cual representa, la probabilidad que se encuentren n puntos en un segmento de longitud S. Si hacemos $S = 1$, obtendremos la distribución del número de puntos por unidad de magnitud (unidad de volumen, de tiempo, de longitud, etc.), es decir, la probabilidad que exactamente hayan n puntos por unidad de magnitud es

$$\mathbb{P}([X = n]) = \frac{\lambda^n}{n!} e^{-\lambda}; n = 0, 1, 2, \cdots,$$

Esta distribución recibe el nombre de "Distribución de Poisson" de parámetro λ.

DEFINICIÓN. Diremos que la variable aleatoria X tiene "Distribución de Poisson" de parámetro λ, si su función de probabilidad es

$$g_P(n; \lambda) = \mathbb{P}([X = n]) = \frac{\lambda^n}{n!} e^{-\lambda}; n = 0, 1, 2, \cdots,$$

EJEMPLOS.

1. La Distribución de Poisson de parámetro λ se obtiene como límite de la Distribución Binomial de parámetros N, p, cuando $N \to \infty$ y $p \to 0$, pero manteniendo el producto $Np = \lambda$. (Ejercicio).

2. En una concurrida intersección de tráfico, la probabilidad que un automóvil tenga un accidente es $p = 10^{-4}$, sin embargo, durante cierta hora del día, entre las 4 pm. y las 6 pm., un gran número de automóviles pasa por la intersección, digamos 1000 automóviles. Supongamos que el valor p es el mismo para todos ellos y que si un automóvil tiene ó no un accidente, no depende de lo que suceda a cualquier otro (lo cual no es una suposición real pero simplifica el probelma). Bajo dichas condiciones, nos preguntamos, cuál es la probabilidad que dos ó más accidentes ocurran durante este período?

   Si llamamos X el número de accidentes entre los 1000 autos que llegan a la intersección, su distribución es Binomial de parámetros $p = 10^{-4}$, $N = 10^3$, luego

   $$\mathbb{P}([X \geq 2]) = 1 - \mathbb{P}([X \leq 1]) = 1 - (0,9999)^{1000} - (0,9999)^{999} \cdot (0,0001)$$

   Hacer este cálculo presenta una dificultad considerable, pero como N es grande y p pequeño, podemos aproximar esta distribución a la de Poisson de parámetro $\lambda = Np = 10^{-1}$, así

   $$\mathbb{P}([X \geq 2]) = 1 - \mathbb{P}([X \leq 1]) \simeq 1 - e^{-0.1} - (0.1)e^{-0.1} \simeq 0.1856463.$$

3. **DESINTEGRACIÓN RADIOACTIVA.**

   Una sustancia radioactiva emite partículas α, el número de partículas que llegan a una pantalla durante un tiempo t, es una variable aleatoria y es uno de los mejores ejemplos de variable aleatoria con distribución de Poisson.

   Aunque una sustancia se desgasta, emitiendo entonces menor cantidad de partículas, para que esta disminución sea notoria, deben pasar años. Por lo tanto, si se hace un estudio mucho más corto, (una semana ó un mes) se puede considerar que las condiciones del experimento no varían.

   En cierta oportunidad, se observó una sustancia radioactiva durante $N = 2608$ intervalos de tiempo, cada uno de $t = 7,5$ segundos.

Se obtuvieron los siguientes resultados:

k	N_k	$N \cdot g_P(k; 3,870)$
0	57	$54,399$
1	203	$210,523$
2	383	$407,361$
3	525	$525,496$
4	532	$508,418$
5	408	$393,515$
6	273	$253,817$
7	139	$140,325$
8	45	$67,882$
9	27	$29,189$
$k \geq 10$	16	$17,075$

$g_P(k\,;\lambda t)$ es la función de probabilidad de la Distribución de Poisson de parámetro λt, es decir,

$$g_P(k\,;\lambda t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!};\ k = 0, 1, 2, 3, \cdots,$$

y representa la probabilidad de encontrar exactamente k partículas α en un intervalo de tiempo t.

Si N es el número de veces que repetimos el experimento y N_k es el número de veces que se observan k partículas, se tiene que

$$N = N_0 + N_1 + N_2 + \cdots + N_k + \cdots$$

El número total de partículas observadas en las N repeticiones de la experiencia es

$$T = 0 \cdot N_0 + 1 \cdot N_1 + 2 \cdot N_2 + \cdots + k \cdot N_k + \cdots$$

El promedio del número de partículas observadas en las N pruebas, es

$$\frac{T}{N}$$

Note que $\dfrac{N_k}{N}$ es la frecuencia con que ocurre el evento " se observaron k partículas " en las N pruebas, luego, se espera que si N es grande, esta frecuencia sea aproximadamente igual a la probabilidad de observar k partículas en el intervalo de tiempo t, es decir

$$\frac{N_k}{N} \simeq g_P(k\,;\lambda t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}.$$

Por lo tanto,

$$\frac{T}{N} = \frac{N_1 + 2 \cdot N_2 + \cdots + k \cdot N_k + \cdots}{N} = \sum_{k=0}^{\infty} \frac{k \cdot N_k}{N} \simeq \sum_{k=0}^{\infty} k e^{-\lambda t}\frac{(\lambda t)^k}{k!} = \lambda t.$$

Así, aproximando λt por $\dfrac{T}{N}$, podemos estimar λ a partir de las observaciones y comparar la teoría con el experimento.

En nuestra tabla, el número de partículas es

$$T = \sum_{k \geq 0} k N_k = 10086$$

el promedio es

$$\frac{T}{N} = 3,870$$

así,

$$\lambda t \simeq \frac{T}{N} = 3,870.$$

Observe que los valores teóricos $N \cdot g_P(k\,; 3,870)$ están bastante cerca de los valores experimentales N_k. Hay métodos estadísticos que permiten decidir si ciertos valores experimentales pueden considerarse como correspondientes a una distribución teórica.

4. Las llamadas que se reciben en una central telefónica, por minuto, tienen distribución de Poisson con parámetro $\lambda = 4$. Si la central puede manejar un máximo de seis llamadas por minuto, cuál es la probabilidad que la central sea insuficiente para atender las llamadas que llegan en un minuto?

   Sea X el número de llamadas que se reciben en un minuto. Por lo tanto

$$\mathbb{P}\left([X > 6]\right) = 1 - \mathbb{P}\left([X \leq 6]\right) = 1 - \sum_{k=1}^{6} e^{-4}\frac{4^k}{k!} = 1 - 0,889 = 0,11.$$

ESPERANZA MATEMÁTICA

INTRODUCCIÓN:

Sean $x_1, x_2, x_3, x_4, \cdots, x_n$, n valores reales, la expresión

$$\frac{x_1 + x_2 + x_3 + x_4 + \cdots + x_n}{n}$$

representa el promedio aritmético ordinario de los n valores.

El conjunto $\{x_1, x_2, x_3, x_4, \cdots, x_n\}$ se puede pensar como el rango de una variable aleatoria X cuya distribución viene dada

por

$$g_X(x) = \mathbb{P}(X = x_i) = \frac{1}{n},\ i = 1, 2, \cdots .n;$$

el promedio de los n valores se puede escribir como

$$\sum_{i=1}^{n} x_i \mathbb{P}(X = x_i).$$

Este promedio puede generalizarse al caso en que los puntos del rango de X no son igualmente probables, definiendo el promedio de esos valores como

$$\sum_{i=1}^{n} x_i \mathbb{P}(X = x_i),$$

donde $\mathbb{P}(X = x_i)$ es diferente a $\dfrac{1}{n}$ para algún x_i.

La esperanza de una variable aleatoria X generaliza la idea de "promedio".

ESPERANZA DE UNA VARIABLE ALEATORIA DISCRETA.

Sea X una variable aleatoria discreta de rango $R(X)$, la esperanza de X (ó valor esperado de X ó media de X ó promedio de X) se denota por $\mathbb{E}(X)$ y se define como:

$$\mathbb{E}(X) = \sum_{x \in R(X)} x \mathbb{P}(X = x)$$

siempre y cuando esta serie sea absolutamente convergente, es decir, si

$$\sum_{x \in R(X)} |x|\, \mathbb{P}(X = x) < \infty,$$

en caso contrario diremos que X no tiene esperanza finita.

PROPIEDADES.

1. Si $X \geq 0$ y $\mathbb{E}(X)$ existe entonces $\mathbb{E}(X) \geq 0$.
2. Si X tiene rango acotado entonces $\mathbb{E}(X)$ existe.
3. Si A es un evento y $X = \mathcal{X}_A$ donde
$$\mathcal{X}_A(\omega) = \left\{ \begin{array}{ll} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \in A^c \end{array} \right\}$$
entonces
$$\mathbb{E}(X) = \mathbb{P}(A)$$
4. Sea $g : \mathbb{R} \to \mathbb{R}$ una función suficientemente regular como para que
$$Y = g(X)$$
sea una variable aleatoria discreta. Si $\mathbb{E}(Y)$ existe, entonces
$$\mathbb{E}(Y) = \sum_{x \in R(X)} g(x)\mathbb{P}(X = x)$$
(Note que esta propiedad permite calcular la esperanza de Y sin necesidad de conocer su distribución, para formalizar la demostración tome en cuenta el caso en que g no sea inyectiva).
5. Si $\mathbb{E}(X)$ existe, entonces
$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$$
(Esto es consecuencia de la desigualdad triangular).
6. Si λ es una constante y $\mathbb{E}(X)$ existe, entonces $\mathbb{E}(\lambda X)$ existe y
$$\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$$
7. (La propiedad que daremos a continuación podrá ser demostrada cuando estudiemos variables aleatorias conjuntas).

   Si X e Y son dos variables aleatorias de esperanza finita, entonces la variable aleatoria $X + Y$ tiene esperanza finita y
$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$
8. Si X e Y son dos variables aleatorias tales que $X \leq Y$, entonces
$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$
si ambas esperanzas existen.
9. Si k es una constante y X tiene esperanza finita y $Y = k$ entonces
$$\begin{array}{l} \mathbb{E}(Y) = \mathbb{E}(k) = k \\ \mathbb{E}(X + k) = \mathbb{E}(X) + k. \end{array}$$

(Dejaremos como ejercicio para la práctica, la demostración de estas propiedades).

MOMENTOS DE UNA VARIBLE ALEATORIA DISCRETA.

Si X es una variable aleatoria discreta y
$$\mathbb{E}(|X|^n) < \infty$$
se define el momento de orden n de X como
$$\mathbb{E}(X^n) = \sum_{x \in R(X)} x^n \mathbb{P}(X = x).$$

VARIANZA DE UNA VARIBLE ALEATORIA DISCRETA.

Si X es una variable aleatoria discreta, definimos la varianza de X como la esperanza de $(X - \mathbb{E}(X))^2$ (si existe), denotamos esta cantidad como

$$V(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \sum_{x \in R(X)} (x - \mathbb{E}(X))^2$$

DESVIACIÓN ESTÁNDAR

Si X es una variable aleatoria discreta de varianza finita, definimos la desviación estándar por

$$D(X) = \sqrt{V(X)}.$$

OBSERVACIÓN.

Sea a un número real

$$h(a) = \mathbb{E}\left((X - a)^2\right).$$

Este valor es una especie de "promedio" de la concentración de los valores de la variable aleatoria X alrededor del punto a.

La función $h(a)$ tienen un mínimo si $a = \mathbb{E}(X)$, en efecto,

$$h'(a) = -2\mathbb{E}(X) + 2a$$

por lo tanto en $a = \mathbb{E}(X)$ hay un mínimo por ser $h(a)$ una parábola con vértice hacia abajo. En ese punto mínimo

$$h(\mathbb{E}(X)) = V(X).$$

Si $V(X) = 0$, entonces $\mathbb{P}(X = \mathbb{E}(X)) = 1$, ya que si $\mathbb{E}\left((X - \mathbb{E}(X))^2\right) = 0$, como $(X - \mathbb{E}(X))^2 \geq 0$, por definición de esperanza, solamente es posible que

$$X - \mathbb{E}(X) = 0$$

con probabilidad 1.

PROPIEDADES.

1. $V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
2. Si k es una constante entonces

$$V(X + k) = V(X).$$
$$D(X + k) = D(X).$$
$$V(kX) = k^2 V(X).$$

(Estas propiedades quedan como ejercicio para la práctica).

EJEMPLOS:

A) Sean X, Y dos variables aleatorias de rangos:

$$R(X) = \{-1, 0, 1\}, \quad R(Y) = \{-5, 0, 5\}$$

con probabilidad de 1/3 para cada valor.

Las esperanzas y varianzas respectivas son:

$$\begin{aligned}
&\mathbb{E}(X) = -1 \cdot 1/3 + 0 \cdot 1/3 + 1 \cdot 1/3 = 0 \\
&V(X) = \mathbb{E}(X^2) - \mathbb{E}(X) = -1 \cdot 1/3 + 0 \cdot 1/3 + 1 \cdot 1/3 - 0 = 2/3 \\
&\mathbb{E}(Y) = -5 \cdot 1/3 + 0 \cdot 1/3 + 5 \cdot 1/3 = 0 \\
&V(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y) = 25 \cdot 1/3 + 0 \cdot 1/3 + 25 \cdot 1/3 - 0 = 50/3
\end{aligned}$$

La variable aleatoria X tiene los valores más concentrados alrededor de 0 que Y pues $V(X) < V(Y)$.

B) Sea X una variable aleatoria con distribución de Bernoulli de parámetro p. Hallemos la esperanza y la varianza de esta variable aleatoria:

$$\mathbb{E}(X) = 1 \cdot \mathbb{P}(X = 1) + 0 \cdot \mathbb{P}(X = 0) = p$$
$$V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 1^2 \cdot p + 0^2 \cdot (1-p) - p^2 = p - p^2 = p(1-p).$$

C) Sea X una variable aleatoria con distribución Binomial de parámetros p, N. Hallemos la esperanza y la varianza de esta variable aleatoria:

$$\mathbb{E}(X) = \sum_{n=0}^{N} n\mathbb{P}(X = n) = \sum_{n=1}^{N} n\binom{N}{n}p^n(1-p)^{N-n} =$$
$$\sum_{n=1}^{N} \frac{N(N-!)!}{(n-1)!\,(N-1-(n-1))!} p \cdot p^{n-1} \cdot (1-p)^{(N-1-(n-1))} =$$
$$p \cdot N \cdot \sum_{k=0}^{N} \binom{N-1}{k} p^k (1-p)^{N-1-k} = pN$$

para hallar la varianza, calculemos primero

$$\mathbb{E}(X^2) = \sum_{n=0}^{N} n^2\mathbb{P}(X = n) = \sum_{n=1}^{N} n^2\binom{N}{n}p^n(1-p)^{N-n} =$$
$$\sum_{n=1}^{N} [n(n-1) + n] \binom{N}{n}p^n(1-p)^{N-n} =$$
$$\sum_{n=1}^{N} n(n-1)\binom{N}{n}p^n(1-p)^{N-n} + \sum_{n=1}^{N} n\binom{N}{n}p^n(1-p)^{N-n} =$$
$$\sum_{n=1}^{N} n(n-1)\binom{N}{n}p^n(1-p)^{N-n} + Np = \sum_{n=2}^{N} n(n-1)\binom{N}{n}p^n(1-p)^{N-n} + Np =$$
$$\sum_{n=2}^{N} n(n-1)\frac{N(N-1)(N-2)!}{n(n-1)(n-2)!\,(N-2-(n-2))!} p^{n-2} \cdot p^2(1-p)^{N-2-(n-2)} + Np =$$
$$N(N-1)p^2 \sum_{k=0}^{N-2} \binom{N-2}{k} p^k (1-p)^{N-2-k} + Np = N(N-1)p^2 + Np$$

De aquí se deduce que

$$V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = N(N-1)p^2 + Np - N^2p^2 = Np(1-p).$$

FUNCIÓN GENERATRIZ O TRANSFORMADA GEOMÉTRICA.

Sea X una variable aleatoria discreta de rango $R(X)$, definimos, para cada número real a, la siguiente función:

$$T(a) = \mathbb{E}(a^X) = \sum_{n \in R(X)} a^n \mathbb{P}(X = n)$$

siempre y cuando esta serie converja absolutamente.

Esta función se denomina transformada geométrica ó función generatriz de momentos de la variable aleatoria X, (veremos a continuación que a partir de ella podemos obtener la esperanza y la varianza de la variable aleatoria).

Por otro lado, puesto que esta función es un polinomio en potencias de a ó una serie de potencias, a partir de ella podemos reconstruir la función de probabilidad de X, basta desarrollar $T(a)$ en potencias de a e igualar el coeficiente de a^n con $\mathbb{P}(X = n)$. Es decir, darse la transformada geométrica de una variable aleatoria discreta equivale a darse la distribución de dicha variable.

PROPIEDADES DE LA TRANSFORMADA

1. $T(1) = \sum_{n \in R(X)}^{N} \mathbb{P}(X = n) = 1$

2. $T'(1) = \sum_{n \in R(X)}^{N} n\mathbb{P}(X = n) = \mathbb{E}(X)$

3. $T''(1) = \sum_{n \in R(X)}^{N} n(n-1)\mathbb{P}(X = n) = \mathbb{E}(X^2) - \mathbb{E}(X)$, de aquí se obtiene que

$$V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = T''(1) + T'(1) - (T'(1))^2 .$$

OBSERVACIONES.

1. Valiéndonos de las propiedades de T, podemos calcular la esperanza y la varianza de una variable aleatoria discreta X, mediante simple derivación, evitando así el cálculo de series.

2. Si $R(X)$ es numerable (no finito), por ejemplo:

$$R(X) = \{0, 1, 2, \cdots\}$$

la serie que define $T(a)$ converge absolutamente si $|a| < 1$ y si $a = 1$, la serie converge ya que $T(1) = 1$.

EJEMPLOS:

1. Sea X una variable aleatoria con distribución de Bernoulli de parámetro p.

$$\begin{array}{l} T(a) = \mathbb{E}(a^X) = (1-p) + ap = p(a-1) + 1 \\ T'(a) = p, \quad T'(1) = p, \quad T''(a) = 0 = \mathbb{E}(X^2) - \mathbb{E}(X) \\ \mathbb{E}(X) = p, \quad V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = p - p^2 = p(1-p). \end{array}$$

2. Sea X una variable aleatoria con distribución Binomial de parámetros p, N.

$$\begin{array}{l} T(a) = \sum_{n=0}^{N} a^n \binom{N}{n} p^n (1-p)^{N-n} = \sum_{n=0}^{N} \binom{N}{n} (ap)^n (1-p)^{N-n} = (ap + 1 - p)^N \\ T'(a) = N(ap + 1 - p)^{N-1} p, \quad T'(1) = Np = \mathbb{E}(X) \\ T''(a) = N(N-1)(ap + 1 - p)^{N-2} p^2, \quad T''(1) = N(N-1)p^2 = \mathbb{E}(X^2) - Np \\ \mathbb{E}(X^2) = N(N-1)p^2 + Np \\ V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = N(N-1)p^2 + Np - N^2p^2 = Np(p-1). \end{array}$$

EJERCICIOS PARA LA PRÁCTICA.

1. Sea X una variable aleatoria con función de probabilidad dada por la siguiente tabla:

x_i	-3	-1	0	1	2	3	5	8
p_i	$0,1$	$0,2$	$0,15$	$0,2$	$0,1$	$0,15$	$0,05$	$0,05$

   (a) Cuál es el rango de X?

   (b) Calcule las probabilidades de los siguientes eventos:

      i. X es negativa ($X < 0$).

      ii. X es par (divisible entre 2).

      iii. X toma valores entre 1 y 8, ambos inclusive.

      iv. $\mathbb{P}(X = -3 \mid X \leq 0)$

      v. $\mathbb{P}(X \geq 3 \mid X > 0)$

2. Determine el valor de la constante A para que las siguientes sean funciones de probabilidad:

   (a) $\mathbb{P}(X = k) = \begin{cases} Ak & k = 1, 2, 3, \cdots n. \\ 0 & \text{en otro caso} \end{cases}$

   (b) $\mathbb{P}(X = k) = \begin{cases} \dfrac{A}{2^k} & k = 1, 2, 3, \cdots n. \\ 0 & \text{en otro caso} \end{cases}$

   (c) $\mathbb{P}(X = k) = \begin{cases} \dfrac{A}{3^k} & k = 1, 3, 5 \cdots n-1. \\ \dfrac{A}{4^k} & k = 2, 4, 6 \cdots n. \\ 0 & \text{en otro caso} \end{cases}$ si n es par .

3. Se lanza un par de dados no cargados, definamos las siguientes variables aleatorias:
   $X =$ "puntaje total"
   $Y =$ "puntaje del primer dado menos puntaje del segundo"
   $Z =$ "diferencia entre el puntaje mayor y el puntaje menor" ($Z = |Y|$).
   $W =$ "puntaje individual mayor".
   Para cada una de las variables aleatorias

   (a) hallar su función de probabilidad; haga el gráfico;

   (b) hallar su función de distribución; haga el gráfico.

4. Una variable aleatoria X discreta, de rango $\{0, 1, 2, 3, \cdots, \}$ tiene por función de probabilidad

$$p_n = \mathbb{P}([X = n]) = e^{-a}\frac{a^n}{n!}, \ a > 0.$$

   Compruebe que

$$\sum_{n=0}^{\infty} e^{-a}\frac{a^n}{n!} = 1.$$

5. Si X es una variable aleatoria discreta con función de probabilidad

$$p_n = \mathbb{P}([X = n]) = 2^{-n}, \ n = 1, 2, 3, \cdots$$

(a) Compruebe que
$$\sum_{n=1}^{\infty} p_n = 1$$
(b) Hallar $\mathbb{P}([X \text{ sea par}])$
(c) Hallar $\mathbb{P}([X \leq 4] \mid X \text{ es par})$.

6. Se realizan pruebas independientes de Bernoulli con probabilidad de éxito p. Si X es el número de intentos necesarios para completar M éxitos, halle la función de probabilidad de X. (Esta distribución se conoce con el nombre de distribución de Pascal, de parámetros M, p).

7. Demuestre que la distribución Binomial, satisface la siguiente fórmula de recurrencia:
$$g_B(n+1; N, p) = \frac{p \cdot (N-n)}{(1-p) \cdot (n+1)} g_B(n; N, p)$$
y mediante ella calcule numéricamente
$$g_B(n; 5, 0.3),\ n = 0, 1, 2, 3, 4, 5.$$

8. Demostrar que la distribución geométrica carece de memoria en el sentido que la distribución de
$$X - m, \text{ dado que } X \geq m,$$
donde m es un entero fijo ($m \geq 0$), es nuevamente geométrica, si X lo es y con el mismo parámetro. Es decir, demuestre que
$$\mathbb{P}([Y = n] \mid X \geq m) = \mathbb{P}([X = n]),\ n = 0, 1, 2, ..$$
donde $Y = X - m$.

9. Para determinar la efectividad de una nueva vacuna contra la gripe, se vacunan 10 personas que son observadas por un período de un año. De ellas, 8 no tuvieron gripe durante ese lapso. Si se sabe que la probabilidad de no tener gripe en un período de un año es $0,5$, cuál es la probabilidad que 8 ó más personas del grupo no hayan sufrido la enfermedad, si la vacuna no es efectiva?

10. Considere un cierto defecto en el metabolismo que ocurre en aproximadamente 1 de cada 100 nacimientos. Si cuatro niños nacen en cierto hospital el mismo día, calcule la probabilidad que

    (a) ninguno tenga el defecto
    (b) no más de uno tenga el defecto.

11. El número de carros que cruzan un puente durante un período fijo de tiempo es una variable aleatoria con distribución de Poisson. Si la probabilidad de que ningún carro cruce el puente en este período es $1/4$, halle una expresión para la probabilidad de que al menos dos carros lo crucen.

12. Un vendedor de periódicos compra cada periódico por Bs.300 y lo vende por Bs.400. Los que no vende los regresa al distribuidor y recibe Bs.200 por ellos. Supongamos que la distribución de la demenda D es
$$\mathbb{P}([D = k]) = e^{-10} \frac{10^k}{k!},\ k = 0, 1, 2, ..$$
Describa la variable aleatoria X que representa su ganancia diaria si compra 10 periódicos cada día.

13. Un llavero tiene cuatro llaves de apariencia similar pero sólo una de ellas, abre la puerta de cierta oficina. Se selecciona al azar una llave y se prueba, si no funciona se selecciona al azar una de las restantes y se prueba de nuevo. Sea X el número de llaves que se prueba antes de encontrar la que abre la puerta. Halle su distribución de probabilidad.

14. Calcular la transformada geométrica de las distribuciones:

(a) Geométrica de parámetro p.

(b) Poisson de parámetro λ

y deduzaca la esperanza y la varianza de cada una de ellas. (Trate de calcular para cada una, la esperanza y la varianza, sin utilizar transformada geométrica).

15. Sea X una variable aleatoria discreta de función de probabilidad

$$g_X(n) = \frac{1}{2^n}, \quad \text{si } n = 1, 2, 3, \cdots$$

Hallar

(a) $\mathbb{E}(X)$, $V(X)$

(b) $\mathbb{P}(X \geq 4 \mid X \text{ es par})$

(c) Si $Y = \dfrac{1}{X}$, hallar $\mathbb{E}(Y)$.

16. Sea X una variable aleatoria con transformada geométrica $P(a)$. En cuentre las transformadas geométricas de $X + 1$ y $2X$.

17. El número de rayos gama emitidos por segundo por cierta sustancia radiactiva es una variable aleatoria con una distribución de Poisson con $\lambda = 5.8$. Si un instrumento de registro queda fuera de operación cuando hay mas de 12 rayos por segundo, cuál es la probabilidad de que este instrumento quede fuera de operación durante cualquier segundo dado?.

Respuesta: 0.007

18. En una ciudad dada , el 6% de los conductores obtienen al menos una boleta de estacionamiento al año. Use la aproximación de Poisson a la distribución Binomial para determinar la probabilidad de que entre 80 conductores (elegidos aleatoriamente en esta ciudad)

(a) cuatro obtengan al menos una boleta de estacionamiento en un año dado (Resp. 0.182)

(b) al menos tres obtengan como mínimo una boleta de estacionamiento en un año dado (Resp. 0.857)

(c) cualquier número de ellos entre tres y seis inclusive obtengan una boleta de estacionamiento en un año dado. (Resp.0.648).

19. En una caja bancaria llegan clientes a una tasa promedio de 1.5 por minuto. Determine las probabilidades de

(a) a lo sumo cuatro clientes llegan en un minuto dado (Resp. 0.981)

(b) al menos tres clientes llegan en un intervalo de 2 minutos
(Resp.0.577)

(c) a lo sumo 15 clientes llegan en un intervalo de 6 minutos.
(Resp. 0.978)

Capítulo 5

VARIABLES ALEATORIAS CONTINUAS

INTRODUCCIÓN:

Las variables aleatorias que hemos estudiado anteriormente, son variables de rango a lo sumo numerable, como por ejemplo, el número de objetos defectuosos en una muestra de tamaño N.

Hay muchas situaciones, en las cuales las variables aleatorias que debemos considerar son continuas en lugar de discretas. Por ejemplo, si escogemos un punto al azar en un disco de radio R, con centro en el origen, interpretando la expresión al azar, como equivalente a "si A y B son subconjuntos del disco, con igual área (A, B en una familia de subconjuntos de $\mathbb{R}^2$, llamados Borelianos, los cuales denotamos por $\mathcal{B}(\mathbb{R}^2)$), si ω es el punto elegido al azar en $\Omega = \left\{\omega = (x, y) : x^2 + y^2 \leq R^2\right\}$, entonces

$$\mathbb{P}(\omega \in A) = \mathbb{P}(\omega \in B),$$

y la probabilidad de que un punto escogido esté en un subconjunto A del disco es proporcional al área de A:

$$\mathbb{P}(\omega \in A) = k\,|A|$$

donde k es una constante de proporcionalidad y $|A|$ es el área de A".

En particular, por ser Ω todo el disco, entonces

$$1 = \mathbb{P}(\Omega) = k\,|\Omega| = k\pi R^2,$$

luego

$$k = \frac{1}{\pi R^2}$$

de donde

$$\mathbb{P}(\{\omega : \omega \in A\}) = \frac{|A|}{\pi R^2}.$$

Definamos la variable aleatoria X como

$$X(\omega) = d(\omega, \vec{0}),$$

donde $d(\omega, \vec{0})$ representa la distancia del punto elegido al azar en el disco, al origen ó centro del disco. Calculemos su función de distribución:

Si $0 \leq x \leq R$, $\{\omega : X(\omega) \leq x\}$ es el disco del plano con centro en el origen y radio x.

Por lo tanto

$$F(x) = \begin{cases} \mathbb{P}([X \leq x] = \dfrac{\pi x^2}{\pi R^2} = \dfrac{x^2}{R^2}, \text{ si } 0 \leq x \leq R \\ \mathbb{P}([X \leq x] = 0, \quad \text{si } x < 0 \\ \mathbb{P}([X \leq x] = 1, \quad \text{si } x > R \end{cases}$$

la cual es una función continua.

DEFINICIÓN:

Diremos que una variable aleatoria X es continua, si su función de distribución

$$F(x) = \mathbb{P}([X \leq x]$$

es continua, lo cual equivale a decir que todos sus saltos son nulos, es decir:

$$\mathbb{P}([X = x] = 0 \qquad x \in \mathbb{R}.$$

OBSERVACIÓN:

De acuerdo a las definiciones que hemos dado, hay variables aleatorias que no son ni continuas ni discretas, pues no toman valores únicamente en un conjunto numerable y cuyas funciones de distribución no son continuas.

DENSIDADES:

Si X es una variable aleatoria y F es una función de distribución, decimos que F es absolutamente continua (ó que tiene densidad) si existe una función f no negativa, tal que

$$F(x) = \int_{-\infty}^{x} f(t)dt \qquad \text{para todo } x \in \mathbb{R}.$$

La función f se llama la densidad de la distribución F ó la densidad de la distribución de probabilidad de la variable aleatoria X.

PROPIEDADES:

1. $\int_{-\infty}^{\infty} f(t)dt = 1$, ya que $\lim\limits_{x\to\infty} F(x) = 1$.
2. Si f es continua en x_0, F es derivable en x_0 y

$$F'(x_0) = f(x_0)$$

(ejercicio, ver primer teorema fundamental del cálculo, Spivak, Cap.14).

OBSERVACIÓN:

La función f puede ser en general, bastante complicada, pero nos limitaremos a considerar funciones f suficientemente regulares como para que el estudiante pueda manejarse con el cálculo diferencial e integral que conoce: tales funciones serán continuas, salvo a lo sumo, en una cantidad numerable de puntos, por lo tanto se tendrá que $f(x) = F'(x)$, salvo a lo sumo en un conjunto numerable de puntos.

EJERCICIO.

Demostrar que si f es no negativa y $\int_{-\infty}^{\infty} f(t)dt = 1$, entonces la función

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

verifica todas las propiedades de una función de distribución (F es una función de distribución).

DEFINICIÓN:

f es una densidad si f es no negativa y $\int\limits_{-\infty}^{\infty} f(t)dt = 1$.

Con la misma notación anterior, si F tiene densidad f, entonces

$$\begin{aligned}&\mathbb{P}([a < X \leq b]) = \mathbb{P}([X \leq b]) - \mathbb{P}([X \leq a]) = \\ &F(b) - F(a) = \int\limits_{a}^{b} f(t)dt.\end{aligned}$$

Geométricamente, la probabilidad que la variable aleatoria X pertenezca al intervalo $(a, b]$ es el área comprendida entre la gráfica de la función f, el eje x y las rectas $x = a, x = b$.

OBSERVACIÓN:

Si F es continua, no tiene saltos, por lo tanto:

$$\mathbb{P}([a \leq X \leq b]) = \mathbb{P}([a < X < b]) = \mathbb{P}([a < X \leq b]) = \mathbb{P}([a \leq X < b])$$

DISTRIBUCIÓN UNIFORME.

Una variable aleatoria X tienen distribución uniforme en el intervalo $[a, b]$, si para cualquier intervalo I contenido en $[a, b]$

$$\mathbb{P}([X \in I])$$

es proporcional a la longitud de I, en particular

$$\mathbb{P}(X \in [a, b]) = \mathbb{P}([a \leq X \leq b]) = k(b - a)$$

siendo k la constante de proporcionalidad y $[a, b]$ el rango de X, por lo tanto

$$k = \frac{1}{b-a}, \quad \text{ya que } \mathbb{P}([a \leq X \leq b]) = 1.$$

La función de distribución correspondiente a la distribución uniforme es

$$F(x) = \begin{cases} \mathbb{P}([X \leq x]) = \mathbb{P}([a \leq X \leq x]) = \dfrac{x-a}{b-a}, & \text{si } a \leq x \leq b \\ 0, \quad \text{si } x < a \\ 1, \quad \text{si } x > b \end{cases}$$

La función de densidad correspondiente a F es

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{si } a \leq x \leq b \\ 0, & \text{si } x < a \\ 0, & \text{si } x > b \end{cases}$$

ya que verifica

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad \text{para todo } x \in \mathbb{R}.$$

DEFINICIÓN:

Diremos que una variable aleatoria X tiene distribución uniforme en el intervalo $[a, b]$ con $(a < b)$, si su densidad es

$$f(x) = \frac{1}{b-a} \quad \text{si } x \in [a, b],$$

usaremos la notación $X \sim \mathcal{U}[a, b]$.

EJEMPLO:

Imaginemos una ruleta cuya circunferencia tiene longitud c y cuyo puntero se puede detener en cualquier punto.

Fijemos un origen de referencia arbitrario sobre el borde, si la ruleta está bien construída, el puntero no tendrá preferencia por unos puntos más que por otros ó por unas zonas más que por otras.

Sea X la distancia sobre la circunferencia entre el punto sobre el cual se detiene el puntero y el origen O de referencia. Debido a la imparcialidad del puntero para detenerse en una zona determinada, la probabilidad de que se detenga en un "segmento" dado de la circunferencia, debe ser proporcional a su longitud.

Si tomamos como segmento toda la circunferencia, la probabilidad de caer en él, debe ser igual a 1 (uno) y como su longitud es c, se concluye que el coeficiente de proporcionalidad debe ser igual a $\frac{1}{c}$. En particular

$$F(x) = \mathbb{P}([X \leq x]) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x}{c} & \text{si } 0 \leq x \leq c \\ 1 & \text{si } x > c \end{cases}$$

La densidad de X es

$$f(x) = \frac{1}{c} \quad \text{si } x \in [0, c]$$

por lo tanto X tiene distribución uniforme en el intervalo $[0, c]$. ($X \sim \mathcal{U}[0, c]$).

DISTRIBUCIÓN EXPONENCIAL.

Sea Y_t una variable aleatoria de distribución de Poisson de parámetro λt, la cual mide el número de acontecimientos instantáneos que ocurren en un tiempo t.

Supongamos que observamos la aparición de dichos acontecimientos a partir de un instante cualquiera, cuánto tiempo tenemos que esperar para que ocurra el primer acontecimiento?. Es claro, que este tiempo de espera es una variable aleatoria X, tratemos de encontrar su distribución. El rango de X es $[0, \infty)$, por lo que $\mathbb{P}([X \leq x]) = 0$ si $x < 0$.

El evento $[X > x], x > 0$, significa que en el intervalo de tiempo $[0, x]$ aun no ha ocurrido ninguno de los acontecimientos en observación, por lo tanto si llamamos Y_x el número de acontecimientos en el intervalo $[0, x]$ se tiene que:

$$\mathbb{P}([X > x]) = \mathbb{P}([Y_x = 0]) = \frac{(\lambda x)^0}{0!} e^{-\lambda x} = e^{-\lambda x} \qquad \text{si } x > 0$$

Entonces

$$F(x) = \begin{cases} \mathbb{P}([X \leq x]) = 1 - e^{-\lambda x} & \text{si} \quad x \geq 0 \\ \mathbb{P}([X \leq x]) = 0 & \text{si} \quad x < 0. \end{cases}$$

Así la densidad de X es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si} \quad x \geq 0 \\ 0 & \text{si} \quad x < 0 \end{cases}$$

DEFINICIÓN:

Una variable aleatoria X tiene distribución exponencial de parámetro $\lambda > 0$ si su densidad es

$$f(x) = \begin{cases} \lambda e^{-\lambda} & \text{si} \quad x \geq 0 \\ 0 & \text{si} \quad x < 0 \end{cases}$$

OBSERVACIÓN: en algunos textos se define la densidad exponencial de parámetro $\mu > 0$ como

$$f(x) = \begin{cases} \frac{1}{\mu} e^{-\frac{x}{\mu}} & \text{si} \quad x \geq 0 \\ 0 & \text{si} \quad x < 0 \end{cases}$$

por lo que es conveniente especificar cuál de las dos definiciones se han de utilizar. Nosotros utilizaremos la primeras a menos que se indique lo contrario.

EJERCICIO:

Grafique la densidad y la función de distribución de la distribución exponencial de parámetro λ y haga un estudio comparativo para diferentes valores del parámetro.

PROPIEDADES:

1. Si X es una variable aleatoria de distribución exponencial de parámetro λ y $x_0 > 0$, entonces la distribución de la variable $Z = X - x_0$ condicionada por el suceso $[X > x_0]$ es de nuevo exponencial de parámetro λ.

   DEMOSTRACIÓN:

   Sea $A = [X > x_0]$
   $$F_A(Z \leq z \mid A)$$
   es la función de distribución de Z condicionada por el suceso A.
   $$F(x) = \mathbb{P}([X \leq x])$$
   es la función de distribución de X que sabemos, es exponencial de parámetro λ. Calculemos explícitamente $F_A(z)$:
   $$F_A(z) = \frac{\mathbb{P}([Z \leq z, X > x_0])}{\mathbb{P}([X > x_0])} = \frac{\mathbb{P}([X \leq z + x_0, X > x_0])}{\mathbb{P}([X > x_0])}$$
   si $z < 0$, el suceso $[X \leq z + x_0, X > x_0] = \emptyset$, de donde $F_A(z) = 0$ en este caso. Si $z \geq 0$
   $$F_A(z) = \frac{\mathbb{P}([x_0 < X \leq z + x_0])}{\mathbb{P}([X > x_0])} = \frac{F(z + x_0) - F(x_0)}{1 - F(x_0)} = \frac{e^{-\lambda x_0} - e^{-(z+x_0)}}{e^{-\lambda x_0}} = 1 - e^{-\lambda z}.$$
   De aquí se deduce que la distribución de Z condicionada por el suceso A es exponencial de parámetro λ.

2. Si X es una variable aleatoria continua, de densidad $f(x)$, $x_0 > 0$,con función de distribución F derivable, salvo a lo sumo, en una cantidad numerable de puntos y talque
   $$\mathbb{P}([X > x_0]) > 0$$
   si la distribución de la variable $Z = X - x_0$ condicionada por el suceso $A = [X > x_0]$ coincide con la distribución de X, entonces X tiene distribución exponencial.

   DEMOSTRACIÓN:

   Sea F_A la función de distribución de Z condicionada por el suceso A y F la función de distribución de X.

   Sea $G(z) = \mathbb{P}([Z > z \mid A]) = 1 - F_A(z) = 1 - F(z)$. Queremos demostrar que existe λ no nulo, tal que
   $$G(z) = \begin{cases} e^{-\lambda z} & \text{si } z \geq 0 \\ 1 & \text{si } z < 0 \end{cases}$$
   de este último resultado deduciríamos que F es la función de distribución exponencial.

   Supondremos que F es derivable salvo a lo sumo en una cantidad numerable de puntos.

   Si $z < 0$, el suceso$[Z > z, X > x_0] = [X > x_0]$, en este caso
   $$G(z) = 1$$
   Si $z \geq 0$,
   $$G(z) = \frac{\mathbb{P}([X > z + x_0])}{\mathbb{P}([X > x_0])} = \frac{1 - F(z + x_0)}{1 - F(x_0)} = \frac{G(z + x_0)}{G(x_0)},$$
   por lo tanto, si $x_0 > 0$, $G(z)$ verifica la siguiente igualdad
   $$G(z + x_0) = G(z)G(x_0)$$

para todo $z \geq 0$. Veamos que en este caso, existe λ tal que

$$G(z) = e^{-\lambda z}, \;\; z \geq 0.$$

G es continua en $\mathbb{R}$ ya que F lo es (la variable aleatoria X es continua) y es derivable salvo a lo sumo en una cantidad numerable de puntos, si $z = 0$

$$G(0) = \frac{G(x_0)}{G(x_0)} = 1$$

si $z > 0$ (es un punto de derivabilidad de G)

$$G'(z) = \lim_{h \to 0^+} \frac{G(z+h) - G(z)}{h} = \lim_{h \to 0^+} G(z)\frac{G(h) - 1}{h} = G(z)G'(0)$$

donde $G'(0)$ es la derivada por la derecha de la función en el punto 0. Sea k este valor. De estos cálculos concluímos que G satisface la siguiente ecuación diferencial:

$$y' - ky = 0$$

con la condición $y(0) = 1$. Cómo resolvemos esta ecuación? Vamos a seguir los siguientes pasos:

(a) multipliquemos ambos miembros de la ecuación por e^{-kz}, para obtener

$$e^{-kz}y' - ke^{-kz}y = 0$$

(b) observe que el término de la izquierda es

$$\frac{\partial}{\partial z}(e^{-kz}y)$$

así, la ecuación diferencial que queremos resolver se transforma en

$$\frac{\partial}{\partial z}(e^{-kz}y) = 0$$

es decir,

$$e^{-kz}y = c$$

donde c es una constante desconocida, por lo tanto la solución de la ecuación es de la forma

$$y = y(z) = ce^{-kz}$$

pero $y(0) = 1$, sustituyendo esta condición en la ecuación se obtiene un valor para la constante:

$$c = 1$$

es decir

$$G(z) = e^{kz}$$

por que G satisface la ecuación diferencial con la condición inicial dada. En este caso, $k = G'(0)$ (donde se entiende que esta derivada es por la derecha).

Ahora bien, $G(z) = 1 - F(z)$, luego

$$\lim_{z \to \infty} G(z) = 0$$

lo cual significa que $k < 0$.Sea $\lambda = -k > 0$, entonces

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0. \end{cases}$$

De aquí se deduce que X tiene distribución exponencial de parámetro $\lambda > 0$.

OBSERVACIÓN:

La primera propiedad de la distribución exponencial, citada anteriormente, se describe diciendo, que la distribución exponencial no tiene memoria.

Pensemos en un Proceso de Poisson que ocurre a través del tiempo, el instante X en el cual ocurre el primer acontecimiento, tiene distribución exponencial, por lo tanto, si al cabo de un tiempo x_0 no ha ocurrido aun el primer acontecimiento, es decir, $X > x_0$, entonces, no importa cuan grande sea x_0, se tiene que el tiempo $Z = X - x_0$, contado a partir de x_0 tiene la misma distribución de X.

Más aun, la segunda propiedad nos dice que la única distribución con esta propiedad es la exponencial.

DISTRIBUCIÓN NORMAL.

La distribución normal con parámetros μ, σ, es aquella que tiene como densidad

$$f(x) = f_N(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R}.$$

Dado que esta función es no negativa, para ver que es efectivamente una densidad, debemos demostrar que

$$\int_{-\infty}^{\infty} f_N(x;\mu,\sigma) = 1.$$

Para obtener este resultado vamos primeramente a hacer el siguiente cambio de variable:

$$z = \frac{x-\mu}{\sigma}$$

sustituyendo en la integral, obtenemos:

$$I = \int_{-\infty}^{\infty} f_N(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz, \qquad z \in \mathbb{R}.$$

Demostrar que $I = 1$ equivale a demostrar que $I^2 = 1$, ó equivalentemente que

$$\left(\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \right)^2 = 2\pi.$$

Observe que

$$\left(\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \right)^2 = \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \cdot \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{z^2+y^2}{2}} dz dy$$

(Todas estas igualdades se han escrito sin justificación, note que se trata de integrales impropias, continuaremos los cálculos sin precisar las justificaciones). Utilizaremos coordenadas polares para resolver la última integral cuya región de integración es todo el plano:

$$\begin{aligned} &y = r\cos\alpha \\ &z = r \ sen \ \alpha \\ &0 \le \alpha \le 2\pi, \ \ 0 \le r < \infty \end{aligned}$$

puesto que el Jacobiano de la transformación es $J = r$, obtenemos:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{z^2+y^2}{2}} dz dy = \int_{0}^{2\pi} \int_{0}^{\infty} r e^{-r^2} dr = 2\pi.$$

Denotaremos la distribución normal de parámetros μ, σ por

$$N(\mu, \sigma), \quad \mu \in \mathbb{R},\ \sigma > 0.$$

En el caso $\mu = 0$, $\sigma = 1$, la densidad $f_N(x; 0, 1)$ se conoce como la densidad normal estándar y se denota usualmente por ϕ, de modo que

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx, \qquad x \in \mathbb{R}.$$

La densidad normal estándar es simétrica respecto al origen. La función de distribución correspondiente a esta densidad se denota usualmente pro Φ :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

Esta función de distribución no puede ser calculada explícitamente ya que el integrando no tiene primitiva expresable como suma de funciones elementales y debe ser calculada por métodos numéricos. Existen tablas mediante las cuales obtenemos $\Phi(x)$ aproximadamente. Por ser $\phi(x) = \phi(-x)$ (simétrica respecto al origen) tenemos que si $x > 0$

$$\Phi(-x) = \int_{-\infty}^{-x} \phi(t)dt = \int_{x}^{\infty} \phi(t)dt = \int_{-\infty}^{\infty} \phi(t)dt - \int_{-\infty}^{x} \phi(t)dt = 1 - \Phi(x)$$

Esta fórmula nos permite obtener el valor de $\Phi(-x)$ a partir del valor de $\Phi(x)$, por lo tanto es suficiente tabular los valores de $\Phi(x)$ para $x \geq 0$.

En el caso de la distribución normal $N(\mu, \sigma)$, $\quad \mu \in \mathbb{R},\ \sigma > 0$, podemos valernos de las tablas correspondientes a la función de distribución normal estándar gracias al resultado que daremos a continuación.

PROPOSICIÓN:

Sea X una variable aleatoria de distribución $N(\mu, \sigma)$, $\quad \mu \in \mathbb{R},\ \sigma > 0$, entonces la variable aleatoria

$$Z = \frac{X - \mu}{\sigma}$$

tiene distribución normal estándar.

DEMOSTRACIÓN:

Sean $F(z)$ y $F_N(x; \mu, \sigma)$ las funciones de distribución de Z y de X respectivamente.

$$F(z) = \mathbb{P}([Z \leq z]) = \mathbb{P}\left(\left[\frac{X - \mu}{\sigma} \leq z\right]\right) = \mathbb{P}([X \leq \sigma z + \mu]) = F_N(\mu + \sigma z; \mu, \sigma).$$

La densidad de Z es $f(z) = F'(z)$,luego:

$$f(z) = \frac{\partial}{\partial z} F_N(\mu + \sigma z; \mu, \sigma) = F_N'(\mu + \sigma z; \mu, \sigma) \cdot \sigma =$$
$$\frac{\sigma}{\sqrt{2\pi}\sigma} e^{-\dfrac{(\mu + \sigma z - \mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\dfrac{z^2}{2}}, \quad z \in \mathbb{R}.$$

Es decir, Z tiene distribución $N(0, 1)$.

COROLARIO:

Si X es $N(\mu, \sigma)$, $\quad \mu \in \mathbb{R},\ \sigma > 0$, entonces

$$\mathbb{P}([a \leq X \leq b]) = \Phi(\frac{b - \mu}{\sigma}) - \Phi(\frac{a - \mu}{\sigma}), \ \ a < b,$$

donde Φ es la función de distribución normal estándar.

DEMOSTRACIÓN:

Sea $Z = \dfrac{X-\mu}{\sigma}$

$$\mathbb{P}\left([a \leq X \leq b]\right) = \mathbb{P}\left(\left[\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right]\right) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma}).$$

EJEMPLOS:

A) Sea X una variable aleatoria $N(\mu, \sigma)$ con $\mu = 5$, $\sigma = 2$; calcular la probabilidad que:

a) X sea positiva

b) X esté entre 4 y 6.

SOLUCIÓN:

a) $\mathbb{P}\left([X > 0]\right) = 1 - .\mathbb{P}\left([X \leq 0]\right) = 1 - \mathbb{P}\left(\left[Z \leq -\frac{5}{2}\right]\right) = 1 - \Phi(-\frac{5}{2})$, donde

$$Z = \frac{X-\mu}{\sigma}$$

$\Phi(-\frac{5}{2}) = 1 - \Phi(\frac{5}{2}) = 1 - 0,9938,$ así

$$\mathbb{P}\left([X > 0]\right) = 0,9938.$$

b)

$$\begin{array}{l} \mathbb{P}\left([4 \leq X \leq 6]\right) = \mathbb{P}\left(\left[-\frac{1}{2} \leq Z \leq \frac{1}{2} =\right]\right) = \\ \Phi(0,5) - \Phi(-0,5) = 2\Phi(0,5) - 1 = 2 \cdot (0,6915) - 1 = 0,3830. \end{array}$$

B) Sea X una variable aleatoria $N(0,1)$. Determine el valor de k para el cual

$$\mathbb{P}\left([X \geq k]\right) = 0,95.$$

SOLUCIÓN:

$$\mathbb{P}\left([X \geq k]\right) = 1 - \mathbb{P}\left([X < k]\right) = 0,95,$$

de donde (en este caso, $\mathbb{P}\left([X = k]\right) = 0$ y por simetría de la densidad):

$$\mathbb{P}\left([X < k]\right) = \mathbb{P}\left([X \leq k]\right) = 0,05 = 1 - \mathbb{P}\left([X \leq -k]\right)$$

porque $k < 0$. Sea $z_0 = -k$, hallemos z_0 tal que

$$\mathbb{P}\left([X \leq z_0]\right) = 1 - 0,05 = 0,95$$

el valor aproximado es $z_0 = 1,6448$. Así $k = -1,6448$.

C) Una fábrica de aluminio produce, entre otras cosas, cierto tipo de láminas de una aleación de aluminio. Por experiencia se sabe que la rigidez, medida en libras por pulgada cuadrada , está normalmente distribuída con $\mu = 2425$ y $\sigma = 115$.

Se desea calcular la probabilidad que una lámina de una aleación de aluminio seleccionada eleatoriamente de este proceso tenga un valor:

a) entre 2250 y 2425

b) menor de 2200.

SOLUCIÓN:

a)

$$\begin{aligned}&\mathbb{P}\left([2250 \leq X \leq 2425]\right) = \Phi(\tfrac{2425-2425}{115}) - \Phi(\tfrac{2250-2425}{115}) = \\ &\Phi(0) - \Phi(-1,52) = \Phi(0) - (1 - \Phi(1,52)) = \\ &0,5 - (1 - 0,9357) = 0,4357.\end{aligned}$$

b)

$$\begin{aligned}&\mathbb{P}\left([X \leq 2200]\right) = \Phi(\tfrac{2200-2425}{115}) = \\ &\Phi(-1,96) = 1 - \Phi(1,96) = 1 - 0,9750 = 0,25.\end{aligned}$$

D) Dada una variable aleatoria real de distribución normal con

$$\mu = 80, \ \ \sigma = 30.$$

Hallar

$$\mathbb{P}\left(|X - 100| > 25\right).$$

SOLUCIÓN: la inecuación

$$|X - 100| \leq 25$$

equivale a

$$100 - 25 \leq X \leq 100 + 25,$$

de aquí se deduce que

$$\begin{aligned}&\mathbb{P}\left(|X - 100| > 25\right) = 1 - \mathbb{P}\left(|X - 100| \leq 25\right) = \\ &1 - \mathbb{P}(75 \leq X \leq 125) = 1 - \mathbb{P}\left(-\tfrac{5}{30} \leq Z \leq \tfrac{45}{30}\right) = \\ &1 - \Phi(1,5) + \Phi(-0,17) = 2 - \Phi(1,5) - \Phi(0,17) = \\ &2 - 0,9332 - 0,5675 = 0,4993.\end{aligned}$$

DISTRIBUCIÓN GAMMA:

Una variable aleatoria X tiene distribución Gamma con parámetros

$$\lambda > 0, \ \ p > 0$$

si su densidad viene dada por

$$f_{\Gamma}(x; \lambda, p) = \frac{\lambda^p}{\Gamma(p)} e^{-\lambda x} x^{p-1} \ , \ \ x \geq 0$$

donde, $\Gamma(p)$ es la función Gamma definida por:

$$\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx, \ \text{si } p > 0.$$

OBSERVACIÓN: note que si $p = 1$, obtendremos $\Gamma(p) = 1$, la densidad de la distribución gamma corresponde a la distribución exponencial de parámetro λ. En algunos textos la densidad de la distribución gamma de parámetros

$$\mu > 0, \ \ p > 0$$

viene dada por

$$f_{\Gamma}(x; \mu, p) = \frac{1}{\mu^p \Gamma(p)} e^{-\frac{x}{\mu}} x^{p-1} \ , \ \ x \geq 0$$

DISTRIBUCIÓN CHI-CUADRADO

Si X tiene distribución Gamma de parámetros

$$\lambda = \frac{1}{2}, \ \ p = \frac{n}{2}$$

diremos que X tiene distribución Chi-Cuadrado con n grados de libertad. Esta distribución se denota por $\mathcal{X}^2_n$, su densidad es:

$$f_{\mathcal{X}^2_n}(x; \frac{1}{2}, \frac{n}{2}) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} e^{-\frac{1}{2}x} x^{\frac{n}{2}-1} \ , \ x \geq 0$$

DISTRIBUCIÓN BETA

Diremos que una variable aleatoria X tiene distribución Beta de parámetros

$$q > 0, \ \ p > 0$$

si su densidad viene dada por

$$f_B(x; p, q) = \frac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1}, \quad 0 \leq x \leq 1$$

donde, $\beta(p, q)$ es la función Beta, definida por:

$$\beta(p, q) = \int_0^1 t^{p-1}(1-t)^{q-1} dt.$$

DISTRIBUCIÓN t-DE STUDENT

Una variable aleatoria X tiene distribución $t-$ de Student con n grados de libertad, si su densidad viene dada por

$$f_S(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \ t \in \mathbb{R}.$$

ESPERANZA DE UNA VARIABLE ALEATORIA CONTINUA CON DENSIDAD.

DEFINICIÓN:

Si X es una variable aleatoria con densidad $f(x)$, definimos

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx \quad \text{si} \quad \int_{-\infty}^{\infty} |x| \, f(x) dx \ < \infty$$

en caso contario diremos que la esperanza de X no existe.

EJEMPLO:

Si X es una variable aleatoria con densidad $f(x) = \lambda e^{-\lambda x}$, si $x \geq 0$, entonces

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx \ = \lambda \int_0^{\infty} x e^{-\lambda x} dx = \frac{1}{\lambda}$$

(lo cual se obtiene fácilmente, por integración por partes).

OBSERVACIÓN:

Todas las propiedades enunciadas para $\mathbb{E}(X)$ en el caso discreto, siguen siendo válidas, la demostración se deja como ejercicio, ya que son consecuencia casi directa de las propiedades de la integral.

DEFINICIÓN:

Sea $g : \mathbb{R} \to \mathbb{R}$ una función tal que

$$Y = g(X),$$

sea una variable aleatoria, (si $g^{-1}(\mathcal{B}(\mathbb{R}))$ es un boreliano de $\mathbb{R}$, donde $\mathcal{B}(\mathbb{R})$ es la sigma álgebra de los borelianos de $\mathbb{R}$), donde X tiene densidad f, entonces

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

si esta integral es absolutamente convergente.(Este resultado se puede demostrar pero nos alejaríamos de los objetivos del curso, por esta razón es presentado como una definición que utilizaremos a continuación).

VARIANZA Y DESVIACIÓN ESTÁNDAR.

Si X tiene densidad f definimos

$$V(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 f(x)dx$$

y la desviación estándar es

$$D(X) = \sqrt{V(X)}$$

(Las propiedades dadas para $V(X)$ en el caso discreto, siguen siendo válidas, la demostración se deja como ejercicio).

TRANSFORMADA DE LAPLACE Ó EXPONENCIAL:

Si X es una variable aleatoria y en la definición de transformada geométrica hacemos $a = e^{-s}$, obtenemos

$$\mathbb{E}(a^X) = \mathbb{E}(e^{-sX})$$

La transformada geométrica es útil para estudiar las características de la distribución de una variable aleatoria discreta, en el caso continuo, cuando X tiene densidad f definimos:

$$L(s) = \mathbb{E}(e^{-sX}) = \int_{-\infty}^{\infty} e^{-sx} f(x)dx, \quad 0 \leq s < \infty$$

si esta integral converge (en particular converge absolutamente pues el integrando es positivo). Esta expresión se conoce con el nombre de transformada de Laplace de $f(x)$.

Si X toma sólo valores positivos ($R(X) \subseteq [0, \infty)$),

$$L(s) = \mathbb{E}(e^{-sX}) = \int_{0}^{\infty} e^{-sx} f(x)dx$$

siempre es convergente ya que $|e^{-sx}| \leq 1$ si $0 \leq s < \infty, 0 \leq x$.

La transformada de Laplace determina unívocamente la densidad f, por lo tanto determina la distribución de X (este hecho lo aceptamos sin demostración).

PROPIEDADES:

1. $L(0)=\int\limits_{-\infty}^{\infty} f(x)dx=1$

2. $L'(0)=-\int\limits_{-\infty}^{\infty} xf(x)dx=-\mathbb{E}(X)$

3. $L''(0)=\int\limits_{-\infty}^{\infty} x^2f(x)dx=\mathbb{E}(X^2)$

   de las dos últimas propiedades, se deduce que

$$V(X)=L''(0)-\left(L'(0)\right)^2.$$

4. $L^{(n)}(0)=(-1)^n\int\limits_{-\infty}^{\infty} x^nf(x)dx=(-1)^n\mathbb{E}(X^n)$.

OBSRVACIÓN:

Estas propiedades nos permiten calcular la esperanza y la varianza de la variable aleatoria X por simple derivación de la transformada de Laplace.

EJEMPLOS:

1. Sea X una variable aleatoria de distribución exponencial de parámetro λ

$$\begin{array}{l} L(s)=\mathbb{E}(e^{-sX})=\int\limits_0^{\infty} e^{-sx}f(x)dx=\lambda\int\limits_0^{\infty} e^{-sx}e^{-\lambda x}dx=\dfrac{\lambda}{s+\lambda} \\ L'(s)=-\dfrac{\lambda}{(s+\lambda)^2},\;\; L''(s)=\dfrac{2\lambda}{(s+\lambda)^3} \\ L'(0)=-\dfrac{1}{\lambda}=-\mathbb{E}(X)\;,\;\; L''(0)=\dfrac{2}{\lambda^2}=\mathbb{E}(X^2) \\ V(X)=\mathbb{E}(X^2)\;-(\mathbb{E}(X)\;)^2=\dfrac{2}{\lambda^2}-\dfrac{1}{\lambda^2}=\dfrac{1}{\lambda^2}. \end{array}$$

2. Sea X una variable aleatoria de distribución Uniforme en $[0,1]$,

$$L(s)=\int_0^1 e^{-sx}dx=\frac{1-e^{-s}}{s},\;\text{ si } s \text{ es no nulo.}$$

   Puesto que $L(s)$ no está definida en $s=0$ pero

$$\lim_{s\to 0^+} L(s)=1,$$

   podemos definir

$$L(s)=\begin{cases} \dfrac{1-e^{-s}}{s} & \text{si } s \text{ es no nulo} \\ 1 & \text{si } s=0 \end{cases}$$

   Podemos hallar directamente $L'(s)$ si s es no nulo por simple derivación, pero en $s=0$, debemos hallar la derivada por definición:

$$\begin{array}{l} L'(s)=\lim\limits_{s\to 0^+}\dfrac{L(s)-L(0)}{s}=\lim\limits_{s\to 0^+}\dfrac{\dfrac{1-e^{-s}}{s}-1}{s}= \\ \lim\limits_{s\to 0^+}\dfrac{1-e^{-s}-s}{s^2}=\lim\limits_{s\to 0^+}\dfrac{e^{-s}-1}{2s}=-\dfrac{1}{2}=-\mathbb{E}(X) \end{array}.$$

Por simple derivación, obtenemos que si s es no nulo

$$L'(s) = \frac{se^{-s} + e^{-s} - 1}{s^2}$$

usando la definición de derivada, podemos hallar a partir de esta expresión, la segunda derivada de L evaluada en $s = 0$

$$L''(s) = \lim_{s \to 0^+} \frac{L'(s) - L'(0)}{s} = \frac{4}{12} = \frac{1}{3} = \mathbb{E}(X^2)$$

de aquí se obtiene que:

$$V(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Función Característica .

Si en la definición de transformada de Laplace hacemos la sustitución $s = -it$ con $i = \sqrt{-1}$, obtenemos la función característica de X :

$$\varphi(t) = \mathbb{E}(e^{itX})$$

donde $e^{itx} = \cos tx + i\, sen\, tx$ para todo $t \in \mathbb{R}$.

Abusando del lenguaje, en lo que sigue llamaremos indistintamente a la función característica de X, Transformada de Fourier de X. Aunque en general la transformada de Fourier es $\varphi(-2\pi t) = \mathbb{E}(e^{-i2\pi tX})$, (es fácil pasar de una a otra cambiando t por $-2\pi t$), sin embargo algunos autores la definen como la función característica.

La ventaja de utilizar esta transformada es que siempre existe para toda variable aleatoria X ya que

$$\left|e^{itx}\right|^2 = \cos^2 tx + \, sen^2\, tx = 1, \mathbb{E}(\left|e^{itX}\right|) = 1 \quad \text{para todo } t.$$

La función característica caracteriza la distribción de X (aceptaremos este hecho sin demostración).

Si X tiene densidad $f(x)$

$$\varphi(t) = \mathbb{E}(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx$$

PROPIEDADES.

1. $\varphi(0) = 1$
2. $\varphi'(0) = i\mathbb{E}(X)$
3. $\varphi''(0) = i^2\mathbb{E}(X^2) = -\mathbb{E}(X^2)$.

EJERCICIOS PARA LA PRÁCTICA

1. Sea X una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} \dfrac{x-a+\delta}{\delta^2}, & \text{si } a-\delta \le x < a \\ \dfrac{a+\delta-x}{\delta^2}, & \text{si } a \le x \le a+\delta \\ 0 & \text{en otro caso,} \end{cases}$$

con $\delta > 0$.

   (a) Hallar la función de distribución F de X. Grafíquela.

   (b) Grafique f. (Esta distribución se denomina Distribución Triangular).

   (c) Hallar $\mathbb{P}(a - \frac{\delta}{2} \le X \le a + \frac{\delta}{2})$

2. Sea X una variable aleatoria con distribución uniforme en $[0,1]$ .Halle la densidad de la variable aleatoria

$$Y = -\frac{1}{\lambda}\ln(1-X), \quad \lambda > 0.$$

(Halle la función de distribución de Y y apartir de ella, halle la densidad de Y).

3. Sea X una variable aleatoria con distribución uniforme en $[0,1]$. Calcule:

   (a) $\mathbb{P}(X^2 \le x)$

   (b) Si $Y = X^2$,cuál es la densidad de Y?

4. Sea X una variable aleatoria con distribución exponencial de parámetro λ.

   (a) Calcule $\mathbb{P}(\dfrac{1}{X} \le x)$, para todo $x \in \mathbb{R}$.

   (b) Sea $Y = \dfrac{1}{X}$. Cuál es la densidad de Y?

5. Sea X una variable aleatoria de distribución Normal de parámetros $\mu = 100,\ \sigma = 20.$ Calcule:

   (a) $\mathbb{P}(X \le 97,3)$

   (b) $\mathbb{P}(X \ge 110)$

   (c) $\mathbb{P}(112,1 \le X \le 115,8)$

   (d) $\mathbb{P}([X \le 87,3] \cup [X \ge 108,5])$

6. Sea X una variable aleatoria de distribución Normal de parámetros $\mu = 12,\ \sigma = 3.$ Halle :

   (a) $\mathbb{P}(X > 3)$

   (b) $\mathbb{P}(|X-12| < 4)$

   (c) $\mathbb{P}(|X-10| > 2)$

7. Sea X una variable aleatoria de distribución Normal de parámetros $\mu = 0,\ \sigma = 1, Y = X^2$. Hallar la densidad de Y. (La variable aleatoria $Y,$ tendrá distribución Chi-cuadrado con $n = 1$ grado de libertad, pues $\sqrt{\pi} = \Gamma(1/2),$ ver el siguiente ejercicio). Explícitamente, la densidad de Y es

$$f_Y(y; \frac{1}{2}, \frac{1}{2}) = \frac{\left(\frac{1}{2}\right)^{\frac{1}{2}}}{\Gamma(\frac{1}{2})} e^{-\frac{1}{2}y} y^{\frac{1}{2}-1} = \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-\frac{1}{2}y},\ y \ge 0$$

8. Sea $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$, si $p > 0$. (Función Gamma). Demuestre que:

   (a) $\Gamma(p) = (p-1)\Gamma(p-1)$, si $p > 1$. (Use integración por partes).

   (b) En particular, si p es un entero $p \geq 1$, $\Gamma(p) = (p-1)!$

   (c) $\int_0^\infty \lambda^p x^{p-1} e^{-\lambda x} dx = \Gamma(p)$.

   (d) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.(Haga el cambio de variable $x = \dfrac{t^2}{2}$).

   (e) $\int_0^\infty x^b e^{-ax} dx = \dfrac{\Gamma(b+1)}{a^{b+1}}$

9. Sea
$$f_B(x; p, q) = \frac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1}, \quad 0 \leq x \leq 1,\, p > 0, q > 0,$$
   donde, $\beta(p, q)$ es la función Beta, definida por:
$$\beta(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1} dt.$$
   Demuestre que:

   (a) $f_B(x; p, q)$ es una densidad.

   (b) $\beta(p, q) = \beta(q, p)$

   (c) $\beta(p, q) = \dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$. (Exprese $\Gamma(p)\Gamma(q)$ como una integral doble y haga el cambio de variable $z = x + y$).

10. Sea
$$f_S(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})} \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}},\ t \in \mathbb{R}$$

   (a) Grafique $f_S(t)$.

   (b) Demuestre que
$$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})} = \frac{1}{\sqrt{n}\,\beta(\frac{1}{2}, \frac{n}{2})}.$$

   (c) Demuestre que $f_S(t)$ es una densidad.
   Ayuda: para demostrar que
$$\int_{-\infty}^{\infty} \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}} dt = \frac{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})}$$
   utilice que $\dfrac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}} = \dfrac{\int_0^\infty x^{\frac{n+1}{2}-1} e^{-(1+\frac{t^2}{n})x} dx}{\Gamma(\frac{n+1}{2})}$.

(d) Si X tiene densidad $f_S(t)$, demuestre que

$$\mathbb{P}(X \leq -x) = \mathbb{P}(X \geq x), \text{si } x \geq 0.$$

11. Sea Y una variable aleatoria de función de distribución

$$F(y) = \begin{cases} 0, & \text{si } y \leq 0 \\ \dfrac{y}{8}, & \text{si } 0 < y < 2 \\ \dfrac{y^2}{16}, & \text{si } 2 \leq y < 4 \\ 1 & \text{si } y \geq 4. \end{cases}$$

(a) Grafique F

(b) Encuentre la función de densidad f de F y grafíquela.

(c) Calcule $\mathbb{P}(Y \geq 1,5)$

(d) $\mathbb{P}(Y \geq 1 \mid Y \leq 3)$.

12. El tiempo de vida de cierto tipo de una lámpara es una variable aleatoria Y cuya distribución viene dada por

$$F(y) = \begin{cases} 0 & \text{si } y < 0 \\ 1 - e^{-y^2} & \text{si } y \geq 0, \end{cases}$$

(el tiempo se mide en cientos de hora, es decir que la lámpara dure 100 horas significa que $Y = 1$).

(a) Demuestre que F tienen las propiedades de una función de distribución.

(b) Calcule la densidad f de F.

(c) Calcule la probabilidad que una lámpara de este tipo dure al menos 200 horas.

13. Se selecciona al azar un punto del círculo de radio $R = 1$ y de centro 0. El espacio muestral es

$$\Omega = \left\{(\omega_1, \omega_2) : \omega_1^2 + \omega_2^2 \leq 1\right\}$$

Sea $X(\omega_1, \omega_2) = \omega_1$. Calcule

(a) $\mathbb{P}(-\frac{1}{2} \leq X \leq \frac{1}{2})$

(b) $\mathbb{P}(\frac{2}{3} \leq X)$.

Calcular la transformada de Lapalace (si existe) de las densidades:

(a) Uniforme en $[a, b]$.

(b) Normal (μ, σ).

(c) Gamma (λ, p).

(d) Beta (p, q).

(e) Chi-Cuadrado con n grados de libertad.

(f) T de Student con n grados de libertad.
y deducir la esperanza y la varianza de cada una de ellas. (Trate de calcular para cada una, la esperanza y la varianza, sin utilizar transformada de Laplace).

14. Sea X una variable aleatoria de densidad

$$f(x) = \begin{cases} x, & x \in [0, 1] \\ 2 - x, & x \in [1, 2] \\ 0 & \text{en otro caso.} \end{cases}$$

(a) Halle la función de distribución de X.

(b) Halle la esperanza y la varianza de X.

(c) Sean $A = \left[\frac{1}{2} \leq X \leq \frac{3}{2}\right]$, $B = [X > 1]$, son A y B eventos independientes?

15. Sea X una variable aleatoria discreta de transformada geométrica

$$T(a) = \mathbb{E}(a^X) = e^{a-1}.$$

Halle su función de probabilidad, su esperanza y su varianza.

16. Sea X una variable aleatoria de densidad

$$f(x) = \begin{cases} 3x^2, & \text{si } -1 \leq x \leq 0 \\ 0 & \text{en otro caso.} \end{cases}$$

Sea $b \in (-1, 0)$. Calcule

$$\mathbb{P}(X > b \mid X < \frac{b}{2}).$$

17. Sea X una variable aleatoria de densidad

$$f(x) = \begin{cases} 6x(1-x), & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otro caso.} \end{cases}$$

(a) Verifique que f es una densidad y grafíquela.

(b) Calcule y grafique su función de distribución.

(c) Determine b tal que

$$\mathbb{P}(X > b) = \mathbb{P}(X < b)$$

(d) Calcule

$$\mathbb{P}(X \leq \frac{1}{2}), \mathbb{P}(\frac{1}{3} < X < \frac{2}{3}).$$

18. Se lanza 4 veces un dado homogéneo, sea

$$Y = \text{“n}^o \text{ de veces que apareció la cara 6”}$$

Encuentre y grafique la función de probabilidad de Y, la función de distribución de Y, la esperanza y la varianza de Y.

19. Un 10% de los utensilios producidos en un cierto proceso de fabricación resulta ser defectuoso. Hallar la probabilidad de que en una muestra de 10 utensilios elegidos al azar encontremos 2 defectuosos, mediante

(a) la distribución Binomial

(b) la aproximación de Poisson a la distribución Binomial.

(c) hallar el promedio de artículos defectuosos.

20. Si la probabilidad de que un individuo sufra una reacción por una inyección de un determinado suero es $0,001$, determine la probabilidad que en un total de 2000 individuos

(a) exactamente 2 tengan reacción

(b) más de 2 tengan reacción

(c) el promedio del número de individuos que tienen reacción.

21. Entre las 2 y las 4 de la tarde, el promedio de llamadas telefónicas que recibe la central de una compañía por minuto es $2,5$. Hallar la probabilidad de que en un determinado minuto haya:

(a) 0 llamadas

(b) 1 llamada

(c) 4 ó menos llamadas

(d) más de 6 llamadas.

22. Una panadería horneó 100 panes de pasas, utilizando un total de 2000 pasas. Si Ud. compra uno de esos panes, cuál es la probabilidad de que su pan tenga 20 pasas, 30 pasas, entre 20 y 30 pasas?

23. Una pareja joven planea tener hijos hasta que nazca el primer varón. Suponga que la probabilidad que un niño nazca varón es $\frac{1}{2}$ y que el resultado de cada nacimiento es independiente de los anteriores, cuál es el valor esperado del número de hijos que tendrán?

24. La media de los pesos de un grupo de 500 estudiantes universitarios es de $68Kg$, con una desviación estándar de $6,8Kg$. Suponiendo que los pesos se distribuyen normalmente, hallar cuántos estudiantes pesan

    (a) entre $54,5$ y $70,5Kg$

    (b) más de $84Kg$

    (c) menos de $58Kg$.

25. El gasto semanal en mantenimiento en una empresa se ha observado por un largo período de tiempo y se ha notado que se modela adecuadamente por una distribución normal de media Bs.80000 y desviación Bs.4000.

    Cuánto se debería presupuestar para mantenimiento semanalmente, de modo que la cantidad presupuestada sea excedida con una probabilidad de tan sólo $0,1$?

26. Los registros de una compañía constructora de pozos indican que la probabilidad que uno de sus pozos nuevos requiera de reparaciones en el término de un año es de 0.20. Cuál es la probabilidad que el sexto pozo construido por esta compañía en un año dado sea el primero en requerir reparaciones en un año?

Capítulo 6

VARIABLES CONJUNTAS.

VARIABLES CONJUNTAS.

Recordemos, que una variable aleatoria es, a groso modo, una función que asocia valores numéricos a los puntos de un espacio muestral, podemos definir dos ó más variables aleatorias.

Por ejemplo, si lanzamos un dado dos veces, podemos definir

$$\begin{array}{ll} X(\omega) = i + j & \text{si } \omega = (i,j) \in \Omega \\ Y(\omega) = \max(i,j) & \text{si } \omega = (i,j) \in \Omega \end{array}$$

donde

$$\Omega = \{(i,j) : 1 \leq i,j \leq 6\}$$

es el espacio muestral.

Ya hemos estudiado la distribución de ambas variables aleatorias, ahora nos interesa la probabilidad de sucesos en donde intervengan ambas variables simultáneamente. Por ejemplo sucesos como:

$$\begin{array}{c} [X \geq Y + 2] \\ [X + Y \leq 6] \end{array}$$

El conocimiento de las distribuciones individuales (ó marginales) de X e Y no es suficiente para calcular estas probabilidades, para ello necesitamos conocer la distribución conjunta de las variables X e Y.

Si X e Y son dos variables aleatorias definidas sobre un mismo espacio muestral Ω, el par $(X(\omega), Y(\omega))$ es un punto en $\mathbb{R}^2$. Es decir, a cada resultado del experimento le asociamos el par $(x,y) \in \mathbb{R}^2$ donde $(x,y) = (X(\omega), Y(\omega))$.

Si $A \subset \mathbb{R}^2$, el conjunto

$$[(X,Y) \in A] = \{\omega : (X(\omega), Y(\omega)) \in A\}$$

es un subconjunto del espacio muestral, si queremos calcular la probabilidad de este subconjunto es necesario que sea un evento (es decir que pertenezca a la $\sigma - \acute{a}lgebra$ de subconjuntos del espacio muestral).

Entre todos los subconjuntos de $\mathbb{R}^2$, restringiremos nuestra atención a aquellos que se puedan obtener mediante uniones numerables, interseccciones numerables, complementario de rectángulos de $\mathbb{R}^2$. Tales conjuntos forman una familia muy extensa de subconjuntos de $\mathbb{R}^2$, reciben el nombre de Borelianos de $\mathbb{R}^2$. Denotaremos esta clase de subconjuntos por $\mathcal{B}(\mathbb{R}^2)$.

Los conjuntos de $\mathbb{R}^2$ que no son borelianos, igual que en el caso unidimensional, son conjuntos patológicos y en la práctica no nos tropezaremos con ellos.

DEFINICION:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidades, X e Y dos variables aleatorias reales definidas sobre Ω. Diremos que X e Y son variables aleatorias conjuntas (ó que (X,Y) es un vector aleatorio bidimensional, si para todo subconjunto A, boreliano de $\mathbb{R}^2$, el conjunto

$$[(X,Y) \in A] \in \mathcal{F}$$

(es decir, es un evento del espacio muestral). Es suficiente pedir en esta definición que el conjunto A, sea un rectángulo cualquiera de $\mathbb{R}^2$.

RANGO CONJUNTO:

Si X e Y son variables aleatorias conjuntas, su rango conjunto se define como el conjunto:

$$R(X,Y) = \left\{(x,y) \in \mathbb{R}^2 : \text{existe } \omega \in \Omega : X(\omega) = x, Y(\omega) = y\right\}.$$

DISTRIBUCIÓN CONJUNTA:

Si X e Y son variables aleatorias conjuntas, su distribución conjunta es una probabilidad definida sobre la $\sigma-$álgebra $\mathcal{B}(\mathbb{R}^2)$ por:

$$\mathbb{P}_{XY}(A) = \mathbb{P}\left((X,Y) \in A\right)$$

donde $A \in \mathcal{B}(\mathbb{R}^2)$.

EJEMPLO:

Se lanzan dos dados distinguibles, podemos definir

$$\begin{array}{ll} X(\omega) = i + j & \text{si } \omega = (i,j) \in \Omega \\ Y(\omega) = \max(i,j) & \text{si } \omega = (i,j) \in \Omega \end{array}$$

donde

$$\Omega = \{(i,j) : 1 \leq i, j \leq 6\}$$

es el espacio muestral.

Calculemos la probabilidad del evento

$$[X \geq Y + 2]$$

Observemos que

$$[X \geq Y + 2] = \bigcup_{k=1}^{6} [X \geq k + 2, Y = k]$$

Necesitamos calcular las probabilidades de los eventos que aparecen a la derecha de la última expresión, para cada uno de los valores de k.

$$\begin{array}{l} CARD\,[X \geq 3, Y = 1] = 0 \\ CARD\,[X \geq 4, Y = 2] = 1 \\ CARD\,[X \geq 5, Y = 3] = 3 \\ CARD\,[X \geq 6, Y = 4] = 5 \\ CARD\,[X \geq 7, Y = 5] = 7 \\ CARD\,[X \geq 8, Y = 6] = 9 \end{array}$$

El cardinal del espacio muestral es

$$CARD(\Omega) = 6^2$$

por lo tanto:

$$\mathbb{P}\left([X \geq Y + 2]\right) = \frac{1}{36} \sum_{k=1}^{6} Card([X \geq k + 2, Y = k] = \frac{25}{36}$$

FUNCIÓN DE DISTRIBUCIÓN CONJUNTA.

Si X e Y son variables aleatorias conjuntas, consideremos el evento

$$[X \leq x, Y \leq y] = \{\omega : X(\omega) \leq x\} \cap \{\omega : Y(\omega) \leq x\}$$

definido para todo $(x, y) \in \mathbb{R}^2$.

La probabilidad de este evento es una función de dos variables. Definimos la función de distribución conjunta de las variables aleatorias Xe Y, como:

$$F_{XY}(x, y) = \mathbb{P}([X \leq x, Y \leq y])$$

y cuando no haya confusión posible escribiremos simplemente $F(x, y)$.

PROPIEDADES DE LA FUNCIÓN DE DISTRIBUCIÓN.

1. $F(x, y)$ es monótona creciente en cada una de las variables. Esto es consecuencia inmediata de la siguiente relación

$$[X \leq a_1, Y \leq a_2] \subset [X \leq b_1, Y \leq b_2]$$

si $a_1 \leq b_1, a_2 \leq b_2$.

2. $F(x, y)$ es continua por la derecha tanto en x como en y, es decir

$$\lim_{x \to a^+} F(x, b) = F(a, b), a \in \mathbb{R}, b \in \mathbb{R}$$
$$\lim_{y \to b^+} F(a, y) = F(a, b), a \in \mathbb{R}, b \in \mathbb{R}$$

En efecto, basta demostrar que si (x_n) es una sucesión decreciente tal que

$$x_n \geq a, \lim_{n \to \infty} x_n = a$$

entonces

$$\lim_{n \to \infty} F(x_n, b) = \lim_{n \to \infty} \mathbb{P}([X \leq x_n, Y \leq b]) = F(a, b) = \mathbb{P}([X \leq a, Y \leq b])$$

Observe que

$$\lim_{n \to \infty} (\mathbb{P}([X \leq x_n, Y \leq b]) - \mathbb{P}([X \leq a, Y \leq b])) =$$
$$\lim_{n \to \infty} \mathbb{P}([a < X \leq x_n, Y \leq b])$$

los eventos $[a < X \leq x_n, Y \leq b]$ decrecen al conjunto vacío, por lo tanto se tiene el resultado.

La continuidad por la derecha en la variable y se obtiene de manera análoga.

3.

$$\lim_{x \to -\infty} F(x, y) = 0, \quad y \in \mathbb{R}$$
$$\lim_{y \to -\infty} F(x, y) = 0, \quad x \in \mathbb{R}$$

En efecto, si (x_n) decrece y tiende a $-\infty$, la sucesión de eventos

$$[X \leq x_n, Y \leq y]$$

decrece y la intersección de ellos es vacía, de dónde se obtiene el primer resultado. El segundo es análogo.

4.

$$\lim_{x \to \infty} F(x, y) = \mathbb{P}([Y \leq y]) = F_Y(y), \quad y \in \mathbb{R}$$
$$\lim_{y \to \infty} F(x, y) = \mathbb{P}([X \leq x]) = F_X(x), \quad x \in \mathbb{R}$$
$$\lim_{(x,y) \to (\infty, \infty)} F(x, y) = 1$$

En efecto

$$\lim_{x \to \infty} F(x, y) = \lim_{x \to \infty} \mathbb{P}([X \leq x, Y \leq y]) = \mathbb{P}([X < \infty, Y \leq y]) = \mathbb{P}([Y \leq y]).$$

Análogamente, obtenemos el segundo resultado. La última afirmación es equivalente a

$$\lim_{(x,y) \to (\infty, \infty)} F(x, y) = \mathbb{P}([X < \infty, Y < \infty]) = \mathbb{P}(\Omega) = 1.$$

PROPOSICIÓN.

Si $a_1 < b_1$, $a_2 < b_2$ entonces

$$\mathbb{P}\left([a_1 < X \leq b_1, a_2 < Y \leq b_2]\right) = F(b_1, b_2) + F(a_1, a_2) - F(a_1, b_2) - F(b_1, a_2)$$

Sean

$$\begin{array}{l} A = [a_1 < X \leq b_1, a_2 < Y \leq b_2] \\ B = [X \leq b_1, Y \leq b_2] \\ C = [X \leq a_1, Y \leq b_2] \\ D = [X \leq b_1, Y \leq a_2] \end{array}$$

El evento B puede escribirse como unión de los otros tres eventos, de aquí se obtiene el resultado. (Note que esta unión no es disjunta).

IMPORTANCIA DE LA FUNCIÓN DE DISTRIBUCIÓN.

La distribución conjunta de dos variables aleatorias, queda determinada por su función de distribución conjunta, puesto que cualquier rectángulo del plano puede expresarse en función de F usando argumentos parecidos al caso unidimensional.

EJEMPLO:

Sea $a \in \mathbb{R}$, $b < c$, el rectángulo degenerado

$$[X = a, b < Y \leq c]$$

se puede obtener intersectando la sucesión decreciente de rectángulos

$$[x_n < X \leq a, b < Y \leq c]$$

donde (x_n) es una sucesión de números reales que crece hacia el punto a. Así:

$$\begin{array}{l} \mathbb{P}\left([X = a, b < Y \leq c]\right) = \lim\limits_{x \to a^-} \mathbb{P}\left([x < X \leq a, b < Y \leq c]\right) = \\ \lim\limits_{n \to \infty} \mathbb{P}\left([x_n < X \leq a, b < Y \leq c]\right) = \\ \lim\limits_{x \to a^-} \left(F(a,c) - F(x,c)\right) + \lim\limits_{x \to a^-} \left(F(x,b) - F(a,b)\right) \end{array}$$

Note que si $F(x,y)$ es continua por la izquierda en $x = a$, esta probabilidad es cero.

EJEMPLO: Sean X e Y dos variables aleatorias cuya función de distribución conjunta es:

$$F(x,y) = \begin{cases} 1 - \frac{1}{x+1} - \frac{1}{y+1} + \frac{1}{1+x+y}, & \text{si } x, y \geq 0 \\ 0 & \text{en otro caso.} \end{cases}$$

Calculemos la probabilidad del evento

$$A = [0 \leq X \leq 1, 1 \leq Y \leq 2]$$

El evento A se puede expresar como la unión disjunta de los siguientes eventos:

$$\begin{array}{l} B = [0 < X \leq 1, 1 < Y \leq 2] \\ C = [0 < X \leq 1, Y = 1] \\ D = [X = 0, 1 < Y \leq 2] \\ E = [X = 0, Y = 1] \end{array}$$

Puesto que $F(x,y)$ es continua en cada una de las variables, se tiene que los eventos: C, D, E tienen probabilidad nula. Por lo tanto

$$\mathbb{P}\left(A\right) = \mathbb{P}\left([0 < X \leq 1, 1 < Y \leq 2]\right) = F(1,2) + F(0,1) - F(0,2) - F(1,1) = \frac{1}{12}.$$

Otras caracterizciones de la distribución conjunta.

Sean X e Y dos variables aleatorias conjuntas

1. **CASO DISCRETO PURO:**

   Si ambas variables son discretas, el rango conjunto

   $$R(X,Y) \subset R(X) \times R(Y)$$

   es finito ó numerable. Para caracterizar la distribución conjunta de X e Y, basta conocer la probabilidad de los eventos $[X = m, Y = n]$ con $(m,n) \in R(X,Y)$. La función

   $$g(m,n) = \mathbb{P}\left([X = m, Y = n]\right), \quad \text{si } (m,n) \in R(X;Y)$$

   se llama función de probabilidad conjunta de las variables X e Y.

   Si A es un boreliano de $\mathbb{R}^2$, se tiene que

   $$\mathbb{P}\left([(X,Y) \in A]\right) = \sum_{(m,n)\in A\cap R(X,Y)} g(m,n).$$

   En particular, la función de distribución conjunta queda unívocamente determinada a partir de la función de probabilidad conjunta, mediante la expresión:

   $$F(x,y) = \sum_{m\le x, n\le y} g(m,n).$$

   PROPIEDADES

   (a) $0 \le g(m,n) \le 1$

   (b)

   $$\sum_{(m,n)\in R(X,Y)} g(m,n) = \sum_{m\in R(X)} \sum_{n\in R(Y)} \mathbb{P}\left([X = m, Y = n]\right) = \sum_{m\in R(x)} \mathbb{P}\left([X = m]\right) = 1$$

2. **CASO CONTINUO PURO**

   Si X e Y son dos variables aleatorias conjuntas, con función de distribución conjunta $F(x,y)$ continua en cada una de las variables, con segundas derivadas parciales mixtas iguales en todos los puntos (x,y), salvo a lo sumo en un número finito ó numerable de puntos, definimos la densidad conjunta de X e Y como:

   $$f(x,y) = \frac{\partial^2 F(x,y)}{\partial y \partial x} = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

   Si se conoce la densidad conjunta, la función de distribución conjunta se determina a partir de ella mediante la expresión:

   $$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(s,t)\,ds\,dt$$

   Más aun, si A es un boreliano de $\mathbb{R}^2$:

   $$\mathbb{P}\left((X,Y) \in A\right) = \iint_A f(x,y)\,dx\,dy$$

   Note que si A tiene área nula, entonces

   $$\mathbb{P}\left((X,Y) \in A\right) = 0,$$

en particular, $\mathbb{P}(X = Y) = 0$.

OBSERVACIÓN: Si $F(x,y)$ es la función de distribución conjunta de las variables aleatorias X e Y, diremos que esta función tiene densidad si existe $f \geq 0$, tal que

$$F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(s,t)dsdt$$

para todo punto (x,y) del plano.

Si $f(s,t)$ es continua en (s,t) y existe la derivada cruzada

$$\frac{\partial^2 F(s,t)}{\partial x \partial y}$$

entonces

$$f(s,t) = \frac{\partial^2 F(s,t)}{\partial x \partial y}.$$

En este curso consideraremos únicamente el caso en que F tiene densidad y existe esta última relación entre la densidad y la función de distribución conjuntas.

PROPIEDADES:

(a) $f(x,y) \geq 0$

(b) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t)dsdt = 1$

(c) La densidad conjunta multiplicada por un elemento infinitesimal de área, $(f(x,y)\Delta x \Delta y)$ representa la probabilidad del rectángulo

$$[x < X \leq x + \Delta x, y < Y \leq y + \Delta y],$$

ya que

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial y \partial x} = \lim_{\Delta y \to 0} \frac{\frac{\partial F}{\partial x}(x,y+\Delta y) - \frac{\partial F}{\partial x}(x,y)}{\Delta y} =$$
$$\lim_{\Delta y \to 0}\lim_{\Delta x \to 0}\left[\frac{F(x+\Delta x, y+\Delta y) - F(x, y+\Delta y)}{\Delta x \Delta y} - \frac{F(x+\Delta x, y) - F(x,y)}{\Delta x \Delta y}\right] =$$
$$\lim_{\Delta y \to 0}\lim_{\Delta x \to 0} \frac{\mathbb{P}([x < X \leq x+\Delta x, y < Y \leq y+\Delta y])}{\Delta x \Delta y}.$$

3. **CASO MIXTO**

Si una de las dos variables es continua y la otra es discreta, por ejemplo, X continua e Y discreta, la distribución conjunta queda determinada por una función $h(x,n)$ con una variable continua y la otra discreta, donde

$$h(x,n) = \lim_{\Delta x \to 0} \frac{\mathbb{P}([x < X \leq x + \Delta x, Y = n])}{\Delta x}$$

La función de distribución conjunta se obtiene mediante $h(x,n)$, a partir de la expresión:

$$F(x,y) = \sum_{n \leq y} \int_{-\infty}^{x} h(s,n)ds.$$

En general, si A es un boreliano de $\mathbb{R}^2$

$$\mathbb{P}([(X,Y) \in A]) = \sum_{(x,n) \in A} \int h(x,n)dx$$

La función $h(x,n)$ se llama densidad-función de probabilidad.

PROPIEDADES:

(a) $0 \leq h(x,n) \leq 1$

(b) $\sum_{n \in R(Y)} \int_{-\infty}^{\infty} h(x,n)dx = 1$

EJEMPLOS:

1. Una caja contiene 12 bolas de las cuales 5 son blancas, 4 negras y 3 rojas. Se extraen al azar cuatro bolas, sin reemplazo. Sea X el número de bolas blancas extraídas e Y el número de bolas negras .

   (a) Hallar la función de probabilidad de cada variable por separado (función de probabilidad marginal).

   (b) Halle la función de probabilidad conjunta.

   (c) Halle la probabilidad del evento $[X = Y]$.

   SOLUCIÓN:

   $CARD(\Omega) = \binom{12}{4}$. Las funciones de distribución marginales de X viene dada por:

$$\mathbb{P}([X = m]) = \frac{\binom{5}{m}\binom{7}{4-m}}{\binom{12}{4}}, \ \ m = 0, 1, 2, 3, 4$$

   y la de Y, es

$$\mathbb{P}([Y = n]) = \frac{\binom{4}{n}\binom{8}{4-n}}{\binom{12}{4}}, \quad n = 0, 1, 2, 3, 4.$$

   El rango conjunto viene dado por:

$$R(X, Y) = \{(m, n) : m = 0, 1, 2, 3, 4; n = 0, 1, 2, 3, 4; \ 1 \leq m + n \leq 4\}$$

   La función de probabilidad conjunta viene dada por:

$$\mathbb{P}([X = m, Y = n]) = \frac{\binom{5}{m}\binom{4}{n}\binom{3}{4-m-n}}{\binom{12}{4}}, (m, n) \in R(X, Y).$$

   Finalmente

$$\mathbb{P}([X = Y]) = \sum_{k=0}^{4} \mathbb{P}([X = k, Y = k]) = \sum_{k=1}^{2} \mathbb{P}([X = k, Y = k]) =$$
$$\frac{\binom{5}{1}\binom{4}{1}\binom{3}{2}}{\binom{12}{4}} + \frac{\binom{5}{2}\binom{4}{2}\binom{3}{0}}{\binom{12}{4}} = \frac{8}{33}.$$

2. Sea (X, Y) un vector aleatorio bidimensional de densidad conjunta

$$f(x, y) = \begin{cases} 2\sqrt{2}, & \text{si } 0 \leq y\sqrt{2} \leq 1 - x \leq 1 \\ 0 & \text{si no.} \end{cases}$$

   Hallar la probabilidad del evento $[X < Y]$.

   SOLUCIÓN: Si graficamos la región de integración (intersección del rango conjunto con la zona donde $y > x$, obtenemos que:

$$\mathbb{P}([X < Y]) = \int_{0}^{\frac{1}{1+\sqrt{2}}} \int_{x}^{\frac{1-x}{\sqrt{2}}} 2\sqrt{2}\,dydx = \frac{1}{1+\sqrt{2}} \simeq 0,4142.$$

DISTRIBUCIONES MARGINALES.

Si X e Y son dos variables aleatorias conjuntas, como tales tienen su distribución conjunta. Cada una de ellas considerada individualmente posee su propia distribución.

Para referirnos a la distribución individual de X ó de Y utilizaremos el término “marginal”, así, diremos distribución marginal de X, función de distribución marginal de X, etc.

Dada la distribución conjunta de X e Y, queremos calcular las distribuciones marginales a partir de ésta. Esto resulta directamente de las propiedades de la función de distribución conjunta $F(x,y)$, pues:

$$F_X(x) = \mathbb{P}([X \le x]) = \lim_{y\to\infty} \mathbb{P}([X \le x, Y \le y]) = \lim_{y\to\infty} F(x,y)$$
$$F_Y(x) = \mathbb{P}([Y \le y]) = \lim_{x\to\infty} \mathbb{P}([X \le x, Y \le y]) = \lim_{x\to\infty} F(x,y)$$

Veamos ahora cómo calcular las distribuciones marginales sin pasar por la función de distribución $F(x,y)$.

1. **CASO DISCRETO:**

   Si X e Y son ambas discretas, con función de probabilidad conjunta

   $$g_{XY}(m,n) = \mathbb{P}([X = m, Y = n]),$$

   obtenemos a partir de ella las funciones de probabilidad marginal, mediante el siguiente cálculo:

   $$g_X(m) = \mathbb{P}([X = m]) = \sum_{n\in R(Y)} \mathbb{P}([X = m, Y = n]) =$$
   $$\sum_{n\in R(Y)} g_{XY}(m,n),\ m \in R(X)$$
   $$g_Y(n) = \mathbb{P}([Y = n]) = \sum_{m\in R(X)} \mathbb{P}([X = m, Y = n]) =$$
   $$\sum_{m\in R(X)} g_{XY}(m,n),\quad n \in R(Y)$$

2. **CASO CONTINUO**

   Si Xe Y son ambas continuas con densidad conjunta $f(x,y)$ entonces puesto que

   $$F_X(x) = \int_{-\infty}^{x} f_X(s)ds = \lim_{y\to\infty} F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(s,t)dtds$$

   concluimos que:

   $$f_X(x) = \int_{-\infty}^{\infty} f(x,t)dt$$

   Análogamente se obtiene que:

   $$f_Y(y) = \int_{-\infty}^{\infty} f(s,y)ds$$

3. **CASO MIXTO:**

   Si X es continua e Y es discreta con densidad-función de probabilidad $h(x,n)$, $(x,n) \in R(X,Y)$, la densidad de X puede obtenerse derivando su función de distribución marginal:

   $$F_X(x) = \lim_{y\to\infty} F(x,y) = \lim_{y\to\infty} \sum_{n\le y} \int_{-\infty}^{x} h(s,n)ds = \sum_{n\in R(Y)} \int_{-\infty}^{x} h(s,n)ds$$
   $$f_X(x) = \frac{dF_X(x)}{dx} = \sum_{n\in R(Y)} h(x,n).$$

   Para hallar la función de probabilidad de Y, hacemos el siguiente cálculo:

   $$g_Y(n) = \mathbb{P}([Y = n]) = \mathbb{P}([-\infty < X < \infty, Y = n]) = \int_{-\infty}^{\infty} h(x,n)dx.$$

DISTRIBUCIONES CONDICIONALES.

Si X e Y son dos variables aleatorias conjuntas y se tiene la información de que $Y = y$, qué podemos decir acerca de la distribución de X condicionada por esta información?

1. **CASO DISCRETO:**

   Si X e Y son dos variables aleatorias conjuntas y

   $$g_{XY}(m,n),\ (m,n) \in R(X,Y),$$

   su función de probabilidad conjunta. Si tenemos la información que sucedió el evento $[Y = n]$, $n \in R(Y)$, lo que se quiere calcular es la probabilidad condicional

   $$\mathbb{P}\left(X = m \mid Y = n\right), \quad m \in R(X),$$

   la cual está perfectamente bien definida ya que $\mathbb{P}\left([Y = n]\right)$ es diferente de cero, si $n \in R(Y)$. Denotamos esta probabilidad por:

   $$g_{X/Y}(m;n) = \mathbb{P}\left(X = m \mid Y = n\right) = \frac{\mathbb{P}\left(X = m, Y = n\right)}{\mathbb{P}\left(Y = n\right)} = \frac{g_{XY}(m,n)}{g_Y(n)},\ m \in R(X), n \in R(Y)$$

   Análogamente se obtiene, la función de probabilidad de Y dado $X = m$:

   $$g_{X/Y}(m;n) = \mathbb{P}\left(Y = n \mid X = m\right) = \frac{g_{XY}(m,n)}{gX(m)},\ m \in R(X), n \in R(Y)$$

2. **CASO CONTINUO**

   Si X e Y son dos variables aleatorias de densidad conjunta $f(x,y)$ y de función de distribución conjunta $F(x,y),$ al condicionar X por el suceso $[Y = y]$, tenemos que ser más cuidadosos, pues este evento tiene probabilidad nula para todo número real y. Sin embargo el suceso

   $$[y < Y \leq y + k]$$

   puede tener probabilidad no nula, entonces vamos a definir la función de distribución condicional de X dado el suceso $[Y = y]$ como:

   $$F_{X/Y}(x;y) = \begin{cases} \lim\limits_{k \to 0} \mathbb{P}\left(X \leq x \mid y < Y \leq y + k\right) & \text{si } \mathbb{P}\left(y < Y \leq y + k\right) \neq 0 \\ 0 & \text{en otro caso.} \end{cases}$$

   Al derivar con respecto a x obtenemos la densidad condicional de X dado el suceso $Y = y$, la cual denotaremos por $f_{X/Y}(x;y)$. Procedemos como sigue:

   $$F_{X/Y}(x;y) = \lim_{k \to 0} \frac{\mathbb{P}\left(X \leq x, y < Y \leq y + k\right)}{\mathbb{P}\left(y < Y \leq y + k\right)} = \lim_{k \to 0} \frac{\frac{F(x,y+k) - F(x,y)}{k}}{\frac{F_Y(y+k) - F_Y(y)}{k}} = \frac{\frac{\partial F(x,y)}{\partial y}}{f_Y(y)},$$

   por lo tanto

   $$f_{X/Y}(x;y) = \frac{\partial F_{X/Y}(x;y)}{\partial x} = \frac{\frac{\partial^2 F_{X/Y}(x;y)}{\partial x \partial y}}{f_Y(y)} = \frac{f(x,y)}{f_Y(y)}.$$

   Análogamente se obtiene la densidad condicional de Y dado el suceso $[X = x]$:

   $$f_{Y/X}(y;x) = \frac{f(x,y)}{f_X(x)}$$

   donde f_Y es la densidad marginal de Y y f_X la de X y F_Y es la función de distribución marginal de Y.

3. **CASO MIXTO:**

Si X es continua e Y es discreta con densidad-función de probabilidad $h(x,n)$, $(x,n) \in R(X,Y)$, mediante razonamientos análogos a los anteriores, obtenemos que la función de distribución de X condicionada por $[Y=n]$, $n \in R(Y)$, es

$$F_{X/Y}(x;n) = \frac{\int\limits_{-\infty}^{x} h(s,n)ds}{g_Y(n)}.$$

La densidad condicional de X dado $[Y=n]$, $n \in R(Y)$, se obtiene derivando respecto de x :

$$f_{X/Y}(x;n) = \frac{h(x,n)}{g_Y(n)},\ n \in R(Y),\ x \in \mathbb{R}.$$

La función de probabilidad de Y dado que $X = x$, por tener el suceso que condiciona probabilidad nula, se define

$$g_{Y/X}(n;x) = \begin{cases} \lim\limits_{h\to 0} \mathbb{P}\left(Y=n \mid x < X \leq x+h\right) & \text{si } \mathbb{P}\left(x < X \leq x+h\right) \neq 0 \\ 0 \quad \text{en otro caso.} \end{cases}$$

de donde

$$g_{Y/X}(n;x) = \lim_{h\to 0} \frac{\mathbb{P}\left(Y=n, x < X \leq x+h\right)}{\mathbb{P}\left(x < X \leq x+h\right)} =$$

$$\lim_{h\to 0} \frac{\frac{\mathbb{P}(Y=n, x<X\leq x+h)}{k}}{\frac{\mathbb{P}(x<X\leq x+h)}{k}} = \frac{h(x,n)}{f_X(x)}.$$

ESPERANZAS Y VARIANZAS CONDICIONALES.

Las esperanzas y varianzas correspondientes a las distribuciones condicionales, se denominan esperanzas y varianzas condicionales. Dependen, en general, de la variable que condiciona.

Las denotaremos por

$$\mathbb{E}(X \mid y),\ V(X \mid y)$$

si se trata de esperanza y varianza condicional de X dado $Y = y$, donde obviamente

$$V(X \mid y) = \mathbb{E}\left((X - \mathbb{E}(\mathbb{X}))^2 \mid y\right) = \mathbb{E}(X^2 \mid y) - \left(\mathbb{E}(X \mid y)\right)^2$$

y en general, la esperanza de una función $h(X)$ dado el suceso $[Y = y]$ se denota por

$$\mathbb{E}(h(X) \mid y).$$

OBSERVACIÓN:

Es importante notar que si $\omega \in \Omega$ e $Y(\omega) = y$, lo que se está expresando es que

$$\mathbb{E}(h(X) \mid y) = \mathbb{E}(h(X) \mid Y(\omega))$$

por lo tanto este condicionamiento no es más que una función de la variable aleatoria Y.

En consecuencia, si queremos calcular la esperanza de esta esperanza condicional, por ser función de la variable aleatoria Y, (vamos a suponer que estamos en el caso continuo, los otros casos son análogos) debemos integrar dicha función respecto a la densidad marginal de Y :

$$\mathbb{E}\left(\mathbb{E}(h(X) \mid y)\right) = \int\limits_{-\infty}^{\infty} \mathbb{E}(h(X) \mid y) f_Y(y)dy = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h(x) f_{X/Y}(x;y) f_Y(y)dxdy =$$

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h(x) f(x,y)dxdy = \int\limits_{-\infty}^{\infty} h(x) f_X(x)dx = \mathbb{E}(h(X)).$$

Todo lo dicho anteriormente, se simetriza al caso

$$\mathbb{E}(Y \mid x), \mathbb{E}(h(Y) \mid x), V(Y \mid x).$$

EJEMPLOS:

1. Sean X e Y dos variables de densidad conjunta
$$f(x,y)\begin{cases} 2 & \text{si } x\geq 0, y\geq 0.x+y\leq 1\\ 0 & \text{en otro caso.}\end{cases}$$
Hallemos las densidades marginales:
$$\begin{array}{l} f_X(x)=\int\limits_{-\infty}^{\infty} f(x,y)dy=\int\limits_{0}^{1-x} 2dy=2(1-x),\quad x\in[0,1]\\ f_Y(y)=\int\limits_{-\infty}^{\infty} f(x,y)dx=\int\limits_{0}^{1-y} 2dy=2(1-y),\quad y\in[0,1]\end{array}$$
donde el rango conjunto es
$$R(X,Y)=\{(x,y): x\geq 0, y\geq 0, x+y\leq 1\}.$$
Las densidades condicionales son:
$$\begin{array}{l} f_{X/Y}(x;y)=\frac{f(x,y)}{f_Y(y)}=\frac{1}{1-y},\quad x\in[0,1-y],\quad 0\leq y<1\\ f_{Y/X}(y;x)=\frac{f(x,y)}{f_X(x)}=\frac{1}{1-x},\quad y\in[0,1-x],\quad 0\leq x<1\end{array}$$
que son uniformes en $[0,1-y]$ y $[0,1-x]$ respectivamente.
El cálculo de las esperanzas, varianzas marginales y condicionales, es elemental, se obtienen los siguientes resultados:
$$\begin{array}{l} \mathbb{E}(X)=\frac{1}{3},\ V(X)=\frac{1}{18},\ \ \mathbb{E}(Y)=\frac{1}{3},\ \ V(Y)=\frac{1}{18}\\ \mathbb{E}(X\mid y)=\frac{1-y}{2},\ V(X\mid y)=\frac{(1-y)^2}{12}, 0\leq y<1;\\ \mathbb{E}(Y\mid x)=\frac{1-x}{2},\ V(Y\mid x)=\frac{(1-x)^2}{12}, 0\leq x<1;\end{array}$$

2. Sean X e Y dos variables aleatorias conjuntas, X continua e Y discreta. Dado $Y=n$, la densidad condicional de X es:
$$f_{X/Y}(x;n)=\frac{e^{-x}x^n}{n!},\ \ x\geq 0$$
Note que la distribución de X condicionada por $Y=n$ es gamma de parámetros $\lambda=1, p=n+1$.
La distribución marginal de Y es geométrica, su función de probabilidad es:
$$g_Y(n)=(1-r)r^n,\ n=0,1,2,\cdots,$$
donde $r\in(0,1)$. Hallar la función de probabilidad condicional de Y dado $X=x$.
SOLUCIÓN:
Primero debemos hallar la densidad-función de probabilidad conjunta:
$$h(x,n)=g_Y(n)f_{X/Y}(x;n)=\frac{(1-r)r^ne^{-x}x^n}{n!}=(1-r)e^{-x}\frac{(rx)^n}{n!},$$
$$x\geq 0,\ n=0,1,2,\cdots$$
La densidad marginal de X se obtiene:
$$\begin{array}{l} f_X(x)=\sum\limits_{n=0}^{\infty} h(x,n)=(1-r)e^{-x}\sum\limits_{n=0}^{\infty}\frac{(rx)^n}{n!}=\\ (1-r)e^{-x}e^{rx}=(1-r)e^{-(1-r)x},\ \ x\geq 0,\end{array}$$
la cual es exponencial de parámetro $(1-r)$.
La función de probabilidad condicional de Y dado $X=x$ es
$$g_{Y/X}(n;x)=\frac{h(x,n)}{f_X(x)}=\frac{(rx)^ne^{-rx}}{n!},\ \ n=0,1,2,\cdots$$
que es de Poisson de parámetro rx.

VARIABLES ALEATORIAS INDEPENDIENTES:

Sean X e Y dos variables aleatorias conjuntas, diremos que X e Y son independientes si

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)$$

para todo $A, B \in \mathcal{B}(\mathbb{R})$.

PROPOSICIÓN:

Las siguientes afirmaciones son equivalentes:

1. X e Y son independientes.
2. La distribución condicional de cualquiera de ellas coincide con su distribución marginal.
3. En el caso continuo puro con densidad:
$$f(x,y) = f_X(x) f_Y(y)$$
En el caso discreto
$$g(m,n) = g_X(m) g_Y(n)$$
En el caso mixto
$$h(x,n) = f_X(x) g_Y(n)$$
4. La función de distribución conjunta es el producto de las marginales
$$F(x,y) = F_X(x) F_Y(y)$$
5. En el caso continuo puro con densidad, ésta es factorizable:
$$f(x,y) = t(x) l(y)$$
En el caso discreto, la función de probabilidad es factorizable
$$g(m,n) = t(m) l(n)$$
En el caso mixto, $h(x,n)$, es factorizable:
$$h(x,n) = t(x) l(n)$$
(La demostración de las equivalencias, se dejan como ejercicio para la práctica).

EJEMPLO:

Sean X e Y dos variables aleatorias de densidad conjunta

$$f(x,y) = 6e^{-(2x+3y)}, \text{ si } x \geq 0, y \geq 0$$

las variables son independientes pues la densidad es factorizable

$$f(x,y) = t(x) l(y)$$

con

$$t(x) = 6e^{-2x}, x \geq 0; \quad l(y) = e^{-3y}, \ y \geq 0.$$

Las densidades marginales son:

$$f_X(x) = \int_0^\infty 6e^{-2x} e^{-3y} dy = 2e^{-2x}, \ x \geq 0$$
$$f_Y(y) = \int_0^\infty 6e^{-2x} e^{-3y} dx = 3e^{-3y}, \ y \geq 0$$

como era de esperarse

$$f(x,y) = f_X(x) f_Y(y).$$

Observe que $t(x)$ debidamente normalizada es $f_X(x)$ y $l(y)$ debidamente normalizada es $f_Y(y)$.

EJERCICIOS PARA LA PRÁCTICA

1. Una caja tiene 5 bolas blancas y 5 bolas negras. Se extraen al azar 3 bolas, elegidas una a una, sin reemplazo. Se definen las siguientes variables aleatorias:

$$X = \begin{cases} 0 & \text{si la } 1^a \text{ bola extraída es blanca} \\ 1 & \text{si la } 1^a \text{ bola extraída es negra} \end{cases}$$
$$Y = \text{n}^o \text{ de bolas negras extraídas}$$

   (a) Halle la función de probabilidad condicional de Y dado X.

   (b) Halle $\mathbb{E}(Y \mid X = 0),\ \mathbb{E}(Y \mid X = 1)$.

   (c) Halle $\mathbb{E}(Y)$, $\mathbb{E}(X)$.

2. Sean X e Y dos variables aleatorias de densidad conjunta

$$f(x, y) = 24xy, \quad \text{si } x \geq 0, y \geq 0, x + y \leq 1$$

   (a) Halle las densidades marginales de cada variable.

   (b) Calcule $\mathbb{E}(X),\ \mathbb{E}(Y),\ V(X),\ V(Y)$.

   (c) Halle las densidades condicionales.

   (d) Calcule $\mathbb{E}(X \mid y),\ \mathbb{E}(Y \mid x),\ V(X \mid y),\ V(Y \mid x)$.

   (e) Son X e Y independientes?

3. Sean X e Y dos variables aleatorias de densidad conjunta

$$f(x, y) = e^{-(x+y)}, \quad \text{si } x \geq 0, y \geq 0.$$

   (a) Halle las densidades marginales

   (b) Halle las densidades condicionales

   (c) Calcule $\mathbb{P}(X < Y \mid X < 2Y)$

   (d) Calcule $\mathbb{P}(X > Y \mid X + Y < 1)$

   (e) Calcule $\mathbb{P}\left((X > 1) \bigcup (Y > 1)\right)$

   (f) Son X e Y independientes?

4. Sean X e Y dos variables aleatorias de densidad conjunta

$$f(x, y) = 4xy, \quad \text{si } 0 \leq x \leq 1, 0 \leq y \leq 1$$

   (a) Halle las densidades marginales.

   (b) Calcule $\mathbb{E}(X),\ \mathbb{E}(Y),\ V(X),\ V(Y)$.

   (c) Son X e Y independientes? Justifique la respuesta.

5. Se extrae una muestra de tamaño 2 con reposición de una bolsa que contiene:dos bolas blancas, una negra y dos rojas. Definimos las siguientes variables aleatorias:

$$X_1 = \begin{cases} 1, & \text{si la } 1^a \text{ bola es blanca} \\ 0, & \text{si la } 1^a \text{ bola no es blanca} \end{cases}$$
$$X_2 = \begin{cases} 1, & \text{si la } 2^a \text{ bola es blanca} \\ 0, & \text{si la } 2^a \text{ bola no es blanca} \end{cases}$$

   (a) Halle la función de probabilidad conjunta de estas variables.

(b) Halle las funciones de probabilidad marginales.

(c) Son independientes estas variables?

(d) Qué sucede si el muestreo se realiza sin reposición?

6. Supongamos que las variables aleatorias X e Y tienen como rango conjunto

$$R(X,Y) = \{(0,0),(-1,1),(1,-1),(-1,-1),(1,1)\}$$

cada uno con probabilidad $\frac{1}{5}$.

(a) Halle las distribuciones marginales.

(b) Determine si estas variables son independientes.

7. En dos lugares de una habitación se mide la intensidad del ruido. Sean X e Y las variables aleatorias que representan la intensidad del ruido en estos puntos y supongamos que la densidad conjunta de estas variables es:

$$f(x,y) = \begin{cases} xy\exp\left(-\frac{1}{2}(x^2+y^2)\right), & \text{si } x>0, y>0 \\ 0 & \text{en otro caso} \end{cases}$$

(a) Calcule las densidades marginales.

(b) Halle $\mathbb{P}(X \leq 1, Y \leq 1)$

(c) $\mathbb{P}(X+Y \leq 1)$

(d) Son independientes estas variables aleatorias?

8. Sean X e Y dos variables aleatorias independientes con la misma función de distribución F. Sea $Z = \max(X,Y)$. Cuál es la función de distribución $G(z)$ de Z?

9. Sea $(X;Y)$ dos variables con densidad conjunta

$$f(x,y) = \begin{cases} 2, & \text{si } (x,y) \in ((0,0.5)\times(0,0.5))\bigcup((-0.5,0)\times(-0.5,0)) \\ 0 & \text{en otro caso} \end{cases}$$

(a) Verifique que f es una densidad.

(b) Demuestre que X e Y no son independientes.

(c) Demuestre que X^2 e Y^2 son independientes.

10. Sean X e Y dos variables aleatorias independientes con función de probabilidad uniforme en

$$\{1,2,3,\cdots,N\}$$

(es decir que $\mathbb{P}(X=i) = \mathbb{P}(Y=i) = \frac{1}{N}$, $i = 1,2,\cdots N$).

Calcule la función de probabilidad de $X+Y$.

Respuesta

$$\begin{array}{l} \mathbb{P}(X+Y=j) = 0, \;\; j<2 \quad \text{ó } j>2N \\ \mathbb{P}(X+Y=j) = \frac{j-1}{N^2}, \;\; 2 \leq j \leq N \\ \mathbb{P}(X+Y=j) = \mathbb{P}(X+Y=N+i) = \frac{N-i+1}{N^2} = \frac{2N-j+1}{N^2} \\ \text{si } \; N+1 \leq j \leq 2N, \text{ con } i = j-N \end{array}$$

11. Sean X e Y dos variables aleatorias independientes con distribución uniforme en $(0,1)$. Halle la densidad de $X+Y$.

12. Sean X e Y dos variables aleatorias independientes con distribución exponencial de parámetro λ. Halle la densidad de la variable $Z = \max(X,Y)$.

Capítulo 7

ESPERANZA DE UNA FUNCIÓN VARIABLES ALEATORIAS CONJUNTAS:

ESPERANZA DE UNA FUNCIÓN VARIABLES ALEATORIAS CONJUNTAS:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidades, X e Y dos variables aleatorias conjuntas:

1. Sea $\mathcal{U} : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $Z = \mathcal{U}(X, Y)$ una variable aleatoria real. Para calcular la esperanza de Z, basta conocer la distribución conjunta de las variables X e Y.

   (a) Si X e Y tienen densidad conjunta $f(x, y)$

   $$\mathbb{E}(Z) = \mathbb{E}(\mathcal{U}(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{U}(x, y) f(x, y) dx dy$$

   (b) Si X e Y son discretas, con función de probabilidad conjunta

   $$g(m, n) = \mathbb{P}(X = m, Y = n)$$

   la esperanza

   $$\mathbb{E}(Z) = \mathbb{E}(\mathcal{U}(X, Y)) = \sum_{m \in R(X)} \sum_{n \in R(Y)} \mathcal{U}(m, n) g(m, n)$$

   (c) Si X es continua e Y es discreta, con densidad-función de probabilidad conjunta $h(x, n)$

   $$\mathbb{E}(Z) = \mathbb{E}(\mathcal{U}(X, Y)) = \sum_{n \in R(Y)} \int_{-\infty}^{\infty} \mathcal{U}(x, n) h(x, n) dx.$$

2. Sea $\mathcal{U} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $\mathcal{U}(X, Y) = (\mathcal{U}_1(X, Y), \mathcal{U}_2(X, Y))$, donde

   $$(Z_1, Z_2) = (\mathcal{U}_1(X, Y), \mathcal{U}_2(X, Y))$$

   son dos variables aleatorias conjuntas. La esperanza de (Z_1, Z_2) , es el vector fila:

   $$\mathbb{E}(\mathcal{U}(X, Y)) = (\mathbb{E}(\mathcal{U}_1(X, Y)), \mathbb{E}(\mathcal{U}_2(X, Y)))$$

3. Si $\mathcal{U}$ es una matriz cuadrada de orden 2 :

$$\mathcal{U}(X,Y)=\begin{pmatrix} \mathcal{U}_{11}(X,Y) & \mathcal{U}_{12}(X,Y) \\ \mathcal{U}_{21}(X,Y) & \mathcal{U}_{22}(X,Y) \end{pmatrix}$$

donde

$$\mathcal{U}_{ij}:\mathbb{R}^2 \longrightarrow \mathbb{R},$$

$\mathcal{U}_{ij}(X,Y)$ es una variable aleatoria, $i=1,2;\ j=1,2.$ Definimos:

$$\mathbb{E}(\mathcal{U}(X,Y))=\begin{pmatrix} \mathbb{E}\left(\mathcal{U}_{11}(X,Y)\right) & \mathbb{E}\left(\mathcal{U}_{12}(X,Y)\right) \\ \mathbb{E}\left(\mathcal{U}_{21}(X,Y)\right) & \mathbb{E}\left(\mathcal{U}_{22}(X,Y)\right) \end{pmatrix}$$

EJEMPLOS:

1. **MOMENTOS:** El momento conjunto de orden i,j de X e Y se define como:

$$\mu_{ij}=\mathbb{E}\left(X^iY^j\right)$$

en particular, si $j=0$

$$\mu_{i0}=\mathbb{E}\left(X^i\right)$$

es el momento de orden i de la distribución marginal de X.

2. **COVARIANZA:** la covarianza de X e $Y,$ se define como:

$$COV(X,Y)=\mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right)$$

la cual puede expresarse, por simple cálculo como:

$$COV(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$$

3. **ESPERANZA CONJUNTA DEL VECTOR ALEATORIO** (X,Y) :

$$\mathbb{E}\left((X,Y)\right)=\left(\mathbb{E}(X),\mathbb{E}(Y)\right)$$

4. **MATRIZA DE COVARIANZA:** Varianza del vector aleatorio (X,Y) : sea $\overrightarrow{Z}=(X,Y),$ se define la varianza del vector aleatorio $\overrightarrow{Z},$ como:

$$Var(\overrightarrow{Z})=\mathbb{E}\left((\overrightarrow{Z}-\mathbb{E}(\overrightarrow{Z}))^T(\overrightarrow{Z}-\mathbb{E}(\overrightarrow{Z}))\right)$$

donde

$$(\overrightarrow{Z}-\mathbb{E}(\overrightarrow{Z}))^T=\left((X-\mathbb{E}(X)),(Y-\mathbb{E}(Y))\right)^T=\begin{pmatrix} X-\mathbb{E}(X) \\ Y-\mathbb{E}(Y) \end{pmatrix}$$

por lo tanto, $Var(\overrightarrow{Z})$ es una matriz cuadrada de orden $2,$ simétrica, sus elementos diagonales son las varianzas marginales de las variables aleatorias X e Y respectivamente y los elementos no diagonales son las covarianzas, esta matriz se llama, matriz de varianzas y covarianzas del vector aleatorio $\overrightarrow{Z}$ y se expresa como:

$$Var(\overrightarrow{Z})=\begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}$$

PROPOSICIÓN:

Sean $\mathcal{U} : \mathbb{R} \longrightarrow \mathbb{R}$ y $\mathcal{V} : \mathbb{R} \longrightarrow \mathbb{R}$, X e Y dos variables aleatorias conjuntas. Si X e Y son independientes, entonces

$$\mathbb{E}(\mathcal{U}(X)\mathcal{V}(Y)) = \mathbb{E}(\mathcal{U}(X))\mathbb{E}(\mathcal{V}(Y))$$

en particular

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

de donde

$$Cov(X,Y) = 0.$$

DEMOSTRACIÓN:

Supongamos que X e Y tienen densidad conjunta $f(x,y)$, (los otros casos son similares). Si X e Y son independientes

$$f(x,y) = f_X(x)f_Y(y), \ \ x,y \in \mathbb{R}.$$

Por lo tanto:

$$\begin{array}{l} \mathbb{E}(\mathcal{U}(X)\mathcal{V}(Y)) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \mathcal{U}(x)\mathcal{V}(y)f_X(x)f_Y(y)dxdy = \\ \int\limits_{-\infty}^{\infty} \mathcal{U}(x)f_X(x)dx \int\limits_{-\infty}^{\infty} \mathcal{V}(y)f_Y(y)dy = \mathbb{E}(\mathcal{U}(X))\mathbb{E}(\mathcal{V}(Y)) \end{array}$$

Esperanza y Varianza de la suma de variables aleatorias:

Sean X e Y dos variables aleatorias conjuntas, entonces:

1. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y), \ \ a,b \in \mathbb{R},$ en particular

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

DEMOSTRACIÓN: Supongamos que X e Y tienen densidad conjunta $f(x,y)$, (los otros casos son similares).

$$\begin{array}{l} \mathbb{E}(aX+bY) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (ax+by)f(x,y)dxdy = \\ a\int\limits_{-\infty}^{\infty} x \int\limits_{-\infty}^{\infty} f(x,y)dydx + b\int\limits_{-\infty}^{\infty} y \int\limits_{-\infty}^{\infty} f(x,y)dxdy = \\ a\int\limits_{-\infty}^{\infty} xf_X(x)dx + b\int\limits_{-\infty}^{\infty} yf_Y(y)dy = a\mathbb{E}(X) + b\mathbb{E}(Y). \end{array}$$

2. para la varianza no hay linealidad

$$Var(aX+bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$$

DEMOSTRACIÓN:

$$\begin{array}{l} Var(aX+bY) = \mathbb{E}\left((aX+bY)^2\right) - \left(\mathbb{E}(aX+bY)\right)^2 = \\ a^2\mathbb{E}(X^2) + b^2\mathbb{E}(Y^2) + 2ab\mathbb{E}(XY) - a^2\left(\mathbb{E}(X)\right)^2 - b^2\left(\mathbb{E}(Y)\right)^2 - 2aba^2\mathbb{E}(X)\mathbb{E}(Y) \\ = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y). \end{array}$$

En particular, si X e Y son independientes

$$Var(aX+bY) = a^2Var(X) + b^2Var(Y).$$

COEFICIENTE DE CORRELACIÓN:

DEFINICIÓN:

Si X e Y son dos variables aleatorias conjuntas, el coeficiente de correlación de X e Y se define como:

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

donde $Var(X)$ y $Var(Y)$ son las varianzas marginales de X e Y.

PROPOSICIÓN:

Si X e Y son dos variables aleatorias conjuntas:

1. $\rho(X,Y)\in[-1,1]$
2. Si
$$Z_1=aX+b,\ Z_2=cY+d$$
entonces
$$\rho(Z_1,Z_2)=\pm\rho(X,Y)$$
3. $|\rho(X,Y)|=1$ si y sólo si existen constantes a y b, con a no nulo, tal que
$$\mathbb{P}\left(Y=aX+b\right)=1$$
(esto significa que entre las variables existe una correlación perfecta).
4. Si X e Y son independientes entonces
$$\rho(X,Y)=0.$$

DEMOSTRACIÓN:

1. Sea $\mathcal{U}_t(X,Y)=\left[(X-\mathbb{E}(X))-t(Y-\mathbb{E}(Y))\right]^2$, obviamente $\mathcal{U}_t\geq 0$ para todo $t\in\mathbb{R}$, luego, $\mathbb{E}(\mathcal{U}_t(X,Y))\geq 0$ para todo t, esto se traduce por:
$$Var(X)-2tCov(X,Y)+t^2Var(Y)\geq 0$$
para todo t. Esta expresión, como función de t, representa una parábola, cuyo discriminante es:
$$4\left(Cov(X,Y)\right)^2-4Var(X)Var(Y)$$
puesto que el coeficiente del término en t^2 de la parábola es
$$Var(Y)>0,$$
debemos exigir que el discriminante
$$4\left(Cov(X,Y)\right)^2-4Var(X)Var(Y)\leq 0$$
de donde se obtiene
$$\rho^2(X,Y)\leq 1,$$
de aquí se obtiene el resultado.
2. $(Z_1-\mathbb{E}(Z_1))=a(X-\mathbb{E}(X)),\ \ (Z_2-\mathbb{E}(Z_2))=c(Y-\mathbb{E}(Y))$, obtenemos que
$$Cov(Z_1,Z_2)=acCov(X,Y).$$
Por otro lado
$$Var(Z_1)=a^2Var(X),\quad Var(Z_2)=b^2Var(Y)$$
por lo tanto
$$\rho(Z_1,Z_2)=\frac{acCov(X,Y)}{|ac|\sqrt{Var(X)Var(Y)}}=\pm\rho(X,Y).$$

3. Supongamos que $|\rho(X,Y)| = 1,$ si definimos

$$Z_1 = \frac{X - \mathbb{E}(X)}{D(X)},\ Z_2 = \frac{Y - \mathbb{E}(Y)}{D(Y)}$$

donde $D(X)$ y $D(Y)$ representan las desviaciones estándar de las variables aleatorias X e Y respectivamente, es decir:

$$D(X) = \sqrt{Var(X)},\ D(Y) = \sqrt{Var(Y).}$$

De la propiedad anterior se deduce que

$$\rho(Z_1, Z_2) = \rho(X,Y) = \pm 1$$

Pero, $\rho(Z_1, Z_2) = Cov(Z_1, Z_2) = \pm 1$ implica en uno de los dos casos que $Var(Z_1 - Z_2) = 0$ y en el otro que $Var(Z_1 + Z_2) = 0$.

Cuando una variable aleatoria Z tiene varianza nula, esto significa que $Z = \mathbb{E}(Z)$ con probabilidad $1,$ en ambos casos se concluye que existen constantes a y b, con a no nulo, tal que

$$\mathbb{P}\left(Y = aX + b\right) = 1.$$

Recíprocamente, si la masa total está concentrada en una recta, es decir

$$\mathbb{P}\left(Y = aX + b\right) = 1$$

con a no nulo, entonces

$$\begin{aligned} Cov(X,Y) = \mathbb{E}(X(aX+b)) - \mathbb{E}(X)\mathbb{E}(aX+b) = \\ a\mathbb{E}(X^2) + b\mathbb{E}(X) - a\left(\mathbb{E}(X)\right)^2 - b\mathbb{E}(X) = aVar(X) \end{aligned}$$

por lo tanto

$$\rho(X,Y) = \frac{aVar(X)}{\sqrt{Var(X)a^2Var(X)}} = \frac{a}{|a|} = \begin{cases} 1, \text{ si } a > 0 \\ -1, \text{ si } a < 0 \end{cases}$$

Conclusión: el coeficiente de correlación toma uno de los valores extremos del intervalo $[-1,1]$ si existe una dependencia lineal completa entre las variables aleatorias. Si $|\rho| \in (-1,1)$ no existe ninguna recta que contenga la masa total de la distribución de X e Y. Así, el coeficiente de correlación mide, en cierto sentido el grado de dependencia lineal entre las variables.

4. Si X e Y son independientes, $Cov(X,Y) = 0,$ en particular también lo es el coeficiente de correlación.

   La recíproca en general no es cierta. Si $\rho(X,Y) = 0$, diremos que X e Y no están correlacionadas.

EJEMPLOS:

1. Sea X e Y dos variables aleatorias discretas de función de probabilidad conjunta

$$g(-1,1) = g(1,1) = \frac{1}{4},\ g(0,-1) = \frac{1}{2}$$

Hallemos, las distribuciones marginales y la covarianza:

$$g_X(m) = \sum_{n \in R(Y)} g(m,n) = \begin{cases} \frac{1}{4}, & \text{si } m = -1 \\ \frac{1}{4}, & \text{si } m = 1 \\ \frac{1}{2}, & \text{si } m = 0 \end{cases}$$

Análogamente obtenemos:

$$g_Y(n) = \sum_{m \in R(X)} g(m,n) = \begin{cases} \frac{1}{2}, & \text{si } n = 1 \\ \frac{1}{2}, & \text{si } n = -1 \end{cases}$$

Para hallar $\mathbb{E}(XY)$ nos valemos de la función de probabilidad conjunta:

$$\mathbb{E}(XY) = \sum_{(m,n)\in R(X,Y)} m \cdot n \cdot g(m,n) = 0$$

Para calcular $\mathbb{E}(X)$ y $\mathbb{E}(Y)$ nos valemos de las funciones de probabildad marginales:

$$\begin{aligned} \mathbb{E}(X) &= \sum_{m\in R(X)} g_X(m) = 0 \\ \mathbb{E}(Y) &= \sum_{n\in R(Y)} g_Y(n) = 0 \end{aligned}$$

De aquí se deduce que $Cov(X,Y) = 0$ y sin embargo las variables aleatorias X e Y no son independientes ya que

$$g(1,1) = \frac{1}{4}, \;\; g_X(1)g_Y(1) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}.$$

Este ejemplo muestra que la condición $Cov(X;Y) = 0$ no es una condición suficiente para obtener la independencia de las variables aleatorias X e Y.

2. Sean X e Y dos variables aleatorias de densidad conjunta:

$$f(x,y) = \begin{cases} 2 & \text{si } 0 \le x \le y \le 1 \\ 0 & \text{si no.} \end{cases}$$

Hallemos el coeficiente de correlación de las variables:

las densidades marginales son:

$$\begin{aligned} f_X(x) &= \int_x^1 2dy = 2(1-x), \quad \text{si } x \in [0,1] \\ f_Y(y) &= \int_0^y 2dx = 2y, \quad \text{si } y \in [0,1] \end{aligned}$$

Las esperanzas de las variables son:

$$\begin{aligned} \mathbb{E}(X) &= \int_0^1 2x(1-x)dx = \tfrac{1}{3} \\ \mathbb{E}(Y) &= \int_0^1 2y^2 dy = \tfrac{2}{3} \end{aligned}$$

Para hallar la varianza de ellas, necesitamos calcular:

$$\begin{aligned} \mathbb{E}(X^2) &= \int_0^1 2x^2(1-x)dx = \tfrac{1}{6} \\ \mathbb{E}(Y^2) &= \int_0^1 2y^3 dy = \tfrac{1}{2} \end{aligned}$$

de aquí obtenemos:

$$\begin{aligned} Var(X) &= \tfrac{1}{6} - \tfrac{1}{9} = \tfrac{1}{18} \\ Var(Y) &= \tfrac{1}{2} - \tfrac{4}{9} = \tfrac{1}{18} \end{aligned}$$

Hallemos la $Cov(X,Y)$ y $\rho(X,Y)$:

$$\begin{aligned} \mathbb{E}(XY) &= \int_0^1 \int_x^1 2xydydx = \tfrac{1}{4} \\ Cov(X,Y) &= \tfrac{1}{4} - \tfrac{2}{9} = \tfrac{1}{36} \\ \rho(X,Y) &= \tfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \tfrac{1}{2}. \end{aligned}$$

ESPERANZA DE LA ESPERANZA:

Sean X e Y dos variables aleatorias conjuntas, entonces si $\mathcal{U}$ es una función definida sobre $\mathbb{R}^2$:

$$\begin{aligned} \mathbb{E}\left(\mathbb{E}(\mathcal{U}(X,Y) \mid y)\right) &= \mathbb{E}(\mathcal{U}(X,Y)) \\ \mathbb{E}\left(\mathbb{E}(\mathcal{U}(X,Y) \mid x)\right) &= \mathbb{E}(\mathcal{U}(X,Y)) \end{aligned}$$

DEMOSTRACIÓN:

Haremos la demostración de la primera igualdad, (la segunda es simétrica) en el caso en que $\mathcal{U} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ y las variables aleatorias X e Y tienen densidad conjunta $f(x,y)$. En otros casos la demostración es análoga.

Para calcular $\mathbb{E}(\mathcal{U}(X,Y) \mid y)$, necesitamos la densidad condicional

$$f_{X|Y}(x;y) = \frac{f(x,y)}{f_Y(y)}$$

así:

$$\mathbb{E}(\mathcal{U}(X,Y) \mid y) = \int_{-\infty}^{\infty} \mathcal{U}(x,y) f_{X|Y}(x;y)dx = h(y)$$

Para hallar la esperanza de $h(Y)$, nos valemos de la densidad marginal de Y :

$$\mathbb{E}\left(\mathbb{E}(\mathcal{U}(X,Y) \mid y)\right) = \int_{-\infty}^{\infty} h(y) f_Y(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{U}(x,y) f_{X|Y}(x;y) f_Y(y)dxdy =$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{U}(x,y) f(x,y)dxdy = \mathbb{E}(\mathcal{U}(X,Y)).$$

DISTRIBUCIÓN NORMAL BIDIMENSIONAL:

Sean X e Y dos variables aleatorias conjuntas de matriz de covarianza Δ. Si el determinante de Δ es no nulo, diremos que la distribución conjunta de X e Y es normal si su densidad conjunta $f(x,y)$ viene dada por:

$$\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right)\right]$$

donde

$$\begin{aligned} &x, y \in \mathbb{R}, \sigma_x^2 = Var(X), \sigma_y^2 = Var(Y) \\ &\mu_x = \mathbb{E}(X), \ \ \mu_y = \mathbb{E}(Y), \ \ \rho = \rho(X,Y). \end{aligned}$$

EJERCICIOS PARA LA PRÁCTICA.

1. Sean X e Y dos variables aleatorias conjuntas de matriz de covarianza Δ. Verifique que
$$Det(\Delta) = Var(X)Var(Y)(1-\rho^2)$$
2. Sean X e Y dos variables aleatorias de densidad conjunta normal bidimensional. Demuestre que si $\rho(X,Y) = 0$ entonces X e Y son independientes.
3. Sea Z una variable aleatoria de distribución $N(\mu,\sigma)$. Demuestre que su transformada de Fourier ó función característica viene dad por:
$$\varphi_Z(t) = \mathbb{E}(e^{itZ}) = e^{it\mu}e^{-\frac{\sigma^2 t^2}{2}} = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)$$
La tranformada de Fourier identifica la distribución de la variable aleatoria, es decir si una variable aleatoria Z tiene esta transformada de Fourier, concluiremos que su distribución es $N(\mu,\sigma)$.
4. Sean X e Y dos variables aleatorias conjuntas de transformadas de Fourier
$$\varphi_X(t) = \mathbb{E}(e^{itX}), \quad \varphi_Y(t) = \mathbb{E}(e^{itY})$$
respectivamente. Si X e Y son independientes halle
$$\varphi_{X+Y}(t).$$
(La transformada de Fourier de la suma de dos variables aleatorias independientes).
5. Sean X e Y dos variables aleatorias independientes de densidad conjunta
$$f(x,y) = f_X(x)f_Y(y).$$
Sea $Z = X+Y$, demustre que la densidad de Z es
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-x)f_Y(x)dx$$
Esta expresión se conoce con el nombre de convoluciones de las funciones
$$f_X \quad \text{y} \quad f_Y.$$
6. Sean X e Y dos variables aleatorias independientes de densidad conjunta normal bidimensional. Demuestre usando el ejercicio anterior que la variable aleatoria
$$Z = X+Y$$
tiene distribución normal $N(\mu,\sigma)$ con
$$\mu = \mu_x + \mu_y, \quad \sigma^2 = \sigma_x^2 + \sigma_y^2$$
7. Sean X e Y dos variables aleatorias independientes de densidad conjunta normal bidimensional. Demuestre usando la transformada de Fourier que la variable aleatoria
$$Z = X+Y$$
tiene distribución normal $N(\mu,\sigma)$ con
$$\mu = \mu_x + \mu_y, \quad \sigma^2 = \sigma_x^2 + \sigma_y^2$$
Para esto siga los siguientes pasos:

(a) Si Z es $N(\mu,\sigma)$, halle su transformada de Fourier o función característica.

(b) Sea $Z = X + Y$, con X e Y independientes, de distribución

$$N(\mu_x,\, \sigma_x) \quad \text{y} \quad N(\mu_y, \quad \sigma_y)$$

respectivamente. Halle la función característica de la variable Z.

(c) Comparando (a) y (b), deduza que Z tiene distribución normal de parámetros

$$\mu = \mu_x + \mu_y, \ \ \sigma^2 = \sigma_x^2 + \sigma_y^2.$$

8. Sean X e Y dos variables aleatorias de densidad conjunta uniforme en el cuadrado de vértices $(1,0),(0,1),(-1,0),(0,-1)$

   (a) Son independientes?

   (b) Halle sus esperanzas, varianzas y covarianzas.

9. Sean X e Y dos variables aleatorias de densidad conjunta uniforme en el triángulo de vértices $(-1,0),(0,1),(1,0)$.

   (a) Escriba la densidad conjunta.

   (b) Calcule $\mathbb{P}\left(X \leq \frac{3}{4}, Y \leq \frac{3}{4}\right)$

   (c) Calcule $\mathbb{P}\left((X - Y) \geq 0\right)$

   (d) Calcule las densidades marginales de X e Y.

   (e) Calcule $\mathbb{E}(X), \mathbb{E}(Y), Var(X), Var(Y)$.

   (f) Calcule $\rho(X,Y)$.

10. Un estudio muestra que el n^o de horas diarias, X, que un adolescente dedica a ver televisión y el n^o de horas diarias, Y, que dedica a estudiar o a hacer tareas tienen aproximadamente la siguiente densidad conjunta:

$$f(x,y) = xye^{-(x+y)}, \qquad x > 0,\ y > 0.$$

   (a) Cuál es la probabilidad de que un adolescente dedique más del doble del tiempo a ver televisión que a sus estudios?

   (b) Son X e Y independientes?

   (c) Halle el promedio de horas diarias dedicadas al estudio.

11. Se lanza independientemente dos veces un dado homogéneo, sea U la variable que nos da el resultado del primer lanzamiento y V la variable que nos da el resultado del segundo lanzamiento.

    Sean

$$X = U + V, \qquad Y = U - V$$

    Demuestre que X e Y son no correlacionadas, es decir que:

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

    sin embargo, no son independientes.

12. Sean X e Y dos variables aleatorias de densidad conjunta normal bidimensional, de parámetros:

$$\begin{aligned} &\mathbb{E}\left((X,Y)\right) = (\mu_x, \mu_y) \\ &Var\left((X,Y)\right) = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \end{aligned}$$

    donde $\rho\sigma_x\sigma_y = COV(X,Y)$.

Demuestre que las distribuciones marginales de las variables X e Y son normales. (Identifique los parámetros)

Capítulo 8

TEOREMA CENTRAL DEL LÍMITE.

8.1 LEY DE GRANDES NÚMEROS. TEOREMA CENTRAL DEL LÍMITE.

VECTORES ALEATORIOS N-DIMENSIONALES:

Todas las definiciones y resultados obtenidos en el caso de variables aleatorias conjuntas se generaliza sin dificultad al caso de n variables aleatorias conjuntas.

Si $(X_1, X_2, \cdots, X_n)$ son n variables aleatorias conjuntas, es decir, el vector

$$(X_1, X_2, \cdots, X_n) : \Omega \to \mathbb{R}^n$$

es aleatorio, (la definición formal es una generalización del caso $n = 2$), su función de distribución conjunta es

$$F(x_1, x_2, \cdots, x_n) = \mathbb{P}\left(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n\right)$$

decir que las variables aleatorias $X_1, X_2, \cdots, X_n$ son independientes equivale a

$$F(x_1, x_2, \cdots, x_n) = F_{X_1}(x_1)F_{X_2(}x_2) \cdots F_{X_n}(x_n)$$

(es el producto de las funciones de distribución marginales).

Si la distribución conjunta tiene densidad

$$f(x_1, x_2, \cdots, x_n) = \frac{\partial^n F(x_1, x_2, \cdots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n)}$$

la independencia de las variables equivale a

$$f(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1)f_{X_2(}x_2) \cdots f_{X_n}(x_n)$$

TRANSFORMADA DE FOURIER DE LA SUMA DE VARIABLES ALEATORIAS.

Si $X_1, X_2, \cdots, X_n$ son n variables aleatorias

$$S_n = X_1 + X_2 + \cdots + X_n$$

es una variable aleatoria.

La transformada de Fourier de S_n es

$$\varphi_{S_n}(t) = \mathbb{E}\left(e^{itS_n}\right) = \mathbb{E}\left(e^{it(X_1+X_2+\cdots+X_n)}\right) = \mathbb{E}\left(\prod_{k=1}^{n} e^{itX_k}\right)$$

En particular, si $X_1, X_2, \cdots, X_n$ son independientes

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \mathbb{E}\left(e^{itX_k}\right) = \prod_{k=1}^{n} \varphi_{X_k}(t)$$

donde

$$\varphi_{X_k}(t) = \mathbb{E}\left(e^{itX_k}\right)$$

es la transformada de Fourier ó Función Característica de la variable aleatoria X_k.

TRANSFORMADA DE FOURIER DE LA DISTRIBUCIÓN NORMAL:

Si Z es una variable aleatoria de distribuciómn $N(\mu, \sigma)$, su transformada de Fourier es

$$\varphi_Z(t) = \mathbb{E}\left(e^{itZ}\right) = e^{i\mu t} e^{-\frac{\sigma^2 t^2}{2}}$$

El cálculo de esta transformada es un ejercicio de la práctica del capítulo anterior).

Si $X_1, X_2, \cdots, X_n$ son independientes, de distribución $N(\mu_k, \sigma_k), k = 1, 2, \cdots, n$;

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \varphi_{X_k}(t) = \prod_{k=1}^{n} \exp(i\mu_k t - \frac{\sigma_k^2 t^2}{2}) = \exp\left(\sum_{k=1}^{n} (i\mu_k t - \frac{\sigma_k^2 t^2}{2})\right)$$

mediante simple identificación de los parámetros, se concluye que S_n tiene distribución normal de parámetros

$$\mu = \sum_{k=1}^{n} \mu_k, \ \sigma^2 = \sum_{k=1}^{n} \sigma_k^2$$

LEY DE GRANDES NÚMEROS.

DESIGUALDAD DE TCHEBICHEV:

Sea X una variable aleatoria de varinaza σ^2 finita y $\mathbb{E}(X) = \mu$, entonces, para todo $\varepsilon > 0$

$$\mathbb{P}\left(|X - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

DEMOSTRACIÓN:

Haremos la demostración en el caso en que la variable aleatoria X tiene densidad $f(x)$, el caso discreto es similar:

$$\mathbb{P}\left(|X - \mu| \geq \varepsilon\right) = \int_{\{x:|X-\mu|\geq\varepsilon\}} f(x)dx \leq$$
$$\frac{1}{\varepsilon^2} \int_{\{x:|X-\mu|\geq\varepsilon\}} |X - \mu|^2 f(x)dx = \frac{\sigma^2}{\varepsilon^2}$$

en particular si $k\sigma = \varepsilon,\ k \geq 1$, la desigualdad se transforma en

$$\mathbb{P}\left(|X - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}$$

LEY DÉBIL DE GRANDES NÚMEROS.

Si $X_1, X_2, \cdots, X_n, \cdots$ es una sucesión de variables aleatorias independientes, definidas sobre el mismo espacio de probabilidad, con

$$\mathbb{E}(X_k) = \mu, \ \ Var(X_k) = \sigma^2, \ k = 1, 2, 3, \cdots$$

entonces para todo $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) = 0$$

DEMOSTRACIÓN: sea

$$S_n = X_1 + X_2 + \cdots + X_n$$

vamos aplicarle la desigualdad de Tchebichev a la variable aleatoria $\dfrac{S_n}{n}$:

$$\begin{array}{l}
\mathbb{E}(\dfrac{S_n}{n}) = \mathbb{E}(\frac{X_1+X_2+\cdots+X_n}{n}) = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(X_k) = \frac{n\mu}{n} = \mu \\
Var(\dfrac{S_n}{n}) = Var(\frac{X_1+X_2+\cdots+X_n}{n}) = \frac{1}{n^2}\sum_{k=1}^{n} Var(X_k) = \frac{\sigma^2}{n} \\
\mathbb{P}\left(\left|\dfrac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \dfrac{\sigma^2}{n\varepsilon^2} \to 0 \quad si \qquad n \to \infty.
\end{array}$$

■

OBSERVACIONES:

1. La ley débil de los grandes números nos dice que en el caso de variables aleatorias independientes (observaciones independientes de un experimento), el promedio

$$\frac{X_1 + X_2 + \cdots + X_n}{n}$$

   es una buena aproximación de la verdadera esperanza (ó valor esperado) de las variables y mientras más grande es el número de observaciones, mejor es la aproximación.

2. La desigualdad nos permite obtener una acotación de la velocidad con que el primer miembro tiende a cero si $n \to \infty$.

EJEMPLO:

Sea A un suceso y repitamos n veces el experimento que consiste en observar A.

Si cada repetición es independiente y nos interesamos en el número de veces que ocurre A, la variable que cuenta el número de veces que A ocurre, es Binomial. Sin embargo, podemos pensar el problema de la manera siguiente, definiendo:

$$X_i = \begin{cases} 1, & \text{si } A \text{ sucede en la } i-\text{ésima repetición} \\ 0 & \text{si no} \end{cases}$$

Obviamente

$$S_n = X_1 + X_2 + \cdots + X_n$$

representa el número de veces que A sucede en las n repreticiones del experimento. Al realizar el experimento n veces, el cociente $\dfrac{k_n}{n}$, representa el cociente entre el número de casos favoreables y el número de veces que se realizó el experimento. Si no conocemos $\mathbb{P}(A)$, usamos este cociente como aproximación de

$$p = \mathbb{P}(A) = \frac{k_n}{n}$$

Esta aproximación está justificada por la ley débil de los grandes números, pues en este caso

$$\mu = \mathbb{E}(\frac{S_n}{n}) = \mathbb{E}(\frac{X_1 + X_2 + \cdots + X_n}{n}) = \frac{np}{n} = p$$

por lo tanto

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - p\right| \geq \varepsilon\right) = 0$$

es decir, $\dfrac{k_n}{n}$, es una buena aproximación de la probabilidad de éxito p.

Recuerde que la varianza de una variable aleatoria binomial de parámetros N, p es $Np(1-p)$, puesto que S_n es binomial de parámetros n, p, tendremos que

$$Var(\frac{S_n}{n}) = \frac{1}{n^2}Var(S_n) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

As, por la desigualdad de Tchebichev

$$\mathbb{P}\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n}\frac{1}{\varepsilon^2}$$

Esta velocidad de convergencia, nos permite resolver problemas como el que enunciamos a continuación:

"Supongamos que p es la probabilidad desconocida de que resulte dañado un objeto producido en una línea de producción y k_n es el número de objetos que resultaron dañados, al elegir al azar una muestra de n objetos con reposición. El verdadero valor p que representa la probabilidad de que un objeto resulte dañado (ó probabilidad de éxito) es desconocido, por esta razón utilizamos el valor

$$\hat{p}_n = \frac{k_n}{n}$$

que sabemos es una aproximación del verdadero valor p. Con frecuencia nos planteamos el problema de hallar el tamaño n de la muestra para que la diferencia entre p y $\hat{p}_n$ sea tolerante. Por ejemplo, supongamos que queremos que la probabilidad de que el verdadero valor p (desconocido) difiera del valor $\hat{p}_n$ (que queremos usar para estimar p) en a lo sumo $0,01$, no sea inferior a 095, es decir, queremos que

$$\mathbb{P}\left(\left|\hat{p}_n - p\right| < 0,01\right) \geq 0,95$$

esto equivale a:

$$\mathbb{P}\left(\left|\hat{p}_n - p\right| \geq 0,01\right) \leq 1 - 0,95 = 0,05$$

Sabemos que

$$\mathbb{P}\left(\left|\hat{p}_n - p\right| \geq 0,01\right) \leq \frac{1}{(0,01)^2}\frac{p(1-p)}{n}$$

Si no tenemos ninguna información adicional sobre p, usamos la acotación para

$$p(1-p),$$

la cual es válida si $0 \leq p \leq 1$:

$$p(1-p) \leq \frac{1}{4}$$

por lo tanto

$$\mathbb{P}\left(\left|\hat{p}_n - p\right| \geq 0,01\right) \leq \frac{1}{(0,01)^2}\frac{p(1-p)}{n} \leq \frac{1}{4}\frac{10^4}{n}$$

puesto que queremos acotar por $0,05$ hallaremos n tal que

$$\frac{1}{4}\frac{10^4}{n} \leq 0,05$$

de aquí se obtiene que

$$n \geq 50000$$

es decir, el tamaño n de la muestra debe ser al menos igual a 50000.

La acotación de la desigualdad de Tchebichev no es muy fina, si pudiésemos utilizar acotaciones más precisas se disminuiría el tamaño de la muestra.

8.2 TEOREMA CENTRAL DEL LÍMITE.

Sea $X_1, X_2, \cdots, X_n, \cdots$ una sucesión de variables aleatorias independientes, con la misma distribución, de esperanza y varianza

$$\mathbb{E}(X_k) = \mu, \ \ 0 < Var(X_k) = \sigma^2 \ \ < \infty \ \ , \ k = 1, 2, 3, \cdots$$

Sea

$$S_n = X_1 + X_2 + \cdots + X_n$$

la esperanza y la varianza de S_n es:

$$\mathbb{E}(S_n) = n\mu, \ \ Var(S_n) = n\sigma^2$$

luego, la variable aleatoria

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

tiene esperanza cero y varianza uno.

El teorema central del límite nos dice que la distribución de

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

tiende cuando $n \to \infty$ a la distribución normal estándar.

DEMOSTRACIÓN:

Denotemos por

$$\varphi_n(t) = \mathbb{E}(e^{itZ_n})$$

la función característica ó transformada de Fourier de la variable aleatoria Z_n.

Si Z es una variable aleatoria y

$$\varphi(t) = \mathbb{E}(e^{itZ})$$

es su función característica, se cumple:

1. $\varphi(0) = 1$
2. $\varphi'(0) = i\mathbb{E}(Z)$
3. $\varphi''(0) = i^2\mathbb{E}(Z^2) = -\mathbb{E}(Z^2)$
   (las demostraciones de estas tres propiedades, son similares a las de la transformada de Laplace)
4. Si Z es normal estándar
   $$\varphi(t) = e^{-\frac{t^2}{2}}$$
5. Si
   $$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k$$
   donde las variables aleatorias Y_k, se definen como
   $$Y_k = \frac{X_k - \mu}{\sigma}$$
   son independientes, idénticamente distribuidas, de esperanza cero y varianza uno.
6. La transformada de Fourier ó función característica de Z_n es
   $$\varphi_n(t) = \mathbb{E}(e^{itZ_n}) = \mathbb{E}(e^{it\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k}) =$$
   $$\mathbb{E}(\prod_{k=1}^{n} e^{it\frac{1}{\sqrt{n}}Y_k}) = \prod_{k=1}^{n}\mathbb{E}(e^{it\frac{1}{\sqrt{n}}Y_k}) = \prod_{k=1}^{n}\varphi_{Y_k}(\tfrac{t}{\sqrt{n}}) = \Big(\varphi(\tfrac{t}{\sqrt{n}})\Big)^n$$
   donde $\varphi(\frac{t}{\sqrt{n}})$ es la transformada de Fourier común a todas las variables Y_k evaluadas en $\dfrac{t}{\sqrt{n}}$.

7. El segundo polinomio de Taylor de la función característica de una variable aleatoria, alrededor de cero, puede expresarse como
$$\varphi(s)=\varphi(0)+\varphi'(0)s+\varphi''(0)\frac{s^2}{2}+o(s^2)$$
donde
$$\lim_{s\to 0} o(s^2)=0.$$
En nuestro caso
$$\varphi(0)=1,\quad \varphi'(0)=0,\quad \varphi''(0)=-1$$
por lo tanto
$$\varphi(s)=1-\frac{s^2}{2}+o(s^2)$$
de donde
$$\varphi(\frac{t}{\sqrt{n}})=1-\frac{t^2}{2n}+o(\frac{t^2}{n})$$

8. Queremos demostrar que
$$\lim_{n\to\infty}=\varphi_n(t)=\lim_{n\to\infty}\left(\varphi(\frac{t}{\sqrt{n}})\right)^n=e^{-\frac{t^2}{2}}$$
lo cual resulta de
$$\lim_{n\to\infty}\left(\varphi(\frac{t}{\sqrt{n}})\right)^n=\lim_{n\to\infty}\left(1-\frac{t^2}{2n}+o(\frac{t^2}{n})\right)^n=\lim_{n\to\infty}\left(1-\frac{t^2}{2n}\right)^n=\exp\left(-\frac{t^2}{2}\right)$$
De aquí se deduce que la distribución límite de Z_n es $N(0,1)$ ya que su transformada de Fourier converge para todo t, cuando $n\to\infty$, a la transformada de Fourier de la distribución Normal Estándar, la cual es $\exp\left(-\frac{t^2}{2}\right)$.

■

COROLARIO: TEOREMA DE MOIVRE-LAPLACE:

Si $X_1, X_2, \cdots, X_n, \cdots$ una sucesión de variables aleatorias independientes, de Bernoulli de parámetro p,
$$S_n=X_1+X_2+\cdots+X_n$$
tiene distribución Binomial de parámetros $p, n;$ se tiene que
$$\mathbb{E}(S_n)=np,\quad Var(S_n)=np(1-p)$$
aplicando el resultado del Teorme Central del Límite, demostrado anteriormente, concluimos que
$$\frac{S_n-np}{\sqrt{np(1-p)}}$$
converge en distribución a la normal estándar, es decir
$$\lim_{n\to\infty}\mathbb{P}\left(a<\frac{S_n-np}{\sqrt{np(1-p)}}\le b\right)=\Phi(b)-\Phi(a)=\int_a^b\phi(t)dt$$
donde
$$\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad x\in\mathbb{R}$$

■

TEOREMA CENTRAL DEL LÍMITE. ENUNCIADO DE LA CONDICIÓN DE LINDEBERG.

Sea $X_1, X_2, \cdots, X_n, \cdots$ una sucesión de variables aleatorias independientes, con

$$\mathbb{E}(X_n) = 0, \quad 0 < \sigma_n^2 = Var(X_n) = \mathbb{E}\left(X_n^2\right) < \infty, \quad n = 1, 2, \cdots$$

si

$$S_n = X_1 + X_2 + \cdots + X_n$$

obtenemos que

$$\mathbb{E}(S_n) = 0, \quad s_n^2 = Var(S_n) = \sum_{i=1}^{n} \sigma_i^2.$$

Diremos que se cumple la condición de Lindeberg si para todo $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^{n} \mathbb{E}\left(X_i^2 \mathcal{X}_{A_{i,n}}\right) = 0$$

donde $\mathcal{X}_{A_{i,n}}$ es la función indicatriz del evento

$$A_{i,n} = \{|X_i| > \varepsilon s_n\}.$$

Bajo esta condición se tiene que

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{s_n} \leq x\right) = \int_{-\infty}^{x} \phi(t)dt$$

donde

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Este enunciado generaliza el dado anteriormente.

La condición $\mathbb{E}(X_n) = 0$, no es restrictiva, pues en caso contrario

$$X_n' = X_n - \mathbb{E}(X_n)$$

cumple las condiciones:

$$\mathbb{E}(X_n') = 0, \quad Var(X_n') = Var(X_n)$$

Para la demostración se puede consultar: M. Loéve (Probablity Theory), W. Feller (An Introducction to Probability Theory, vol.2), P. Billingsley (Convergence and Probability Measures).

Considerando, en particular, $X_1, X_2, \cdots, X_n, \cdots$ variables aleatorias independientes, con la misma distribución y

$$\begin{aligned} X_n' &= X_n - \mathbb{E}(X_n), \ \mathbb{E}(X_n') = 0, \\ 0 &< \sigma^2 = Var(X_n') = \mathbb{E}\left(X_n'^2\right) < \infty, \quad n = 1, 2, \cdots \end{aligned}$$

es fácil verificar que la sucesión $X_1', X_2', \cdots, X_n'$ satisface la condición de Lindeberg, por lo tanto obtenemos el Teorema Central del Límite demostrado anteriormente, para este caso.

EJEMPLOS:

1. Se arroja un dado 6000 veces. Aproximar mediante la distribución normal la probabilidad de obtener el número 6 entre 990 y 1010 veces.

   Si $X =$ "el n^o de veces que obtenemos 6 en 6000 lanzamientos", sabemos que la distribución de X es Binomial de parámetros $N = 6000$, $p = \frac{1}{6}$. Podemos aproximar la probabilidad de este suceso, usando el teorem Central del Límite (Moivre-Laplace):

$$\mathbb{P}(990 \leq X \leq 1010) = \mathbb{P}\left(\frac{990-1000}{\sqrt{6000\frac{1\cdot 5}{6\cdot 6}}} \leq \frac{X-1000}{\sqrt{6000\frac{1\cdot 5}{6\cdot 6}}} \leq \frac{1010-1000}{\sqrt{6000\frac{1\cdot 5}{6\cdot 6}}}\right) \simeq$$
$$\frac{1}{\sqrt{2\pi}} \int_{-0,346}^{0,346} e^{-\frac{t^2}{2}} dt = \Phi(0,346) - \Phi(-0,346) = 2\Phi(0,346) - 1 \simeq 0,27.$$

2. Un antropólogo desea estimar la altura media de los hombres de cierto grupo étnico (se supone que la altura se distribuye normalmente). Si la desviación estándar es $\sigma = 2,5$ pulgadas y se escogen al azar 100 hombres, calcule la probabilidad de que la diferencia entre la media de la muestra y la media real de la población no exceda $0,5$ pulgadas. (Se supone que cada individuo se elige de manera independiente).

   Solución: si μ es la media real de la población y
$$\frac{S_n}{n}$$
   es la media de la muestra, lo que se quiere calcular es
$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \leq 0,5\right)$$
   Ahora bien, S_n tiene distribución $N(n\mu, \sigma_n)$, con
$$\sigma_n^2 = n\sigma^2$$
   La variable $\dfrac{S_n}{n}$ tiene distribución $N(\mu, \dfrac{\sigma}{\sqrt{n}})$, por lo tanto la variable
$$\frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$
   tiene distribución Normal Estándar, por lo tanto
$$\begin{aligned}&\mathbb{P}\left(\left|\tfrac{S_n}{n} - \mu\right| \leq 0,5\right) = \mathbb{P}\left(-0,5 \leq \tfrac{S_n}{n} - \mu \leq 0,5\right) = \\ &\mathbb{P}\left(-\tfrac{0,5}{\frac{2,5}{\sqrt{100}}} \leq \tfrac{S_n - n\mu}{\sigma\sqrt{n}} \leq \tfrac{0,5}{\frac{2,5}{\sqrt{100}}}\right) = \mathbb{P}\left(-2 \leq \tfrac{S_n - n\mu}{\sigma\sqrt{n}} \leq 2\right) \simeq 0,8544\end{aligned}$$
   (En este ejemplo, se supone que la distribución es exactamente normal estándar, si no conocemos la distribución, nos valemos del teorema central del límite para aproximar
$$\frac{S_n - n\mu}{\sigma\sqrt{n}}$$
   por la distribución normal estándar.

3. Si en el ejercicio anterior se quiere que la diferencia entre la media de la muestra y la media de la población sea menor que $0,4$ pulgadas con probabilidad $0,95$; de qué tamaño debe ser la muestra?

   Solución: en el caso anterior tomamos $n = 100$, en este caso el tamaño de la muestra es desconocido. Queremos hallar n tal que
$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \leq 0,4\right) = 0,95$$
   pero
$$\begin{aligned}&\mathbb{P}\left(\left|\tfrac{S_n}{n} - \mu\right| \leq 0,4\right) = \mathbb{P}\left(-\tfrac{0,8\sqrt{n}}{5} \leq \tfrac{S_n - n\mu}{\sigma\sqrt{n}} \leq \tfrac{0,8\sqrt{n}}{5}\right) = \\ &= 2\Phi\left(0,16\sqrt{n}\right) - 1 = 0,95\end{aligned}$$
   despejando, obtenemos:
$$\Phi\left(0,16\sqrt{n}\right) = 0,975$$
   en la tabla hallamos que
$$z_0 = 0,16\sqrt{n} = 1,96$$
   de donde
$$\sqrt{n} = 12,25, \quad n \simeq 151$$

Tarea:

1. Demostrar que si X es una variable aleatoria, $X \geq 0$, con esperanza finita entonces

$$\mathbb{P}\left(X \geq a\right) \leq \frac{1}{a}\mathbb{E}\left(X\right)$$

(Esta desigualdad se conoce con el nombre de Desigualdad de Markov). Sugerencia:

$$X \geq X \cdot \mathcal{X}_A \geq a \cdot \mathcal{X}_A$$

donde el evento $A = [X \geq a]$.

2. Si X es una variable aleatoria de Distribución Binomial de parámetros N y $p = \frac{1}{2}$ entonces

$$\mathbb{P}\left(X - Np \geq j\right) = \mathbb{P}\left(X - Np \leq -j\right)$$

deducir de aquí que

$$\mathbb{P}\left(X - Np \geq j\right) = \frac{1}{2}\mathbb{P}\left(|X - Np| \geq j\right)$$

(Sugerencia: tratar los casos N par y N impar por separado y utilizar que $\binom{N}{k} = \binom{N}{N-k}$)

3. Demostrar que si X es $N(0,1)$

$$\mathbb{P}\left(X \geq a\right) = 1 - \Phi(a) \leq \frac{1}{\sqrt{2\pi}a}e^{-\frac{a^2}{2}}$$

y que

$$\lim_{a\to\infty} \frac{1-\Phi(a)}{\frac{1}{\sqrt{2\pi}a}e^{-\frac{a^2}{2}}} = 1$$

(Este último resultado nos dice que

$$\frac{1}{\sqrt{2\pi}a}e^{-\frac{a^2}{2}}$$

es asintóticamente equivalente a $1 - \Phi(a)$, por lo tanto si a es grande podemos valernos de esta cota para obtener una buena aproximación de $1 - \Phi(a)$).

4. Se lanza una moneda 5000 veces, si suponemos que la moneda no está cargada, cuál es la probabilidad de que el número de caras sea mayor que 2800 ?. (Utilizar la igualdad $\mathbb{P}\left(S_n - Np \geq j\right) = \frac{1}{2}\mathbb{P}\left(|S_n - Np| \geq j\right)$ y la desiguladad de Tchebychev, luego haga el cálculo utilizando la aproximación normal).

EJERCICIOS PARA LA PRÁCTICA:

1. Sea X una variable aleatoria con distribución normal de parámetros

$$\mu \in \mathbb{R}, \quad \sigma > 0$$

   Demuestre que

   (a) $\mathbb{E}(X) = \mu, \quad Var(X) = \sigma^2$.

   (b) $Z = \dfrac{X-\mu}{\sigma}$ tiene distribución normal estándar.

2. Una línea aérea sabe que el 5% de las personas que hacen reservaciones en un cierto vuelo, al final no se presentan. Si la aerolínea vende 160 boletos para este vuelo, y sólo hay 155 asientos en el avión, cuál es la probabilidad de que todo pasajero con reservación que se presente al aeropuerto tenga un puesto en el vuelo?. Respuesta $=\ 0,8621$.

3. En una empresa se ha observado que el gasto semanal en mantenimiento y reparaciones es una variable aleatoria con distribución aproximadamente normal de media $\mu = Bs.\ 24000$ y desviación $\sigma = Bs.1200$. Cuánto se debe presupuestar semanalmente para mantenimiento y reparaciones de modo que el monto presupuestado sea excedido con una probabilidad de a lo sumo $0,1$?

   Respuesta: $x_0 \geq 25537,8$

4. Un encuestador cree que el 20% de los votantes de una zona está a favor del candidato A. Si se escogen 24 votantes de la zona, aproxime la probabilidad de que la fracción de votantes de la muestra que favorece al candidato A, no difiera de la verdadera fracción (en toda la zona) en más de $0,06$.

   Respuesta: $0,5408$

5. Una máquina se manda a reparar si una muestra de 100 artículos escogidos al azar de su gran producción diaria, revela un 15% ó mas de defectuosos. Si la máquina en realidad sólo produce un 10% de defectuosos, calcule aproximadamente la probabilidad de que la manden a reparar.

   Respuesta: $0,0475$

6. La vida activa de un cierto fármaco sigue una distribución $N(1200, 40)$ días. Se desea enviar un lote de medicamentos, de modo tal que la vida media del lote no sea inferior a 1180 días con probabilidad $0,95$. Qué tamaño debe tener la muestra?.Respuesta: $n \simeq 11$

7. Encuentre una aproximación de la probabilidad, de que el número de veces que salga 1, esté comprendido entre 1900 y 2150 veces, al lanzar un dado perfecto 12000 veces. Respuesta: $0,99275$

8. Se toma una muestra al azar con reposición, a efectos de estimar la fracción p de hembras en una población. Encontrar un tamaño de muestra que asegure que la estimación se hará con un error de menos de $0,005$, al menos con una probabilidad de $0,99$. Respuestas: Por Tchebichev ($n \geq 10000$) y por TCL (66348)

9. Se desea estimar la probabilidad de falla p, en un proceso de producción, mediante la observación de n objetos producidos, elegidos independientemente. Se sabe que p está entre $0,1$ y $0,3$ por información previa. Halle el tamaño n de la muestra para que la probabilidad de que la frecuencia relativa de objetos fallados en la muestra, difiera del verdadero valor p en más de $0,01$ sea menor que $0,05$.

   Respuestas: Por Tchebichev ($n \geq 42000$) y por TCL ($n \geq 8068$)

10. El porcentaje de individuos daltónicos de una población es P desconocido. Se desea estimar este procentaje P a partir del porcentaje observado en una muestra de tamaño n. Calcular el tamaño que debe tener la muestra a fin de que el error cometido sea inferior al 1% con probabilidad al menos de $0,90$ en los casos:

    (a) No se sabe nada acerca de P. Respuesta $n \geq 25000$

(b) Se sabe que P es inferior al 16%. Respuesta $n \geq 13440$
(Por Tchebichev)

11. Se ha observado que las notas de un examen de admisión siguen una distribución aproximadamente normal, de media 78 y varianza 36.(Las notas están entre 1 y 100).

    (a) Si un grupo de estudiantes va apresentar dicho examen, qué porcentaje de ellos espera Ud. que obtenga notas entre 70 y 90?. Respuesta: 88,54%

    (b) Cuál es la probabilidad de que una persona que tome el examen obtenga más de 72? Resp.0,8413

    (c) Suponga que los estudiantes cuyas notas se encuentran en el 10% superior de la distribución serán admitidos inmediatamente. Cuál debe ser la nota mínima que debe tener un estudiante para ser admitido inmediatamente?

        Respuesta:85,689

12. La duración de un tipo de bombillos sigue una distribución Normal de media $\mu = 1000$ horas y desviación $\sigma = 100$ horas.

    Se desea enviar una muestra de bombillos de manera que la duración media de la muestra no difiera de μ en más de 50 horas con una probabilidad de $0,95$.

    (a) Hallar el tamaño que debe tener la muestra. Resp. $n \geq 16$

    (b) Resuelva el problema, si se desconoce la distribución :

        i. Usando Tchebichev. Resp. $n \geq 80$

        ii. Usando el Teorema Central del Límite. Resp. $n \geq 16$.

13. Una compañía tiene 90 ejecutivos. Supongamos que la probabilidad de que un ejecutivo necesite una secretaria al comenzar su día de trabajo es $\frac{1}{10}$. Si queremos que con un 95% de certeza haya una secretaria disponible para cada ejecutivo que la solicite, cuántas secretarias deben contratarse para un centro secretarial que sirva al grupo de 90 ejecutivos?

14. Los médicos opinan que los niveles altos de colesterol en la sangre están asociados a una mayor incidencia de enfermedades cardíacas.

    Suponga que el logaritmo (ln) del nivel de colesterol, para los hombres en un cierto grupo de edad, es una variable aleatoria con distribución normal de media $2,35$ y desviación $1,2$

    (a) Si un nivel de colesterol en la sangre dentro del rango $150 - 250\, mg/ml$ se considera normal, qué porcentaje de la población antes mencionada esperaría Ud. tener en este rango? Qué porcentaje estará por encima de 250?

    (b) Los niveles superiores a $300\, mg/ml$ se consideran de alto riesgo, qué porcentaje de la población antes mencionada esperaría Ud. tener en este rango?

15. Un fabricante de cereales afirma que el peso medio de una caja del cereal que vende es de $330,4\, grs.$ con una desviación de $21\, grs$. Se desea verificar si su afirmación es cierta. Para esto se va a elegir una muestra aleatoria de cajas del cereal y calcular el peso promedio de la muestra. Cuántas cajas debemos tener en la muestra para que el peso promedio se encuentre a menos de $7\, grs.$ de la verdadera media con una probabilidad de $0,99$? (Suponga que la distribución del peso de cada caja es normal). Respuesta: $n \geq 60$

Bibliografía

[1] Obregón Sanin Iván. *Teoría de la Probabilidad* Ed. Limusa (1975)

[2] Meyer *Probabilidad y aplicaciones Estadísticas* Fondo Educativo Interamericano C.A. (1970).

[3] Feller W. *Introducción a la Teoría de Probabilidad y sus aplicaciones.* Vol. I Ed. Limusa (1975)

[4] Maravall Dario *Filosofía de las Matemáticas* Ed. Dossat S.A. (1961)