Probabilidad
Licienciado en Matemáticas Aplicadas
Apuntes de Clase

Paul Ramírez De la Cruz

2 de mayo de 2007

Índice general

1. PROBABILIDAD

1.1. Espacios de probabilidad

- Consideremos el lanzamiento de una moneda para observar la cara que queda hacia arriba. ¿Podemos afirmar con certeza cuál será dicha cara?
- Pensemos en que elegimos a un alumno en el campus para preguntarle cuántos libros ha solicitado a préstamo en la biblioteca universitaria durante el último mes. ¿Podemos asegurar cuál será su respuesta?
- Supongamos que en una prueba de control de calidad se examina un componente electrónico haciéndolo funcionar de manera ininterrumpida hasta que falla, y entonces registramos el tiempo transcurrido desde el inicio de la prueba. ¿Podríamos garantizar que su duración será exactamente un cierto valor?

Los ejemplos anteriores se refieren a situaciones en las que resulta necesario trabajar con el concepto de probabilidad. Antes de continuar, establezcamos de manera formal algunos conceptos:

Definición 1 (Experimento) *Es cualquier procedimiento mediante el cual obtenemos una observación.*

En particular, para el estudio de la probabilidad nos interesa observar aquellos experimentos cuyo resultado no es pronosticable con certeza, es decir, aquellos en que existe aleatoriedad.

Definición 2 (Espacio Muestral) *El el conjunto de todos los resultados individuales que puede tener un experimento. A cada uno de los elementos del espacio muestral se le llama punto muestral.*

Definición 3 (Evento) *Es cualquier resultado posible al realizar un experimento.*

Si un evento tiene un único elemento, se le llama evento simple. En general, a cualquier subconjunto del espacio muestral se le llama evento compuesto.

Notación 4 *Sea Ω el espacio muestral de un experimento R. Si Ω es a lo más numerable, denotaremos a los eventos simples de R por*

$$E_i,\ i \in \{1, ..., |\Omega|\}.$$

Ejemplo 5 *Al lanzar un dado de seis caras legal (que no esté cargado), observamos que puede ocurrir que la cara que se muestra hacia arriba tenga un punto, o dos puntos, ..., o seis puntos. Generalmente asignamos valores numéricos a los resultados individuales, así que el conjunto de resultados posibles del experimento de lanzar un dado, es decir, su espacio muestral, es* $\Omega = \{x \mid x \in \mathbb{N},\ x \leq 6\}$.

Entonces decimos que $E_i = \{i\}$, $i \in \{1, ..., 6\}$ *son los eventos simples del experimento* R.

Ejemplo 6 *Sea* R *el experimento de contar el número de cabellos en la cabeza de una persona. Entonces podemos decir que* $\Omega = \{0, 1, ...\}$ *y* $E_i = \{i\}$, $i \in \{0, 1, ...\}$, *si bien sabemos que existirá un* $M \in \mathbb{N}$ *tal que sería imposible encontrar a una persona con* M *cabellos o más.*

Ejemplo 7 *Supongamos que medimos la resistencia eléctrica de un componente electrónico, entonces* $\Omega = \{x \mid x \in \mathbb{R}^+\}$ *y algunos ejemplos de eventos compuestos son:*

$$\begin{aligned} A &= \{x \mid x > a\},\ a \text{ constante} \\ B &= \{x \mid a < x \leq b\},\ a, b \text{ constantes} \\ C &= \{x \mid x = 2k - 1\},\ k \in \mathbb{N}. \end{aligned}$$

Ejercicio 8 *Considere el experimento de lanzar una moneda y sea*

$$X = \begin{cases} 0, & \text{si cae "águila"} \\ 1, & \text{si cae "sol"} \end{cases}.$$

1. *Antes de lanzar la moneda ¿cómo podríamos medir la posibilidad de ocurrencia del evento* $A = \{X = 1\}$*?*

2. *Repita el experimento diez veces. A partir de los datos que obtuvo ¿Cómo valoraría ahora la posibilidad de que ocurra* A*?*

3. *Suponga que un especialista mide la densidad de la moneda en ambas caras y determina que, debido a la diferencia de estas, las posibilidades son de ocho águilas y dos soles en diez lanzamientos. ¿Qué piensa de dicha asignación?*

Notemos que los eventos son elementos del conjunto potencia de Ω, 2^Ω. Generalmente, nos concentraremos en un subconjunto de 2^Ω con ciertas características deseable. Consideremos entonces la siguiente:

Definición 9 ($\sigma-$álgebra) *Sea* Ω *un conjunto. Una* $\sigma-$*álgebra (léase sigma-álgebra)* $\mathcal{A}$ *sobre* Ω *es un subconjunto no vacío de* 2^Ω *que cumple:*

1. $\Omega \in \mathcal{A}$

2. *Si* $E \in \mathcal{A}$, *entonces* $E^c \in \mathcal{A}$.

3. *Si* $\{E_i\}_{i\in\mathbb{N}}$, *es una sucesión numerable de eventos en* $\mathcal{A}$, *entonces*

$$\bigcup_{i=1}^{\infty} E_i \in \mathcal{A}.$$

Proposición 10 *Sean* Ω *un conjunto no vacío y* $\mathcal{A}$ *una sigma-álgebra en* Ω. *Sea* $\{A_i\}_{i=1}^{\infty}$ *una sucesión de subconjuntos de* Ω. *Entonces* $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$.

Ejercicio 11 *Sea* Ω *un conjunto. Muestre que*

1. $S = \{\Omega, \emptyset\}$ *es una* $\sigma-$*álgebra. A* S *se le denomina* $\sigma-$*álgebra trivial.*

2. 2^{Ω} *es una* $\sigma-$*álgebra.*

Proposición 12 *Sea* A *un subconjunto propio no vacío de* Ω, *y sea* $F = \{\emptyset, \Omega, A, A^c\}$. *Entonces* F *es la* $\sigma-$*álgebra más pequeña que contiene a* A.
Demostración. *Notamos que* $\emptyset, \Omega \in F$. *Sea* E *cualquier elemento de* F, *entonces observamos que* $E^c \in F$. *Si consideramos cualquier sucesión numerable de eventos en* F, *notamos que el conjunto que resulta de la unión de dicha sucesión es:*

$$\begin{cases} \emptyset, \ o \\ A, \ o \\ A^c, \ u \\ \Omega \end{cases}$$

luego cualquier sucesión numerable de eventos en F *está en* F. *Por tanto,* F *es una sigma-álgebra.*

Ahora, sea G *cualquier* $\sigma-$*álgebra que contiene a* A, *entonces* $\emptyset, \Omega, A, A^c \in G$, *por tanto* $F \subset G$.

En consecuencia, F *es una* $\sigma-$*álgebra y es la más pequeña de todas las* $\sigma-$*álgebras que contienen a* A. ■

Si $A_1, A_2, ..., A_n$ es una colección arbitraria de subconjuntos de Ω, puede determinarse la más pequeña $\sigma-$álgebra que la contiene, como veremos enseguida.

Proposición 13 (La intersección de $\sigma-$álgebras es $\sigma-$álgebra) *Sea* Ω *un conjunto no vacío dado arbitrariamente y sea* F_j, $j \in J$, $J \neq \emptyset$ *una familia de* $\sigma-$*álgebras en* Ω, *donde* J *es un conjunto de índices cualquiera. Entonces*

$$F := \bigcap_{j\in J} F_j$$

es una $\sigma-$*álgebra.*

Demostración. *Sea* $F = \bigcap_{j \in J} F_j$. *Notemos que* $\emptyset, \Omega \in F_j$ *para todo* $j \in J$, *de modo que* $\emptyset, \Omega \in \bigcap_{j \in J} F_j$.

Ahora, sea $A \in F$, *entonces* $A \in F_j$ *para todo* $j \in J$. *Luego, como cada* F_j *es una* $\sigma-$*álgebra,* $A^c \in F_j$, *y en consecuencia,* $A^c \in F$.

Finalmente, sea $\{A_i\}_{i \in \mathbb{N}}$ *una sucesión de eventos en* Ω *que está contenida en cada* F_j, $j \in J$, *entonces está en* $\bigcap_{j \in J} F_j$. *Como la sucesión está en* F_j *y cada* F_j *es una* $\sigma-$*álgebra, entonces*

$$\bigcup_{i \in \mathbb{N}} A_i \in F_j, \ \forall j \in J$$

por lo cual

$$\bigcup_{i \in \mathbb{N}} A_i \in \bigcap_{j \in J} F_j.$$

Debido a lo anterior, tenemos que $\bigcap_{j \in J} F_j$ *es una* $\sigma-$*álgebra.* ■

Proposición 14 *Sea* Ω *cualquier conjunto no vacío y sea* $\mathcal{G}$ *un sistema cualquiera de subconjuntos de* Ω. *Entonces existe la mínima* $\sigma-$*álgebra* $\sigma(\mathcal{G})$ *sobre* Ω *tal que*

$$G \subset \sigma(\mathcal{G}).$$

Demostración. *Sea*

$$J = \{\mathcal{C} \text{ es una } \sigma - \text{álgebra sobre } \Omega \text{ tal que } \mathcal{G} \subset \mathcal{C}\}.$$

Sabemos que J *es no vacío porque* $\mathcal{G} \subset 2^{\Omega}$ *y* 2^{Ω} *es una* $\sigma-$*álgebra. Luego*

$$\sigma(\mathcal{G}) := \bigcap_{C \in J} C$$

es una $\sigma-$*álgebra de acuerdo con la proposición anterior.*

Ahora, sea $\mathcal{F}$ *cualquier* $\sigma-$*álgebra tal que* $\mathcal{G} \subset \mathcal{F}$, *entonces por la forma en que definimos a* $\sigma(\mathcal{G})$, *tenemos que* $\sigma(\mathcal{G}) \subset \mathcal{F}$. ■

Notación 15 *Si* $\mathcal{G}$ *es una clase de subconjuntos de* Ω, *a la mínima* $\sigma-$*álgebra que contiene los conjuntos de* $\mathcal{G}$ *se le denota por*

$$\sigma(\mathcal{G})$$

y se le llama la mínima $\sigma-$*álgebra sobre* Ω. *También se le llama la* $\sigma-$*álgebra generada por* $\mathcal{G}$.

Proposición 16 *Sean* $a, b \in \mathbb{R}^{\#},\ a < b,$ *entonces*

$$(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n} \right)$$

y

$$[a, b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b \right).$$

Definición 17 (Reales extendidos) *La colección* $\mathbb{R}^{\#} = \mathbb{R} \cup \{-\infty, \infty\}$ *se denomina sistema de los números reales extendidos.*

Proposición 18 *Sean*[1]*:*

$\mathcal{G}_0$ *el sistema de todos los subconjuntos abiertos de* $\mathbb{R}$,
$\mathcal{G}_1$ *el sistema de todos los subconjuntos cerrados de* $\mathbb{R}$,
$\mathcal{G}_2$ *el sistema de todos los intervalos* $(-\infty, b]\,, b \in \mathbb{R}$,
$\mathcal{G}_3$ *el sistema de todos los intervalos* $(-\infty, b)\,, b \in \mathbb{R}$,
$\mathcal{G}_4$ *el sistema de todos los intervalos* $(a, b]\,, -\infty < a < b < \infty$,
$\mathcal{G}_5$ *el sistema de todos los intervalos* $(a, b)\,, -\infty < a < b < \infty$,
Entonces $\sigma(\mathcal{G}_0) = \sigma(\mathcal{G}_1) = \sigma(\mathcal{G}_2) = \sigma(\mathcal{G}_3) = \sigma(\mathcal{G}_4) = \sigma(\mathcal{G}_5)$.

Definición 19 ($\sigma-$álgebra de Borel) *A la* $\sigma-$*álgebra presentada en la proposición anterior se le llama* $\sigma-$*álgebra de Borel y se le denota por* $\mathcal{B}(\mathbb{R})$.

Definición 20 (Medida) *Una medida sobre una* $\sigma-$*álgebra* $\mathcal{A}$ *es una función real extendida no negativa* μ *en* $\mathcal{A}$ *tal que si* $\{A_n\}_{n=1}^{\infty}$ *es una sucesión de conjuntos ajenos dos a dos en* $\mathcal{A}$, *entonces*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Definición 21 (Medida de probabilidad) *Si* μ *es una medida tal que* $\mu(\Omega) = 1$, *entonces a* μ *se le llama* ***medida de probabilidad****, o simplemente* ***probabilidad****.*

Notación 22 *Si tenemos una medida de probabilidad, se le suele representar con* $P(\cdot)$.

Ejemplo 23 (Medida de longitud) *Sea* A *un intervalo en* $\mathbb{R}$ *con extremos* $a, b,\ a < b$, *y definamos* $\mu(A) = b - a$. *Entonces* μ *es una medida en la* $\sigma-$*álgebra de Borel llamada "medida de longitud".*

Ejemplo 24 (Medida de conteo) *Sea* Ω *cualquier conjunto y sea* 2^{Ω} *su conjunto potencia. Definamos* $\mu(A)$ *como el número de puntos en* A, *de modo que si* A *tiene* n *elementos,* $n = 0, 1, 2, ...,$ *entonces* $\mu(A) = n$, *mientras que si* A *es no numerable entonces* $\mu(A) = \infty$. *Entonces la función* μ *es una medida sobre* 2^{Ω} *llamada "medida de conteo".*

[1] Puede verse la demostración en **[Geiss and Geiss, 2006]**

Ejemplo 25 (Medida de probabilidad de conteo generalizada) *Sea* $\Omega = \{x_1, x_2, ...\}$ *un conjunto a lo más numerable y para cada* $i \in \mathbb{N}$ *sea* p_i *la medida del evento* $\{x_i\}$, *donde se cumple que* $\sum_{i=1}^{\infty} p_i = 1$.

Para cada subconjunto A *de* Ω, *definamos*

$$P(A) = \sum_{x_i \in A} p_i.$$

Entonces a P *se le llama la medida de probabilidad de conteo generalizada.*

Definición 26 (Espacio de probabilidad) *Es una tripleta* (Ω, F, P), *donde* Ω *es un conjunto,* F *es una* $\sigma-$*álgebra sobre* Ω *y* P *es una medida de probabilidad.*

Teorema 27 (Propiedades de la probabilidad) *Sea* (Ω, F, P) *un espacio de probabilidad. Entonces:*

1. $P(\emptyset) = 0$.
2. $\forall A, B \in F : P(A \cup B) + P(A \cap B) = P(A) + P(B)$.
3. *Si* $A, B \in F$ *y* $B \subset A$, *entonces* $P(A) = P(B) + P(A - B)$.
4. $\forall A_1, A_2, ... \in F$ *se cumple que* $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$.

Demostración. 1. Sea A un evento en Ω. Entonces

$$\begin{aligned} P(A) &= P(A \cup \emptyset) = P(A) + P(\emptyset) \\ &\Rightarrow P(\emptyset) = 0. \end{aligned}$$

2. Notemos que $A = (A \cap B) \cup (A - B)$, luego $P(A) = P(A \cap B) + P(A - B)$. Similarmente, $P(B) = P(A \cap B) + P(B - A)$.

Si sumamos las dos ecuaciones anteriores, tenemos que

$$\begin{aligned} P(A) + P(B) &= P(A \cap B) + [P(A - B) + P(A \cap B) + P(B - A)] \\ &= P(A \cap B) + P(A \cup B). \end{aligned}$$

3. Podemos escribir $A = B \cup (A - B)$, luego $P(A) = P(B) + P(A - B)$.

4. Notemos que $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \left(A_1^c \cap ... \cap A_{n-1}^c \cap A_n\right)$ y los conjuntos en el lado derecho de la igualdad son ajenos y por el punto 3 anterior, $P\left(A_1^c \cap ... \cap A_{n-1}^c \cap A_n\right) \leq P(A_n)$ para cada n, entonces se cumple que $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$. ■

Definición 28 (Aditividad finita y numerable) *Sean* Ω *un conjunto no vacío y* $\mathcal{A}$ *una sigma-álgebra en* Ω, $\mu : \mathcal{A} \to [-\infty, \infty]$ *una función y* $\{A_n\}_{n=1}^{\infty}$ *una sucesión de subconjuntos de* $\mathcal{A}$, *ajenos dos a dos. Se dice que*

1. μ *posee aditividad finita si para* $k \in \mathbb{N}$
$$\mu\left(\bigcup_{n=1}^{k} A_n\right) = \sum_{n=1}^{k} \mu\left(A_n\right).$$

2. μ *posee aditividad numerable si*
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right).$$

Observación 29 *Cualquier función de probabilidad es aditiva numerable.*

Teorema 30 (Propiedad de Continuidad de la Probabilidad) [2] *Sea* P *una función de probabilidad definida en una* $\sigma-$*álgebra* $\mathcal{A}$.

1. *Si* $\{A_n\}_{n=1}^{\infty}$ *es una sucesión de subconjuntos en* $\mathcal{A}$, *tales que* $A_1 \subset A_2 \subset A_3 \subset ...$ *y* $A_n \uparrow A$, *entonces*
$$\lim_{n\to\infty} P\left(A_n\right) = P\left(A\right).$$

2. *Si* $\{A_n\}_{n=1}^{\infty}$ *es una sucesión de subconjuntos en* $\mathcal{A}$, *tales que* $A_1 \supset A_2 \supset A_3 \supset ...$, $\mu\left(A_1\right) < \infty$ *y* $A_n \downarrow A$, *entonces*
$$\lim_{n\to\infty} P\left(A_n\right) = P\left(A\right).$$

Demostración.

1. *Dado que los* A_n *forman una sucesión creciente, podemos expresarlos como*
$$A = A_1 \cup \left(A_2 - A_1\right) \cup \cdots \cup \left(A_n - A_{n-1}\right) \cup \cdots$$
Entonces por el inciso [3] *del Teorema* [27] *tenemos que*
$$\begin{aligned} P\left(A\right) &= P\left(A_1\right) + P\left(A_2\right) - P\left(A_1\right) + P\left(A_3\right) - P\left(A_2\right) + \cdots P\left(A_n\right) - P\left(A_{n-1}\right) + \cdots \\ &= \lim_{n\to\infty} P\left(A_n\right). \end{aligned}$$

2. *Similar.*

■

[2] Tomada de [**Ash, 2000**]

Definición 31 (Ocurrencia de eventos) *Sea Ω el espacio muestral de un experimento y sea A un evento en Ω. Al realizar el experimento se observa un resultado, digamos E. Se dice que ocurrió A si y sólo si $E \subset A$.*

Ejemplo 32 *Consideremos el experimento de lanzar un dado legal. Sea $A = \{\text{Cae un número par}\} = \{2, 4, 6\}$. Supongamos que el resultado fue $\{4\}$, entonces como $\{4\} \subset A$, decimos que ocurrió A. Por el contrario, si el resultado hubiera sido, digamos, $\{5\}$, como $\{5\} \not\subset A$, decimos que no ocurrió A.*

Ejemplo 33 *Sea el experimento de medir la temperatura ambiente a medio día en la universidad y definamos $B = \{\text{Hay más de } 25°C\} = \{x \mid x \geq 25\}$. Si se observa el resultado $E = \{23.5°C\}$, decimos que no ocurrió B, porque $E \not\subset B$. Por otro lado, si se observara por ejemplo que el resultado fue $\{32.1°C\}$, entonces decimos que ocurrió B porque $E \subset B$.*

Ejercicio 34 *Resuelva lo siguiente:*

1. *Muestre que $A \subset B$ si y sólo si la ocurrencia de A implica la de B.*
2. *Muestre que $A \cup B$ es el evento que ocurre si y sólo si ocurre A u ocurre B.*
3. *Muestre que $A \cap B$ es el evento que ocurre si y sólo si ocurre A y ocurre B.*

Ejercicio 35 *Para cualesquiera tres eventos A, B y C, muestre que $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.*

Ejercicio 36 (Desigualdad de Bonferroni) *Demuestre que si $A_1, A_2, ..., A_k$ son eventos, entonces*

$$P\left(\bigcap_{i=1}^{k} A_i\right) \geq 1 - \sum P\left(A_i^c\right).$$

1.1.1. Ejercicios suplementarios

Tomados de [Bain and Engelhardt, 1992].

1. Una máquina expendedora de gomas de mascar da piezas verdes, rojas y azules. El experimento consiste en obtener una goma de mascar de la máquina

   a) Describa el espacio muestral.

   b) Liste todos los posibles eventos.

c) Si R es el evento de que la goma obtenida es roja, liste los resultados en R^c.

d) Si V es el evento de que la goma obtenida es verde, entonces ¿cuál es el evento $R \cap V$?

2. Un experimento consiste en sacar gomas de mascar de la máquina del ejercicio anterior hasta que se obtiene una azul. Describa el espacio muestral para el experimento.

3. Suponga que se examina una muestra de material radiactivo. Proporcione el espacio muestral del experimento si:

   a) Se mide el número de partículas alfa que emite la muestra en un intervalo fijo de tiempo.

   b) Se mide el tiempo que transcurre hasta que se emite la primera partícula alfa.

4. Se conduce un experimento en el que se mide la proporción de oro en una pieza de metal. Dé el espacio muestral.

5. Sacamos 100 gomas de mascar de una máquina y obtenemos 20 rojos (R), 30 verdes (V) y 50 azules (A). Proporcione una función que indique la probabilidad de sacar una goma de cada color.

1.2. Técnicas de conteo de puntos muestrales

Definición 37 (Principio de multiplicación) *Si una actividad consta de m etapas, y la $i-$ésima etapa se puede completar de n_i formas, $i = 1, ..., m$, entonces la actividad completa se puede completar de n formas donde*

$$n = \prod_{i=1}^{m} n_i.$$

Ejemplo 38 *Suponga que se realiza el lanzamiento de dos monedas y se observa las caras mostradas hacia arriba. Calcule el número de puntos en el espacio muestral.*

Ejemplo 39 *Se lanza una moneda y después se saca una canica de una caja que tiene canicas rojas, verdes, negras y azules. Calcule el número de puntos en el espacio muestral.*

Ejemplo 40 *Un experimento consiste en repetir r ocasiones un ensayo que tiene n resultados posibles. ¿Cuántos puntos muestrales tiene el experimento?*

1.3. Probabilidad condicional e independencia de eventos

1.3.1. Probabilidad condicional

Definición 41 (Probabilidad condicional) *Sea* $(\Omega, \mathcal{F}, P)$ *un espacio de probabilidad y sean* A *y* B *eventos en* $\mathcal{F}$ *con* $P(B) \neq 0$. *La probabilidad condicional de* A *dado* B *se representa por* $P(A \mid B)$ *y está dada por*

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

Corolario 42 *Si* A *y* B *son dos eventos y* $P(A) \neq 0, P(B) \neq 0$, *entonces*

$$\begin{aligned} P(A \cap B) &= P(A \mid B) P(B) \\ &= P(B \mid A) P(A). \end{aligned}$$

Ejemplo 43 *Suponga que se clasifica a las personas de cierta carrera que han satisfecho los requisitos para graduarse de la universidad, considerando su género y si cuentan actualmente con trabajo o no. Suponga que se elige a una persona aleatoriamente.*

	Empleado (E)	***Desempleado*** (E^c)	***Total***
Hombre	*460*	*40*	*500*
Mujer	*140*	*260*	*400*
Total	*600*	*300*	*900*

.

1. *Calcule la probabilidad de que sea hombre.*
2. *Calcule la probabilidad de que esté desempleado.*
3. *Calcule la probabilidad de que sea hombre y esté desempleado.*
4. *Calcule la probabilidad de que sea hombre dado que está desempleado.*
5. *Calcule la probabilidad de que esté desempleado dado que es hombre.*

Ejemplo 44 *La probabilidad de que un vuelo de programación regular despeque a tiempo es 0.83, la probabilidad de que llegue a tiempo es 0.82, y la probabilidad de que despegue y llegue a tiempo es 0.78. Encuentre la probabilidad de que un avión:*

1. *Llegue a tiempo dado que despegó a tiempo.*
2. *Haya despegado a tiempo dado que llegó a tiempo.*

Ejemplo 45 *La probabilidad de que un hombre casado vea cierto programa de televisión es de 0.4 y de que una mujer casada lo vea, es 0.5. La probabilidad de que un hombre vea el programa dado que su esposa lo ve es 0.7. Encuentre la probabilidad de que:*

1. *Una pareja de casados vea el programa.*

2. *Una esposa vea el programa dado que su esposo lo hace.*

3. *Al menos uno de los integrantes del matrimonio ve el programa.*

Ejemplo 46 *Entre 100 estudiantes que se hallan en el aula 50 conocen el inglés, 40 el francés y 35 el alemán. 20 estudiantes conocen el inglés y el francés, 8 el inglés y el alemán, 10 el francés y el alemán. 5 estudiantes dominan todos los tres idiomas. Un estudiante sale del aula. Calcule la probabilidad de que el estudiante que salió:*

1. *Hable inglés, dado que habla francés.*

2. *Hable francés, dado que habla inglés.*

3. *No hable alemán, dado que habla francés.*

4. *No hable inglés, dado que habla francés.*

5. *Hable inglés, dado que habla francés y alemán.*

1.3.2. Independencia de eventos

De manera intuitiva, decimos que dos eventos son independientes entre sí, si la ocurrencia de uno de ellos no cambia la probabilidad de que ocurra el otro, y viceversa. Formalmente, podemos expresar esta idea en la siguiente

Definición 47 (Eventos independientes) *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sean A y B eventos en $\mathcal{F}$. Decimos que A y B son independientes si y sólo si*

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

Corolario 48 *Si $P(A \mid B) = P(A)$, entonces $P(B \mid A) = P(B)$.*

Corolario 49 *Dos eventos, A y B son independientes si y sólo si*

$$P(A \mid B) = P(A)\, P(B).$$

Teorema 50 *Si en un experimento pueden ocurrir los eventos $A_1, A_2, ..., A_k$; entonces*

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_k) &= P(A_1) P(A_2 \mid A_1) P(A_3 \mid A_1 \cap A_2) \\ &\quad \cdots P(A_k \mid A_1 \cap A_2 \cap \cdots \cap A_{k-1}). \end{aligned}$$

Definición 51 (Partición) *Sea A un conjunto y sean $A_1, A_2, ..., A_k$ subconjuntos de A. Decimos que los A_i, $i \in \{1, ..., k\}$ forman una partición para A, si son tales que:*

1. $\forall i \in \{1, ..., k\} : A_i \neq \emptyset$.
2. $A_i \cap A_j = \emptyset$, *para todo* $i \neq j$, $i, j \in \{1, ..., k\}$.
3. $\displaystyle\bigcup_{i=1}^{k} A_i = A$.

Teorema 52 (Probabilidad Total) *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sea $\{A_1, A_2, ..., A_k\}$ una partición de Ω con $A_i \in \mathcal{F}$, $i = 1, ..., k$. Sea $B \in \mathcal{F}$ dado arbitrariamente, entonces*

$$P(B) = \sum_{i=1}^{k} P(B \mid A_i) P(A_i).$$

Demostración. *Notemos que al ser los A_i, $i \in \{1, ..., k\}$ una partición para Ω, en particular tenemos que son ajenos dos a dos. Por tanto, podemos escribir B como*

$$B = \bigcup_{i=1}^{k} B \cap A_i,$$

donde las intersecciones del lado derecho de la igualdad anterior son ajenas dos a dos, luego

$$\begin{aligned} P(B) &= \sum_{i=1}^{k} P(B \cap A_i) \\ &= \sum_{i=1}^{k} P(B \mid A_i) P(A_i) \end{aligned}$$

■

Teorema 53 (de Bayes) *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\{A_1, A_2, ..., A_k\}$ una partición de Ω con $A_i \in \mathcal{F}$, $i = 1, ..., k$. Sea $B \in \mathcal{F}$ con $P(B) \neq 0$, entonces para todo $j \in \{1, ..., k\}$ se cumple que*

$$P(A_j \mid B) = \frac{P(B \mid A_j) P(A_j)}{\displaystyle\sum_{i=1}^{k} P(B \mid A_i) P(A_i)}.$$

Demostración. *Por la definición de probabilidad condicional tenemos que*

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)}.$$

Por otro lado, dado que $P(A_j) \neq 0$, *tenemos que*

$$P(A_j \cap B) = P(B \mid A_j)\, P(A_j)$$

y por el Teorema de Probabilidad Total,

$$P(B) = \sum_{i=1}^{k} P(B \mid A_i)\, P(A_i).$$

■

En el contexto de la estadística Bayesiana, a $P(B \mid A_j)$ se le llama la probabilidad "*a priori*" de que ocurra B, dado que ha ocurrido A_j, mientras que a $P(A_j \mid B)$ se le llama la probabilidad "*a posteriori*" de que haya ocurrido A_j dado que ocurrió B.

Ejemplo 54 *En una fábrica de tornillos las máquinas A, B y C producen, respectivamente, 25, 35 y 40 por ciento del total. De tales producciones, 5, 4, y 2 por ciento son tornillos defectuosos. De la producción del día, se toma una pieza y se observa que es defectuosa. ¿De cuál de las tres máquinas es más probable que provenga? Véase [Feller, 1978], p. 141.*

Ejemplo 55 *Suponga que 5 de cada 100 hombres y 25 de cada 10,000 mujeres son daltónicos. Se elige aleatoriamente una persona daltónica. ¿Cuál es la probabilidad de que sea hombre? (Suponga que las probabilidades de elegir un hombre o una mujer son iguales) Véase [Feller, 1978], p. 141.*

Ejemplo 56 *Pepito estaba sospechando desde hace timepo que está sufriendo de la rarísima y terrible enfermedad llamada Barbanorexia Picuda. Entonces hizo un exámen de sangre y salió positivo. Se espantó y fue a buscar más información en la internet. Resulta que 0.01 % de la población (uno en 10 mil) tiene esta enfremedad, pero resulta también que el exámen que hizo no es 100 % confiable, sino que tiene 1 % de probabilidades de ser "falso positivo"(o sea, hay una probabilidad de 1 en 100 que a alguien sin la enfermedad la prueba le indique que está enfermo), y 1 % de chance de "falso negativo". Pepito se deprimió mucho hasta que se sentó a calcular probablididades...*

La pregunta entonces es: ¿Cuál es la probabilidad de que Pepito esté sano (o sea, que no tenga la Barbanorexia Picuda)?

Vea [Centro de Investigación En Matemáticas, A. C., 1998].

Ejemplo 57 *Para realizar un estudio clínico, se recluta a pacientes en dos grupos: de control (a los cuales se les administra un placebo) y tratamiento (a quienes se les administra el medicamento que está a prueba). La probabilidad de que un resultado adverso se presente en el grupo de control es* p_0*, y en el grupo en tratamiento,* p_1*. A los pacientes se les asigna aleatoriamente a uno de los dos grupos y sus resultados son independientes. ¿Cuál es la probabilidad de que el primer resultado adverso ocurra en el grupo de control? Vea [Young, 2007].*

1.3.3. Ejercicios suplementarios

Tomados de [Feller, 1978].

1. Las letras en el código Morse se forman por una sucesión de guiones y puntos, donde se permite que haya repeticiones. ¿Cuántas letras se puede formar si se permite usar diez símbolos o menos?

2. Cada pieza de dominó está marcada con dos números. Las piezas son simétricas en el sentido de que cada par de números no está ordenado. ¿Cuántas piezas distintas se puede obtener usando los dígitos $1, 2, ..., n$?

3. Considere un grupo de n personas.

   a) Calcule la probabilidad de que al menos dos de ellas compartan fecha de cumpleaños.

   b) Con ayuda de R, calcule el valor que debe tener n para que la probabilidad considerada sea de al menos 0.5.

2. VARIABLES ALEATORIAS

La noción intuitiva de variable aleatoria es la de alguna característica de los elementos de interés que puede medirse o clasificarse. Formalmente, tenemos la siguientes definiciones:

Definición 58 (Función medible) *[Ash, 2000] Sean Ω_1 y Ω_2 espacios muestrales y sean $\mathcal{F}_1$ y $\mathcal{F}_2$ $\sigma-$álgebras sobre Ω_1 y Ω_2, respectivamente. Sea $h : \Omega_1 \to \Omega_2$. Se dice que h es medible con respecto a las $\sigma-$álgebras $\mathcal{F}_1$ y $\mathcal{F}_2$ si y sólo si $h^{-1}(A) \in \mathcal{F}_1$ para todo $A \in \mathcal{F}_2$.*

Definición 59 (Variable aleatoria) *[Ash, 2000] Una variable aleatoria X en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ es una función Borel-medible de Ω a $\mathbb{R}$.*

Muchas veces es conveniente permitir que X tome los valores $\pm\infty$. Se dice que X es una variable aleatoria extendida si es una función Borel-medible de Ω a $\mathbb{R}^{\#}$.

Si X es una variable aleatoria en $(\Omega, \mathcal{F}, P)$, la medida de probabilidad inducida por X es la medida de probabilidad P_X en la $\sigma-$álgebra de Borel, $\mathcal{B}(\mathbb{R})$, dada por

$$P_X(B) = P(\{\omega : X(\omega) \in B\}), \ B \in \mathcal{B}(\mathbb{R}).$$

Los valores que asume P_X caracterizan por completo a la variable aleatoria X en el sentido de que proporcionan las probabilidades de todos los eventos que involucran a X. Es de gran utilidad saber que dicha información puede darse en forma de una única función de $\mathbb{R}$ a $\mathbb{R}$.

Definición 60 (Función de distribución) *La función de distribución de una variable aleatoria X es la función F_X, o simplemente $F : \mathbb{R} \to [0,1]$ dada por*

$$\begin{aligned} F(x) &\equiv F_X(x) \\ &= P(X \leq x), \ x \in \mathbb{R}. \end{aligned}$$

F_X tiene las siguientes propiedades:

1. *Es continua por la derecha en $\mathbb{R}$, es decir que está definida en toda la recta real y $\forall a \in \mathbb{R} : \lim_{x \to a^+} F_X(x) = F_X(a)$.*

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

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

Definición 61 (Variable aleatoria discreta) *Sea X una variable aleatoria en $(\Omega, \mathcal{F}, P)$. Decimos que X es una variable aleatoria:*

1. *Degenerada, si toma un único valor.*

2. *Simple, si toma solamente una cantidad finita de valores.*

3. *Discreta, si asume una cantidad numerable de valores.*

En general, a las variables simples se les suele llamar también discretas.

Definición 62 (Función de masa de probabilidad) *Sea X una variable aleatoria discreta. A la función $p : \mathbb{R} \to [0, 1]$ dada por $p(x) = P(X = x), \quad x \in \mathbb{R}$, se le llama función de masa de probabilidad (fmp) de X.*

Nótese que la fmp de una v.a.d.X es tal que:

1. $0 \leq P(X = x) \leq 1, \quad \forall x.$

2. $\displaystyle\sum_x P(X = x) = 1.$

Si la variable aleatoria X es a lo más discreta, y ordenamos los valores que toma, digamos $\{x_n\}$, de modo que $x_n < x_{n+1}$ para todo n, entonces F_X es una función escalonada con una discontinuidad en cada x_n de magnitud $p_n = P(X = x_n)$; F_X es constante entre las x_n y toma el valor superior en cada discontinuidad.

Definición 63 (Variable aleatoria continua) *Sea X una variable aleatoria en $(\Omega, \mathcal{F}, P)$ con función de distribución F_X. Decimos que X es una variable aleatoria continua si y sólo si F_X es continua en todo $\mathbb{R}$.*

Definición 64 (Función de densidad de probabilidad) *Sea X una variable aleatoria en $(\Omega, \mathcal{F}, P)$ con función de distribución F_X. Se dice que X es continua si y sólo si existe una función continua no negativa f_X tal que*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt \qquad \forall x \in \mathbb{R}.$$

A tal f_X se le llama función de densidad de probabilidad (fdp) de X.

2.1. Valor esperado y varianza de una variable aleatoria

Definición 65 (Valor esperado) *Si X es una variable aleatoria discreta en el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y X toma los valores $x_1, x_2, ... \in D$ con probabilidades $p_i = P(X = x_i)$, $i \in \mathbb{N}$, entonces el valor esperado de X o esperanza matemática de X, denotado por $E(X)$ o μ_X, se define como*

$$\begin{aligned} E(X) &\equiv \mu_X = \sum_{x_i \in D} x_i P(X = x_i) && (2.1) \\ &= \sum_{x_i \in D} x_i p_i. && (2.2) \end{aligned}$$

siempre y cuando la suma converja.

Si X es una variable aleatoria continua con función de densidad de probabilidad f_X, entonces el valor esperado de X es

$$E(X) \equiv \mu_X = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

si la integral converge.

El valor esperado de X representa "el valor promedio" de la variable, por ejemplo, para una v.a.d. si consideamos que en un gran número de repeticiones del experimento, digamos N, observaremos x_i en Np_i ocasiones, así que el promedio de tales valores es $\frac{1}{N}(Np_1x_1 + Np_2x_2 + \cdots + Np_kx_k) = \sum_{i=1}^{k} x_i p_i$.

Ejemplo 66 *Considere el experimento de lanzar dos monedas. Sea X el número de "águilas" observadas. Calcule $E(X)$.*

Ejemplo 67 *La distribución de probabilidad de las calificaciones en un examen de estadística está dada por la tabla siguiente:*

Y	$p(y)$
5	0.05
6	0.10
7	0.40
8	0.20
9	0.20
10	0.05

¿Cuál es la calificación promedio del grupo?

Ejemplo 68 *[Mendenhall, W. Estadística Para Administradores. Problema 3.40] Una compañía manufacturera envía sus productos mediante camiones de remolque (trailers) de dos tamaños, uno de* $8 \times 10 \times 30$ *pies y de* $8 \times 10 \times 40$ *pies. Si 30 % de sus remesas se hace mediante camiones de 30 pies y 70 % utilizando el de 40 pies, calcule el volumen promedio de carga enviado por camión de remolque (suponga que cada vehículo va siempre lleno).*

Proposición 69 *Si la variable aleatoria* X *está definida en el conjunto* $D \subset \mathbb{R}$*, entonces el valor esperado de cualquier función* $h(X)$ *es*

$$E(h(X)) = \mu_{h(X)} = \sum_{x \in D} h(x) P(X = x),$$

si X *es discreta; y*

$$E(h(X)) = \mu_{h(X)} = \int_{x \in D} h(x) f_X(x) \, dx,$$

si X *es continua y* f_X *es su fmp.*

Proposición 70 *Sea* X *una variable aleatoria y sean* $a, b \in \mathbb{R}$*. Entonces*

$$E(aX + b) = aE(X) + b.$$

Proposición 71 *Sean* $X_1, X_2, ..., X_n$ *variables aleatorias, entonces*

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n).$$

Definición 72 (Momentos y momentos centrales) *Sea* X *una variable aleatoria discreta en* $(\Omega, \mathcal{F}, P)$ *con* $E(X) = \mu_X$*. Si* $k > 0$*, entonces:*

1. *Al número* $E\left(X^k\right)$ *se le llama* $k-$*ésimo momento de* X*,*
2. *A* $E\left(|X|^k\right)$ *se le llama* $k-$*ésimo momento absoluto de* X*,*
3. $E\left((X - \mu_X)^k\right)$ *se llama* $k-$*ésimo momento central de* X*,*
4. *El número* $E\left(|X - \mu|^k\right)$ *se llama* $k-$*ésimo momento central absoluto de* X*.*

Nótese que el valor esperado de X es el primer momento de la variable.

El primer momento central de X es de printipal interés en estadística, como vemos en la siguiente

Definición 73 (Varianza) *Al segundo momento central de una variable aleatoria X se le llama varianza de X. Se le representa por $Var(X)$ o por σ_X^2 y, de acuerdo con la definición previa, se calcula como sigue:*

Para una v.a.d.:

$$\begin{aligned} Var(X) &\equiv \sigma_X^2 & (2.3)\\ &= E\left((X-\mu_X)^2\right) & (2.4)\\ &= \sum_x (x-\mu_X)^2 p(x), & (2.5)\end{aligned}$$

si la suma converge y donde p es la fmp de X.

Para una v.a.c.:

$$\begin{aligned} Var(X) &\equiv \sigma_X^2 & (2.6)\\ &= E\left((X-\mu_X)^2\right) & (2.7)\\ &= \int_x (x-\mu_X)^2 f_X(x)\,dx, & (2.8)\end{aligned}$$

si la integral converge y donde f_X es la fdp de X.

Corolario 74 *Si X es una variable aleatoria con valor esperado μ_X y varianza $\sigma_X^2 < \infty$, entonces*

$$\sigma_X^2 = E\left(X^2\right) - \mu_X^2. \tag{2.9}$$

Ejemplo 75 *[Bain, L. J.; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Ejercicio 2.26. p. 87] En una tienda de cómputo, la demanda anual por un paquete de software particular es una variable discreta X. El dueño de la tienda ordena cuatro copias del paquete a \$100.00 cada copia y las vende a los clientes a \$350.00 cada una. Al final del año el paquete es obsoleto y el propietario pierde la inversión hecha en las copias no vendidas. Las fdp's para X, el número de copias vendidas, y Y, las ganancias del dueño de la tienda, están dadas en las siguientes tablas:*

x	0	1	2	3	4
$p(x)$	0.1	0.3	0.3	0.2	0.1

y	-400	-50	300	650	1000
$p(y)$	0.1	0.3	0.3	0.2	0.1

a) Encuentre el número de copias que esperaría vender, $E(X)$.
b) Encuentre la varianza del número de copias vendidas, $Var(X)$.
c) Calcule la ganancia esperada por el propietario de la tienda, $E(Y)$.
d) Calcule la varianza de la ganancia, $Var(Y)$.

Ejemplo 76 *[Johnson, R. Estadística Elemental. Ejercicio 5.16. p. 191] Tomando como base datos históricos, la distribución de ventas diarias de paquetes de queso cottage de una libra de peso (0.454 kg) en una tienda es como sigue:*

No. de paquetes vendidos, x	10	11	12	13	14	Total
$p(x)$	0.2	0.4	0.2	0.1	0.1	1.0

a) Calcule el número promedio de paquetes vendidos.
b) Encuentre la varianza del número de paquetes vendidos.

Ejemplo 77 *Con frecuencia se hace uso de los datos provenientes de censos para obtener distribuciones de probabilidad de diversas variables aleatorias. Cierto estudio en los EUA sobre este tipo de datos, relativos a familias con ingresos totales anuales iguales o mayores de \$50,000, indica que el 20 % no tiene hijos, 30 % tiene uno, 40 % tiene dos y el restante 10 % tiene tres. Utilizando esta información:*

a) Construya una tabla de la distribución de probabilidad del número de hijos para estas familias.
b) Calcule el número promedio de hijos.
c) Calcule la varianza del número de hijos para este tipo de familias.

2.2. Algunas distribuciones discretas de probabilidad

2.2.1. Distribución de probabilidad uniforme discreta

Definición 78 *Si la variable aleatoria discreta X toma los valores $x_1, x_2, ..., x_k$ con idénticas probabilidades, entonces X sigue una distribución uniforme discreta con parámetro k, lo cual representamos con $X \sim UD(k)$, y su función de masa de probabilidad está dada por*

$$u(x; k) = \frac{1}{k}; \quad x = x_1, x_2, ..., x_k. \tag{2.10}$$

Teorema 79 *La media y la varianza de una variable aleatoria uniforme discreta están dadas por*

$$E(X) = \mu_X = \frac{1}{k} \sum_{i=1}^{k} x_i \tag{2.11}$$

y

$$Var(X) = \sigma_X^2 = \frac{1}{k} \sum_{i=1}^{k} (x_i - \mu_X)^2 . \tag{2.12}$$

Demostración. *Sabemos que $E(X) = \sum_x xP(X = x)$, luego*

$$\begin{aligned} E(X) &= \sum_{i=1}^{k} x_i \frac{1}{k} \\ E(X) &= \frac{1}{k} \sum_{i=1}^{k} x_i . \end{aligned}$$

$$\textit{Además, } Var\left(X\right)=E\left(\left(X-\mu_{X}\right)^{2}\right)=\sum_{x}\left(x-\mu_{X}\right)^{2}P\left(X=x\right), \textit{ luego}$$

$$\begin{aligned}
Var\left(X\right) &= \sum_{i=1}^{k}\left(x_{i}-\mu_{X}\right)^{2}\frac{1}{k}\\
Var\left(X\right) &= \frac{1}{k}\sum_{i=1}^{k}\left(x_{i}-\mu_{X}\right)^{2}.
\end{aligned}$$

■

Ejemplo 80 *Suponga que se lanza un dado legal y llamemos X al número de puntos mostrado por la cara superior del dado. Entonces $X\sim UD\left(6\right)$, y $f\left(x;6\right)=\frac{1}{6}$; $\quad x=1,2,3,4,5,6.$*

a) Calcule $E\left(X\right)$.

c) Obtenga el valor de $Var\left(X\right)$.

Ejemplo 81 *Considere el experimento de extraer un foco de una caja que contiene cuatro de ellos, uno de 40 watts (w), uno de 60 w, uno de 75 w y uno de 100 w. Sea $Y=$ Potencia del foco extraído.*

a) Elabore una gráfica de la fmp de Y.

b) Calcule $E\left(Y\right)$.

c) Calcule $Var\left(Y\right)$.

2.2.2. Distribución de probabilidad Bernoulli

Definición 82 (Ensayo Bernoulli) *Se le llama ensayo Bernoulli a una realización de un experimento en el cual sólo puede ocurrir uno de dos eventos, digamos E y su complemento E'. De manera arbitraria al evento E se le llama éxito y al evento E' se le llama fracaso. Se representa a la probabilidad de que ocurra el evento E con p y a la probabilidad de que ocurra el evento E' con q. En consecuencia se tiene que*

$$P\left(E\right)=p$$

y por lo tanto

$$P\left(E'\right)=q=1-p.$$

Definición 83 (Distribución Bernoulli) *Si un experimento puede resultar únicamente en "éxito" (E) o "fracaso" (F), y si definimos la variable X como*

$$X=\begin{cases} 0 & \textit{si ocurre } E' \\ 1 & \textit{si ocurre } E \end{cases}$$

entonces la variable X sigue una distribución Bernoulli con parámetro p, lo cual denotamos $X\sim Ber\left(p\right)$.

Teorema 84 *Sea $X \sim Ber(p)$. Entonces la función de masa de probabilidad de X está dada por*

$$Ber(x; p) = p^x q^{1-x}; \quad x = 0, 1. \tag{2.13}$$

El valor esperado y varianza de una variable aleatoria Bernoulli son, respectivamente

$$E(X) = p \tag{2.14}$$

$$Var(X) = pq. \tag{2.15}$$

Demostración. *Sea $X \sim Ber(p)$, entonces X puede tomar solamente los valores 0 o 1. Notemos que*

$$\begin{aligned} P(X=0) &= q \\ &= p^0 q^{1-0} \end{aligned}$$

y que

$$\begin{aligned} P(X=1) &= p \\ &= p^1 q^{1-1} \end{aligned}$$

por tanto la fmp de X está dada por $Ber(x; p) = p^x q^{1-x}; \quad x = 0, 1.$

El valor esperado de X es

$$\begin{aligned} E(X) &= \sum_x x P(X = x) \\ &= 0 \cdot P(X=0) + 1 \cdot P(X=1) \\ &= p. \end{aligned}$$

Finalmente, la varianza de X es

$$\begin{aligned} Var(X) &= \sum_x (x - \mu_X)^2 P(X = x) \\ &= (0-p)^2 (1-p) + (1-p)^2 p \\ &= p^2 (1-p) + (1-p)^2 p \\ &= p(1-p)(p+1-p) \\ &= p(1-p). \end{aligned}$$

■

Ejemplo 85 *Suponga que cierto componente se somete a una prueba de choque dada. Sea*

$$X = \begin{cases} 0 & \text{si el componente sobrevive a la prueba} \\ 1 & \text{en caso contrario} \end{cases}$$

y suponga además que $P(X = 1) = 0{,}8$.

a) Obtenga una expresión para $f(x; p)$.
b) Calcule $E(X)$.
c) Calcule $Var(X)$.

Ejemplo 86 *Sea el experimento de clasificar a las personas en una fila de espera como hombre o mujer. Sea*

$$X = \begin{cases} 0 \text{ si la persona es hombre} \\ 1 \text{ si la persona es mujer} \end{cases}.$$

Suponga que $P(X=1) = 0{,}5$.
a) Exprese $f(x;p)$ *como una función.*
b) Calcule $E(X)$.
c) Calcule $Var(X)$.

2.2.3. Distribución de probabilidad binomial

Definición 87 (Distribución binomial) *Consideremos el experimento de realizar* n*, de ensayos Bernoulli independientes, cada uno con la misma probabilidad de éxito igual a* p*, y sea* $X =$ *Número de éxitos observado en los* n *ensayos. Entonces* X *sigue la distribución binomial con parámetros* n *y* p*, lo cual representamos por*

$$X \sim Bin(n,p).$$

Teorema 88 (Distribución binomial) *Si* $X \sim Bin(n,p)$ *entonces la fmp de* X *está dada por*

$$b(x;n,p) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0,1,2,...,n \tag{2.16}$$

Corolario 89 *Si* $b(x;n,p) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0,1,2,...,n,$ *entonces*

$$\sum_{x=0}^{n} b(x;n,p) = 1.$$

Demostración. *Por el teorema del binomio, tenemos que*

$$\begin{aligned} \sum_{x=0}^{n} b(x;n,p) &= \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \\ &= [p + (1-p)]^n = 1^n \\ &= 1. \end{aligned}$$

■

Teorema 90 (Valor esperado y varianza de una v.a. binomial) *Sea* $X \sim Bin(n,p)$*. El valor esperado de* X *está dado por*

$$E(X) = np. \tag{2.17}$$

La varianza de X está dada por

$$Var(X) = npq. \tag{2.18}$$

Demostración. *Observemos que podemos escribir X como $X = X_1 + X_2 + \cdots + X_n$, donde $X_i \sim Ber(p)$, $\quad i = 1, ..., n$, y entonces*

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \cdots + X_n) \\ &= E(X_1) + E(X_2) + \cdots + E(X_n) \\ &= p + p + \cdots + p \\ &= np. \end{aligned}$$

y, dado que las X_i son independientes,

$$\begin{aligned} Var(X) &= Var(X_1) + \cdots + Var(X_n) \\ &= pq + \cdots + pq \\ &= npq. \end{aligned}$$

■

Ejemplo 91 *La probabilidad de que cierta clase de componente sobreviva a una prueba de choque dada es 3/4. Si se prueban seis componentes de manera consecutiva,*

a) Encuentre la probabilidad de que sobrevivan exactamente dos de los seis componentes.

b) Encuentre la probabilidad de que sobrevivan a lo más dos componentes.

c) Encuentre la probabilidad de que sobrevivan más de dos componentes.

d) Calcule el número esperado de componentes que sobrevivirán entre los seis probados.

Ejemplo 92 *Se realiza el lanzamiento de tres monedas y se observa $X =$ Número de "águilas" en los lanzamientos.*

a) Calcule la probabilidad de observar cero "águilas".

b) Calcule la probabilidad de que se observe un "águila", $X = 1$

c) Obtenga la probabilidad de observar cuando mucho un "águila"

d) Obtenga la probabilidad de observar más de un "águila".

e) Calcule el número esperado de "águilas" en los tres lanzamientos y la varianza del número de "águilas".

Ejemplo 93 *Parte de cierto examen de estadística contiene cinco preguntas que deben ser respondidas con verdadero o falso. Suponga que un alumno responde aleatoriamente a las cinco preguntas. ¿Cuál es la probabilidad de que acierte*

a) a ninguna respuesta?

b) cuando mucho a una respuesta?

c) a más de una respuesta?

d) a las cinco respuestas?

e) Calcule el número esperado de respuestas correctas y la varianza del número de respuestas correctas.

Ejemplo 94 *Los registros muestran que 30 % de todos los pacientes admitidos a una clínica paga sus facturas con tarjeta de crédito. Suponga que durante cierta semana fueron admitidos 10 pacientes. Determine:*

a) La probabilidad de que ningún paciente pague con tarjeta de crédito.

b) La probabilidad de que menos de cinco pacientes paguen con tarjeta de crédito.

c) La probabilidad de que todos los pacientes paguen con tarjeta de crédito.

d) Que todos los pacientes, menos uno, paguen con tarjeta de crédito.

e) Calcule el valor esperado y la varianza del número de pacientes que pagarán con tarjeta de crédito.

2.2.4. Distribución de probabilidad hipergeométrica

Supongamos que una población o colección consta de un número finito de elementos, digamos N, y que hay r elementos del tipo 1 (en los cuales estamos interesados) y $N-r$ elementos del tipo 2. Supongamos además que se realiza la extracción de $0<n<N$ elementos sin reemplazo, y denotemos por X el número de elementos del tipo 1 entre los extraídos. La variable X, entonces, sigue una distribución de probabilidad hipergeométrica con parámetros N, r y n, lo cual representamos por

$$X \sim Hip\,(N, r, n)$$

Teorema 95 (Distribución de probabilidad hipergeométrica) *Sea* $X \sim Hip\,(N, r, n)$. *La fmp de* X *está dada por*

$$h(x; N, r, n) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}; \qquad \text{máx}\,(0, n-(N-r)) \leq x \leq \text{mín}\,(n, r)\,. \quad (2.19)$$

Demostración. Supongamos que tenemos N elementos, de los cuales r son de tipo 1 y $N-r$ de tipo 2, y que se realiza la extracción sin reemplazo de n de los elementos. Sea X el número de objetos de tipo 1 entre los n extraídos, entonces:

- El número de grupos de tamaño n que se puede obtener de los N objetos es $\binom{N}{n}$.
- La cantidad de grupos de tamaño x que se puede obtener de los r objetos de tipo 1 es $\binom{r}{x}$.
- El número de subconjuntos de tamaño $n-x$ de un conjunto con $N-r$ elementos es $\binom{N-r}{n-x}$.
- En consecuencia, la probabilidad de extraer x elementos de tipo 1 entre los n escogidos es $\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$

Ahora, el número máximo de elementos de tipo 1 en la extracción es n, a menos que $r < n$, por tanto $x < \text{mín}(n, r)$. Además, el número mínimo de elementos de tipo 1 en la extracción es 0, a no ser que haya menos de n elementos de tipo 2, de modo que en la extracción necesariamente habrá al menos uno de tipo 1. Por tanto $\text{máx}(0, n - (N - r)) < x$.

Por lo anterior, tenemos que la fmp de X es

$$h(x; N, r, n) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad \text{máx}(0, n - (N - r)) \leq x \leq \text{mín}(n, r).$$

■

Teorema 96 (Valor esperado y varianza de una v.a. hipergeométrica) *El valor esperado y la varianza de una variable aleatoria hipergeométrica X están dados por*

$$E(X) = \frac{nr}{N} \tag{2.20}$$

y

$$Var(X) = \frac{rn(N-r)(N-n)}{N^2(N-1)}. \tag{2.21}$$

Ejemplo 97 *En un lote de cinco computadoras hay dos de ellas que son defectuosas. Se seleccionan tres de las computadoras aleatoriamente. Calcule la probabilidad de que entre las tres seleccionadas:*
a) no haya ninguna defectuosa.
b) haya cuando mucho una defectuosa.
c) haya más de una defectuosa.
d) Calcule el número esperado de computadoras defectuosas entre las tres seleccionadas.
e) Calcule la varianza del número de computadoras defectuosas en la selección.

Ejemplo 98 *Un producto industrial particular se envía en lotes de 20 artículos. La prueba para determinar si un artículo es defectuoso es costosa, así que el fabricante obtiene muestras de la producción en vez de utilizar un plan de inspección al 100 %. Un plan de muestreo diseñado para minimizar el número de artículos defectuosos empleados necesita que se extraigan 5 artículos de cada lote y un lote es rechazado si entre las cinco piezas hay más de una defectuosa. Suponga que un lote en particular contiene cuatro artículos defectuosos. Calcule:*

a) La probabilidad de que el lote sea rechazado.

b) El número esperado de unidades defectuosas en la selección.

c) La varianza del número de unidades defectuosas en la selección.

d) La probabilidad de que un lote con sólo 2 unidades defectuosas sea rechazado.

e) La probabilidad de que un lote con 5 piezas defectuosas no sea rechazado.

2.2.5. Distribución de probabilidad Poisson

En ocasiones se tiene una situación física en la cual cierto evento es recurrente, tal como las llamadas que se reciben en cierto intervalo de tiempo en una central telefónica o el número de defectos en un tramo de cable. Si la probabilidad de ocurrencia de un evento de este tipo es proporcional al tamaño del intervalo, área, volumen, etc. entonces la variable aleatoria X, el número de eventos observados en una unidad (de tiempo, de longitud, etc.) sigue una distribución conocida como Poisson.

Definición 99 (Experimento Poisson) *Es un experimento en el cual interesa observar la variable X, el número de ocurrencias de cierto evento durante un cierto intervalo de tiempo (o longitud, o área, o volumen, etc). Un experimento de este tipo tiene las siguientes características [Walpole and Myers, 1992]:*

1. *El número de resultados en un intervalo de tiempo o región específicos es independiente del número que ocurre en cualquier otro intervalo disjunto o región del espacio disjunta. A esto se le llama la propiedad de falta de memoria de la distribución Poisson.*

2. *La probabilidad de que un único resultado ocurra en un intervalo de tiempo muy corto o en una región pequeña es proporcional a la longitud del intervalo de tiempo o al tamaño de la longitud y no depende del número de resultados que ocurre fuera de ese intervalo o región.*

3. *La probabilidad de que más de un resultado ocurra en ese intervalo de tiempo tan corto o en esa pequeña región tiende a cero.*

Teorema 100 *Si cierto evento aleatorio ocurre con una intensidad λ por cada unidad de tiempo (longitud, área, volumen, etc.) y la variable X cuenta el número de eventos que ocurre en un intervalo de tiempo (longitud, área, volumen, etc.) de magnitud t, entonces X sigue la distribución Poisson con parámetro λt, lo cual representamos por $X \sim Poi(\lambda t)$, y su función de masa de probabilidad es*

$$p(x; \lambda t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \qquad x = 0, 1, 2, ... \tag{2.22}$$

Teorema 101 (Valor esperado y varianza de una variable aleatoria Poisson) *Para una variable aleatoria Poisson X con parámetro λ, se tiene que*

$$E(X) = \lambda \tag{2.23}$$

y

$$Var(X) = \lambda. \tag{2.24}$$

Demostración. *Por definición,* $E(X)=\sum_{x=0}^{\infty}xP(X=x)=\sum_{x=0}^{\infty}x\frac{\lambda^x e^{-\lambda}}{x!}$. *Como en el primer término* $x=0$, *podemos escribir*

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty}x\frac{\lambda^x e^{-\lambda}}{x!}=\sum_{x=1}^{\infty}\frac{\lambda\lambda^{x-1}e^{-\lambda}}{(x-1)!} \\ &= \lambda\sum_{x=1}^{\infty}\frac{\lambda^{x-1}e^{-\lambda}}{(x-1)!}. \end{aligned}$$

Si hacemos $y=x-1$, *entonces tenemos que*

$$E(X)=\lambda\sum_{y=0}^{\infty}\frac{\lambda^y e^{-\lambda}}{y!}.$$

Notamos que los términos de la serie corresponden a las probabilidades de una variable aleatoria $Y\sim Poi(\lambda)$, *por lo cual* $\sum_{y=0}^{\infty}\frac{\lambda^y e^{-\lambda}}{y!}=1$, *y en consecuencia*

$$E(X)=\lambda.$$

Para calcular la varianza de X, *recordemos que podemos expresarla como* $Var(X)=E(X^2)-[E(X)]^2$ *y notemos que* $E(X(X-1))=E(X^2-X)=E(X^2)-E(X)$, *de donde* $E(X^2)=E(X(X-1))+E(X)$.

Consideremos entonces $E(X(X-1))$ *y observemos que*

$$E(X(X-1))=\sum_{x=0}^{\infty}x(x-1)\frac{\lambda^x e^{-\lambda}}{x!}$$

y como los dos primeros términos se cancelan, podemos escribir

$$\begin{aligned} E(X(X-1)) &= \sum_{x=2}^{\infty}x(x-1)\frac{\lambda^x e^{-\lambda}}{x!} \\ &= \sum_{x=2}^{\infty}\frac{\lambda^2\lambda^{x-2}e^{-\lambda}}{(x-2)!} \\ &= \lambda^2\sum_{x=2}^{\infty}\frac{\lambda^{x-2}e^{-\lambda}}{(x-2)!}. \end{aligned}$$

Si hacemos ahora $y=x-2$, *tenemos que*

$$E(X(X-1))=\lambda^2.$$

Por tanto,

$$\begin{aligned} Var(X) &= E\left(X^2\right)-[E(X)]^2 \\ &= E(X(X-1))+E(X)-[E(X)]^2 \\ &= \lambda^2+\lambda-[\lambda]^2 \\ Var(X) &= \lambda. \end{aligned}$$

■

Ejemplo 102 *[Bain, L.J; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Problema 3.21] El número de llamadas que se reciben en un conmutador durante una hora sigue una distribución Poisson con media $\lambda = 10$. Encuentre la probabilidad de ocurrencia durante una hora de cada uno de los siguientes eventos:*

a) Se reciben exactamente dos llamadas.

b) Se reciben a lo más dos llamadas.

c) Se reciben más de dos llamadas.

Ejemplo 103 *Suponga que el número de errores mecanográficos en un documento sigue una distribución Poisson con una media de 3 errores por página. Calcule la probabilidad de que:*

a) en una página dada haya 3 errores.

b) en dos páginas haya 3 errores.

c) en tres páginas haya 3 errores.

Ejemplo 104 *[Bain, L.J; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Problema ?.??] Cierta línea de ensamblado produce componentes electrónicos a razón de 500 por hora, y durante ese tiempo se tiene un promedio de 5 componentes defectuosos. Para una hora dada:*

a) ¿Cuál es la probabilidad de que el númeor de componentes defectuosos sea cuando mucho dos?

b) ¿Cuál es la probabilidad de que el número de componentes defectuosos sea mayor que dos?

Ejemplo 105 *Si X sigue una distribución Poisson y $P(X = 0) = 0{,}2$, encuentre $P(X > 4)$.*

Ejemplo 106 *En una intersección ocurren en promedio tres accidentes de tránsito por mes. ¿Cuál es la probabilidad de que en esta intersección*

a) ocurran exactamente cinco accidentes en el próximo mes?

b) ocurran exactamente cinco accidentes en los dos próximos meses?

c) ocurran exactamente cinco accidentes en los siguientes quince días?

Ejemplo 107 *Cierta área del este de Estados Unidos resulta, en promedio, afectada por seis huracanes al año. Encuentre la probabilidad de que para cierto año esta área resulte afectada por*

a) menos de cuatro huracanes.

b) cualquier cantidad entre seis y ocho huracanes.

2.2.6. Distribución de probabilidad binomial negativa

Consideremos una sucesión de ensayos Bernoulli independientes cada uno con probabilidad de éxito p. Supongamos que observamos X, el número de fracasos antes de lograr k éxitos. Entonces se dice que X sigue la distribución binomial negativa[1] con parámetros p y k, $X \sim BN(k,p)$.

Teorema 108 *Sea $X \sim BN(k,p)$. Entonces la función de masa de probabilidad de X está dada por*

$$bn(x;k,p) = \binom{x+k-1}{k-1} p^k (1-p)^x ; \quad x = 0,1,...$$

Demostración. *Supongamos que $X \sim BN(k,p)$ y que en una realización del experimento se reqiere de $x+k$ ensayos, el último de los cuales es un éxito y que también hay $k-1$ éxitos entre los $x+k-1$ ensayos previos. Una forma de observar esto es mediante el evento A, en el que los primeros $k-1$ ensayos resulten ser éxitos y los siguientes x sean fracasos (el último ensayo necesariamente es un éxito):*

$$A = \underbrace{EEE...E}_{k-1 \text{ éxitos}} \underbrace{FF..F}_{x \text{ fracasos}} E$$

luego, como los ensayos son independientes,

$$\begin{aligned} P(A) &= \underbrace{P(E)P(E)...P(E)}_{k-1}\underbrace{P(F)...P(F)}_{x}P(E) \\ &= p^{k-1}(1-p)^x p \\ &= p^k (1-p)^x . \end{aligned}$$

Observemos ahora que cualquier otra forma de observar un evento con x éxitos y x fracasos en los primeros $x+k-1$ ensayos es simplemente un reacomodo de los ensayos en el evento A (con excepción del último), y el número de ellos corresponde al número de formas de acomodar $k-1$ éxitos en $x+k-1$ ensayos, o equivalentemente, al número de formas de extraer $k-1$ objetos de entre $x+k-1$, es decir $\binom{x+k-1}{k-1}$. Por tanto

$$bn(x;k,p) = \binom{x+k-1}{k-1} p^k (1-p)^x ; \quad x = 0,1,...$$

■

[1] Se le llama binomial negativa en contraste con la distribución binomial, ya que en la binomial el número de éxitos n es fijo, mientras que el número de éxitos, x, es aleatorio. En la distribución binomial negativa el número de éxitos k es fijo y el número de ensayos, $x+k$, es aleatorio

Teorema 109 *El valor esperado y la varianza de una variable aleatoria binomial negativa con parámetros k y p son*

$$E(X) = \frac{k(1-p)}{p}$$

y

$$Var(X) = \frac{k(1-p)}{p^2}.$$

2.2.7. Distribución de probabilidad geométrica

Consideremos el caso particular de una distribución binomial negativa con parámetros p y $k = 1$. Entonces X cuenta el número de fracasos antes del primer éxito. En este caso, se dice que X sigue la distribución geométrica con parámetro p, $X \sim Geo(p)$.

Corolario 110 *Sea X una variable aleatoria que sigue la distribución geométrica con parámetro p. Entonces la función de masa de probabilidad de X es*

$$g(x;p) = p(1-p)^{x-1}; \quad x = 1, 2, ...$$

Su valor esperado y su varianza son

$$E(X) = \frac{1-p}{p}$$

y

$$Var(X) = \frac{1-p}{p^2}.$$

Corolario 111 (Función de distribución de probabilidad geométrica) *Sea $X \sim Geo(p)$. Entonces su función de distribución de probabilidad es*

$$F(x) = 1 - q^x,$$

donde $q = 1 - p$.

Teorema 112 ***Demostración.*** *Por definición, la función de distribución de probabilidad es*

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

y en el caso particular de una variable geométrica, tenemos que

$$\begin{aligned} F(x) &= \sum_{i=1}^{x} p(1-p)^{i-1} \\ &= p\sum_{i=1}^{x} q^{i-1}, \qquad q := 1-p. \end{aligned}$$

Si hacemos el cambio de índice $j = i - 1$, tenemos que

$$F(x) = p \sum_{j=0}^{x-1} q^j. \tag{2.25}$$

Sea $s_n := \sum_{i=0}^{x-1} q^j$*. Observemos que si multiplicamos ambos miembros de la definición de* s_n *por* $1 - q$*, obtenemos:*

$$\begin{aligned} (1-q)\, s_n &= (1-q) \\ &= \sum_{j=0}^{x-1} \left(q^j - q^{j+1}\right) \\ &= 1 - q^x, \end{aligned}$$

porque se trata de una suma telescópica. Así que, despejando s_n*, tenemos*

$$s_n = \frac{1-q^x}{1-q}.$$

Si sustituimos en [2.25], tenemos

$$\begin{aligned} F(x) &= p\frac{1-q^x}{1-q} \\ &= p\frac{1-q^x}{p} \end{aligned}$$

luego

$$F(x) = 1 - q^x. \tag{2.26}$$

■

Teorema 113 (Falta de memoria de la distribución geométrica) *Sea* $X \sim Geo(p)$ *y sean* $s, t \in \mathbb{N}$*. Entonces*

$$P(X > t + s \mid X > t) = P(X > s).$$

Demostración. *Por definición,*

$$P(X > t + s \mid X > t) = \frac{P(X > t + s, X > t)}{P(X > t)}$$

pero como $s > 0$*, tenemos que* $X > t + s \Rightarrow X > t$*, así que* $\{X > t + s\} \subset \{X > t\}$*. Por tanto*

$$P(X > t + s, X > t) = P(X > t + s)$$

de donde

$$P(X > t + s \mid X > t) = \frac{P(X > t + s)}{P(X > t)}. \tag{2.27}$$

Al sustituir esta última expresión en [2.25], resulta:

$$
\begin{aligned}
P(X > t+s \mid X > t) &= \frac{1-F(t+s)}{1-F(t)} \\
&= \frac{1-(1-q^{t+s})}{1-(1-q^{t})} \\
&= \frac{q^{t+s}}{q^{t}} = q^{s} \\
&= 1-(1-q^{s}) \\
&= 1-F(s) \\
&= P(X > s).
\end{aligned}
$$

■

2.3. Algunas distribuciones continuas de probabilidad

2.3.1. Introducción

La noción de variable aleatoria discreta proporciona un medio adecuado para modelar probabiliddes en una amplia clase de problemas; sin embargo, en muchas otras situaciones una variable aleatoria discreta no es un modelo adecuado.

Las variables aleatorias continuas surgen de experimentos relacionados con algún proceso de medición de cantidades físicas como la longitud, el tiempo, la masa o las obtenidas a partir de ellas. Los siguientes son ejemplos de variables aleatorias continuas:

1. La magnitud, en escala Richter, de un sismo.
2. La altura que alcanza el agua, medida con respecto al fondo de una presa.
3. El tiempo transcurrido entre llegadas sucesivas de vehículos a una intersección.
4. La fuerza requerida para doblar una barra de acero reforzado.

Recordemos que la función de distribución (fd), F_X, de una variable aleatoria X se define como $F_X(x) = P(X \leq x)$ y que la función de densidad de probabilidad (fdp) de una variable aleatoria continua X es una función continua f_X tal que

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

De lo anterior y del Teorema Fundamental del Cálculo tenemos una forma para obtener la función de densidad de probabilidad a partir de la función de distribución:

$$f_X(x) = \frac{d}{dx} F_X(x) = F'_X(x),$$

cuando la derivada existe.

Teorema 114 *Una función f_X es una función de densidad de probabilidad para alguna variable aleatoria continua X si y sólo si satisface las propiedades siguientes:*

1. La función f_X evaluada en cualquier valor x de la v. a. X es mayor o igual que cero:

$$f_X(x) \geq 0, \quad \forall x \in \mathbf{R}.$$

2. El área comprendida bajo la curva f_X y sobre el eje horizontal es igual a 1:

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1.$$

Cuando se conoce la fdp para una v.a.c. X se puede calcular la probabilidad de que X se encuentre en un intervalo dado

$$\begin{aligned} P(a \leq X \leq b) &= P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a) \\ &= \int_{-\infty}^{b} f_X(x)\, dx - \int_{-\infty}^{a} f_X(x)\, dx \\ &= \int_{a}^{b} f(x)\, dx, \end{aligned}$$

mientras que la probabilidad de que una v.a.c. X tome un valor específico cualquiera, es igual a cero:

$$\begin{aligned} P(X = a) &= P(a \leq X \leq a) \\ &= \int_{a}^{a} f_X(x)\, dx \\ &= 0. \end{aligned}$$

Corolario 115 *Sea X una variable aleatoria continua. Entonces se cumplen las siguientes igualdades:*

$$\begin{aligned} P(a < X < b) &= P(a \leq X < b) \\ &= P(a < X \leq b) \\ &= P(a \leq X \leq b). \end{aligned}$$

Recordemos si X es una variable aleatoria continua con función de densidad de probabilidad f_X, entonces el valor esperado de X es

$$\begin{aligned} E(X) &\equiv \mu_X \\ &= \int_{-\infty}^{\infty} x f_X(x)\, dx \end{aligned}$$

si la integral converge absolutamente, y que al segundo momento central de una variable aleatoria X se le llama varianza de X; se le representa por $Var(X)$ o por σ_X^2 y, para una v.a.c. X con fdp f_X, se calcula como sigue:

$$\begin{aligned} Var(X) &\equiv \sigma_X^2 = E\left((X-\mu_X)^2\right) \\ &= \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\, dx, \end{aligned}$$

si la integral converge absolutamente.

Además de la media y la varianza, existen otras medidas de interés para variables aleatorias, que presentamos a continuación:

Definición 116 (Cuantil o percentil) *Si $0 < p < 1$, entonces un $100p-$ésimo percentil de la distribución de una variable aleatoria continua X es una solución x_p de la ecuación*

$$F_X(x_p) = p. \tag{2.28}$$

Notemos que si se tratara de una v.a. discreta, la ecuación [2.28] no tendría solución en general. Esto puede remediarse si replanteamos la definición del $100p-$*ésimo* cuantil de una distribución como un valor x_p tal que:

$$P(X \leq x_p) \geq p$$

y

$$P(X \geq x_p) \leq 1-p.$$

Definición 117 (Cuartiles y mediana) *Algunos percentiles muy utilizados en la descripción de datos son los "cuartiles", llamados así porque dividen la distribución en cuatro partes de probabilidad $\frac{1}{4}$:*

- *Al cuantil 25 % se le llama **primer cuartil**.*
- *El cuantil 50 % es la mediana, que corresponde con el segundo cuartil.*
- *Al cuantil 75 % se le llama tercer cuartil.*

Definición 118 (Moda) *Sea X una variable aleatoria continua. Si la fmp de X, f_X, tiene un máximo en $x = \widehat{\mu}_0$, entonces se dice que la moda de la función de masa de probabilidad de X es $\widehat{\mu}_0$.*

Ejemplo 119 *[Bain, L. J.; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Ejemplo 2.3.3. p. 65] Una máquina produce alambre de cobre, y ocasionalmente existe una imperfección en algún punto a lo largo del cable. La longitud del cable (en metros) producida entre imperfecciones sucesivas es una variable aleatoria continua X con fdp de la forma*

$$f_X(x) = \begin{cases} c(1+x)^{-3}, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

a) Calcule el valor de c.
b) Obtenga la expresión para F_X.
c) Calcule $P(0{,}40 < X < 0{,}45)$.
d) Obtenga el valor esperado de X.
e) Calcule la varianza de X.

Ejemplo 120 *[Bain, L. J.; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Ejercicio 2.12. p. 85] Una variable aleatoria continua tiene una fdp dada por $f(x) = c(1-x)x^2$, si $0 < x < 1$ y cero en otro caso.*
a) Encuentre la constante c.
b) Calcule $P(0{,}5 < x < 0{,}6)$.
c) Calcule $E(X)$ y $Var(X)$.

2.3.2. Distribución de probabilidad normal o gaussiana

Puede considerarse a la distribución normal o gaussiana como una de las más importantes dentro de la probabilidad y la estadística por varias razones, entre las que podemos mencionar las siguientes:

1. Numerosos fenómenos continuos parecen seguirla o se pueden aproximar mediante ella.

2. Se puede utilizar para aproximar varias distribuciones discretas de probabilidad y de esta forma evitar cálculos tediosos.

3. Proporciona la base para la inferencia estadística por su relación con el Teorema Central del Límite.

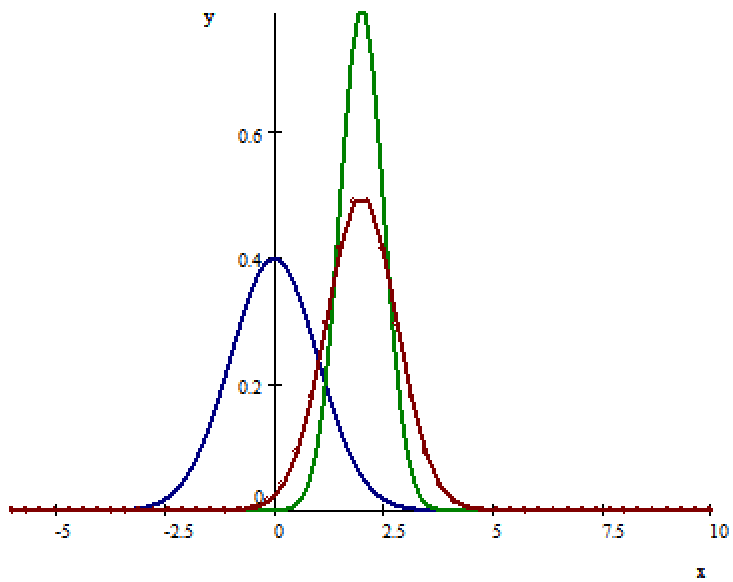
4. Muchos métodos de análisis estadístico suponen que la variable de interés sigue una distribución normal.

Función de densidad de probabilidad normal

Una variable aleatoria X sigue la distribución normal con media μ y varianza σ^2 (o, equivalentemente, con desviación estándar σ) lo cual representamos con $X \sim N(\mu, \sigma^2)$ si tiene la función de densidad de probabilidad

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad -\infty < x < \infty. \tag{2.29}$$

La gráfica de la función de densidad de probabilidad normal tiene forma acampanada y es simétrica con respecto a su media. La forma y la posición de la gráfica dependen de los valores de μ y σ, por lo cual habrá una gráfica diferente para cada combinación de μ y σ. Por ejemplo, la siguiente gráfica muestra tres distintas distribuciones normales:



Función de distribución de probabilidad normal

Nótese que fd de una variable aleatoria normal está dada por

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \int_{-\infty}^{x} \left\{\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right]\right\} dt. \tag{2.30}$$

Sin embargo, la integral anterior no puede evaluarse de manera cerrada, de modo que para calcular las probabilidades asociadas con una distribución normal, debe evaluarse la integral (2.30) numéricamente. Esto se ha hecho para un caso particular de la distribución normal, que tratamos a continuación:

Distribución normal estándar

Es una distribución normal que tiene una media $\mu = 0$ y una desviación estándar $\sigma = 1$. Se suele representar a una variable que sigue una distribución normal estándar con Z. Si Z se distribuye normal estándar, esto se representa con la notación $Z \sim N(0,1)$. La función de densidad de probabilidad de una variable normal estándar está dada por la expresión

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] \qquad -\infty < z < \infty \tag{2.31}$$

Tablas de la distribución normal estándar

Como hemos dicho, existen tablas para obtener las probabilidades asociadas a los valores de una variable normal estándar. Existen tres tipos de tablas para la distribución normal estándar:

1. Probabilidad acumulada desde $-\infty$ hasta el valor de Z mostrado en la tabla, z_0 (Tabla de cola izquierda).

2. Probabilidad entre 0 y el valor de Z mostrado en la tabla, z_0 (Tabla de probabilidad entre cero y z_0).

3. Probabilidad acumulada desde z_0 hasta ∞ (Tabla de cola derecha).

Aunque es posible utilizar cualquiera de los tres tipos de tablas para calcular probabilidades, en este curso utilizaremos la tabla del tipo 3. Para calcular probabilidades con este tipo de tabla debe hacerse lo siguiente:

1. Localizar en la primera columna de la tabla (que tiene como encabezado la letra z) la parte entera y el primer decimal de z_0.

2. Sobre ese renglón, ir hacia la columna cuyo encabezado sea el segundo decimal de z_0.

3. El número que se encuentra en el cruce de dichos renglón y columna representa la probabilidad de que Z sea mayor (o igual) que z_0, es decir $P(Z \geq z_0)$. Esto representa el área bajo la curva y sobre el eje horizontal entre z_0 e ∞.

Ejemplo 121 *Suponga que Z sigue una distribución normal estándar. Utilice una tabla que indique las probabilidades para dicha distribución y calcule lo siguiente:*
a) $P(Z \geq 1{,}64)$
b) $P(Z \geq 2{,}57)$
c) $P(Z < 1{,}64)$
d) $P(Z < 2{,}57)$
e) $P(0 \leq Z \leq 1{,}64)$
f) $P(0 \leq Z \leq 2{,}57)$
g) $P(Z \leq -1{,}64)$
h) $P(Z \leq -1{,}96)$
i) $P(1{,}33 \leq Z \leq 2{,}50)$

j) $P(0{,}25 \leq Z \leq 1{,}45)$
k) $P(-2{,}33 \leq Z \leq 0{,}53)$
l) $P(-0{,}53 \leq Z \leq 2{,}33)$

Ejemplo 122 *Suponga que Z sigue una distribución normal estándar. Utilice una tabla que indique las probabilidades para dicha distribución y calcule el valor de z_0 para tener las probabilidades indicadas:*
a) $P(Z \geq z_0) = 0{,}0202$
b) $P(Z \geq z_0) = 0{,}2877$
c) $P(Z \leq -z_0) = 0{,}0301$
d) $P(Z \leq -z_0) = 0{,}0073$
e) $P(Z \geq z_0) = 0{,}2855$
f) $P(Z \geq z_0) = 0{,}0156$
g) $P(-z_0 \leq Z \leq z_0) = 0{,}6826$
h) $P(-z_0 \leq Z \leq z_0) = 0{,}9812$

Estandarización de una distribución normal

Supongamos que la variable aleatoria X se distribuye normal con media μ y varianza σ^2, entonces la variable Z, definida como

$$Z = \frac{X - \mu}{\sigma} \tag{2.32}$$

se distribuye normal estándar.

El hecho de que la variable Z definida en 2.32 se distribuya normal estándar nos permite utilizar la tabla de esta distribución para calcular probabilidades asociadas con una variable X que siga una distribución normal cualquiera. Como se tiene que

$$P(X \leq x) = P\left(Z \leq \frac{X - \mu}{\sigma}\right), \tag{2.33}$$

entonces para calcular probabilidades para X, primero se estandarizan sus valores y luego se utiliza la tabla para la distribución normal estándar.

Ejemplo 123 *[Hoel, Paul G. Introduction to Mathematical Statistics. Ejercicio 3.46] Si X se distribuye normalmente con $\mu = 2$ y $\sigma = 1/3$ encuentre:*
a) $P(X > 3)$
b) $P(2 < X < 3)$
c) El valor de x_0 si $P(X > x_0) = 0{,}05$

Ejemplo 124 *[Hoel, Paul G. Introduction to Mathematical Statistics. Ejercicio 3.47] Suponga que la estatura, X, de estudiantes universitarios se distribuye de manera normal con $\mu = 175$ cm y $\sigma = 7{,}62$ cm. Calcule el porcentaje esperado de estudiantes tal que:*
a) $X < 165$.
b) $X > 183$.
c) $163 < X < 188$.

Ejemplo 125 *[Hoel, Paul G. Introduction to Mathematical Statistics. Ejercicio 3.48] Suponga que su calificación en un examen, en unidades estándar, z, es 0.8 y que las calificaciones se distribuyen de manera normal.*
a) ¿Qué porcentaje de los estudiantes que toman dicho examen se esperaría tuvieran una calificación mayor que la de usted?
b) Si 250 estudiantes presentaron el examen, ¿cuántos se esperaría tuvieran una calificación mayor que la suya?

Ejemplo 126 *[Hoel, Paul G. Introduction to Mathematical Statistics. Ejercicio 3.50] El coeficiente intelectual (CI) de una población de individuos se distribuye de manera aproximadamente normal con media 100 y desviación estándar 16. ¿Cuál es la probabilidad de que un individuo seleccionado aleatoriamente tuviera un CI*
a) de menos de 80?
b) de más de 140?
c) entre 90 y 110?
d) ¿Qué intervalo de valores, centrado en 100, incluiría al 50 % de los individuos?

Ejemplo 127 *[Hoel, Paul G. Introduction to Mathematical Statistics. Ejercicio 3.52] Los paquetes de cierto cereal indican que su contenido es de 12 onzas. La máquina de llenado está sujeta a errores, con una desviación estándar de 0.1 onzas.*
a) Si la máquina es ajustada para llenar 12.1 onzas de cereal, ¿a qué porcentaje de los paquetes le faltará al menos una onza?
b) Si el productor desea que a lo más al 2 % de los paquetes les falte una onza o más, ¿qué valor medio de llenado debería utilizarse?

2.3.3. Distribución gamma

Es una distribución contginua que ocurre frecuentemente en ciertas aplicaciones. Su nombre resulta de su relación con una función llamada función gamma.

Definición 128 ***Función Gamma.*** *La función gamma, denotada por $\Gamma(\alpha)$ para todo $\alpha > 0$, está dada por*

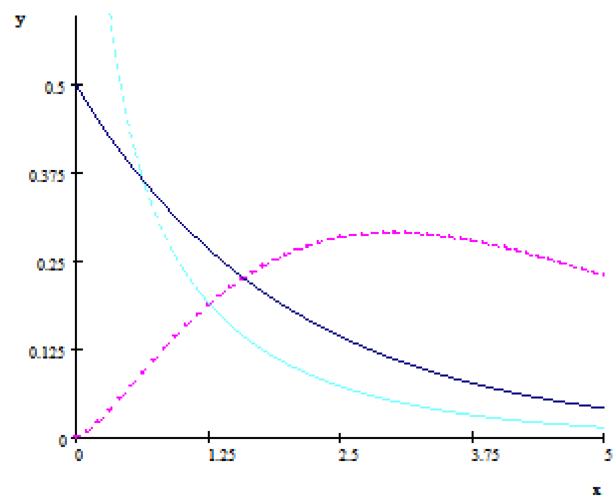
$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

Teorema 129 ***Propiedades de la función gamma.*** *La función gamma tiene las siguientes propiedades:*
1. $\Gamma(\kappa) = (\kappa - 1)\Gamma(\kappa - 1) \qquad \kappa > 0.$
2. $\Gamma(n) = (n-1)! \qquad n = 1, 2, ...$
3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$

Definición 130 *Distribución gamma. Una variable aleatoria X sigue la distribución gamma con parámetros $\alpha > 0$ y $\beta > 0$ si tiene la fdp*

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta} \qquad x > 0.$$

La media y la varianza de una variable aleatoria gamma están dadas por $E(X) = \alpha\beta$ y $Var(X) = \alpha\beta^2$.



2.3.4. Distribución exponencial

Es un caso particular de la distribución gamma cuando $\alpha = 1$.

Definición 131 ***Distribución exponencial.*** *La variable aleatoria continua X tiene una distribución exponencial, con parámetro β, si su fdp es*

$$f(x;\beta) = \frac{1}{\beta}e^{-x/\beta} \qquad \beta > 0,\ x > 0.$$

El valor esperado de una variable aleatoria exponencial es $E(X) = \beta$ y su varianza es $Var(X) = \beta^2$.

La distribución exponencial guarda relación con la distribucion Poisson. Recordemos que un experimento Poisson es aquel en que ciertos eventos de interés ocurren aleatoriamente en una unidad de tiempo, distancia, masa, etc. Por lo tanto, la distribución Poisson se utiliza la probabilidad de que en un intervalo dado (de tiempo, distancia, masa, etc.) ocurra una cantidad x de eventos (llegadas a un crucero, fallas de componentes, accidentes de tránsito, etc.). La distribución exponencial, por otro lado, se utiliza para calcular la probabilidad de que entre dos ocurrencias sucesivas transcurra un intervalo de tamaño x o mayor o menor. Los parámetros λ, de la distribución Poisson, y β, de la distribución exponencial, se relacionan mediante la igualdad $\lambda = \frac{1}{\beta}$.

Ejemplo 132 *[Walpole, Ronald E; Myers, Raymond H.; Myers, Sharon L. Probabilidad y estadística para ingenieros. Ejemplo 6.18* $\approx$*] Suponga que las llamadas telefónicas que llegan a un conmutador particular siguen un proceso de Poisson con un promedio de cinco llamadas por minuto. ¿Cuál es la probabilidad de que entre llamadas sucesivas transcurran*
a) menos de un minuto? (0.1813)
b) más de un minuto? (0.8187)
c) más de dos minutos? (0.6703)

Ejemplo 133 *[Bain, L. J.; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Ejemplo 3..3.2. p .116] Suponga que cierto componente de estado sólido tiene un tiempo de falla (en horas)* $X \sim Exp\,(100)$ *. Encuentre la probabilidad de que el componente sobreviva*
a) menos de 50 horas. (0.3935)
b) entre 50 y 100 horas. (0.6321-0.3935)
c) más de 100 horas. (0.3679)

Ejemplo 134 *[Bain, L. J.; Engelhardt, M. Introduction to Probability and Mathematical Statistics. Ejemplo 3..3.2. p .116* $\approx$*] Suponga que el tiempo (en horas) hasta la falla de un transistor es una variable aleatoria* $X \sim Exp\,(250)$ *. Halle la probabilidad de que*
a) $X > 40$. *(0.8521)*
b) $X > 275$. *(0.3329)*

Ejemplo 135 *[Walpole, Ronald E; Myers, Raymond H.; Myers, Sharon L. Probabilidad y estadística para ingenieros. Ejercicio 6.10.4* $\approx$*] En cierta ciudad, el consumo diario de agua (en millones de litros) sigue aproximadamente una distribución exponencial con* $\beta = 1/3$. *Si la capacidad diaria de dicha ciudad es igual a nueve (millones de litros de agua), ¿cuál es la probabilidad de que en cualquier día el suminstro de agua sea inadecuado?. (0.0498)*

Ejemplo 136 *[Walpole, Ronald E; Myers, Raymond H.; Myers, Sharon L. Probabilidad y estadística para ingenieros. Ejercicio 6.10.4] El tiempo, en horas, que toma reparar una bomba de calor es una variable aleatoria* X *que tiene un dsitribución exponencial con parámetro* $\beta = 1/2$. *¿Cuál es la probabilidad de que la siguiente llamada de servicio requiera*
a) a lo más una hora para reparar la bomba de calor? (0.3935)
b) al menos de dos horas para reparar la bomba de calor? (0.3679).

Bibliografía

[Ash, 2000] Robert B. Ash. *Probability and Measure Theory*. Harcourt Academic Press, Estados Unidos de América, 2000.

[Bain and Engelhardt, 1992] Lee J. Bain and Max Engelhardt. *Introduction to Probability and Mathematical Statistics*. PWS Kent Publishing Company, EUA, 1992.

[Centro de Investigación En Matemáticas, A. C., 1998] Centro de Investigación En Matemáticas, A. C. Rincón de problemas. elementales. http://www.cimat.mx:88/ gil/rincon/elementales/, 1998. Consultado el 16 de febrero de 2007.

[Feller, 1978] William Feller. *Introducción a la Teoría de Probabilidades Y Sus Aplicaciones*. Limusa, México, 1978.

[Geiss and Geiss, 2006] Christel Geiss and Stefan Geiss. An introduction to probability theory. http://www.math.jyu.fi/ geiss/scripts/introduction-probability.pdf, 2006.

[Walpole and Myers, 1992] Ronald E. Walpole and Raymond H. Myers. *Probabilidad Y Estadística*. McGraw-Hill Interamericana de México, México, 1992.

[Young, 2007] Alister Young. Probability and statistics II. worked examples. http://www.ma.ic.ac.uk/ ayoung/m2s1/WorkedExamples1.pdf, 2007. Consultado el 16 de febrero de 2007.