Tema 1. Introducción a la teoría de la probabilidad

UGR

Granada, noviembre/diciembre 2024

Índice

1. *Perspectivas sobre la teoría de la probabilidad*

Perspectivas:

- frecuentista
- subjetiva

Perspectiva frecuentista: equipara la probabilidad de ocurrencia de un cierto evento de un experimento a la frecuencia relativa de aparición del mismo en una secuencia de repeticiones del experimento.

Ejemplo 1: lanzamiento de dado sin trucar. Se considera el evento obtener un 3. Si se lanza el dado un número suficiente de veces ($\boldsymbol{n}$) se observa que este resultado aparece $\boldsymbol{n}_3$ veces. Es decir:

$$\frac{n_3}{n} \approx \frac{1}{6}$$

Perspectiva frecuentista

Ejemplo 2: sacar una pica de una baraja francesa al extraer de forma aleatoria una carta. La baraja se compone de 52 cartas y hay 13 de cada palo. Si se realiza el experimento n veces se obtiene un resultado positivo en n_p ocasiones. La frecuencia relativa de ocurrencia del evento es por tanto:

$$p_p = \frac{n_p}{n} \approx \frac{13}{52}$$

Esta interpretación parte de *asumir que hay un proceso estocástico que puede repetirse tantas veces como deseemos y del que pueden contarse las apariciones de los eventos de interés*.

Perspectiva subjetiva: hay situaciones en que no es posible especificar la frecuencia relativa de ocurrencia de un determinado evento. Por ejemplo:

- imaginemos que queremos estimar la probabilidad de que un alumno supere una determinada materia (suponemos que no aporta información considerar los registros de años anteriores).
- un profesor puede estimar que esta probabilidad es p, en base a su juicio personal y a la información sobre la forma en que el alumno ha ido trabajando (experiencia + información indirecta).

Estas valoraciones se conocen como probabilidades subjetivas.

Una forma alternativa de interpretar la probabilidad subjetiva de que el alumno supera la materia indicada es considerar las dos apuestas siguientes:

- si el alumno supera la asignatura (con probabilidad estimada $p = n/100$), se reciben 100 euros
- hay una urna con 100 bolas de las que n son rojas y las demás son negras. Si se realiza una extracción y la bola es roja, se reciben 100 euros

Si la estimación subjetiva es fiable (me merece confianza) entonces tengo que sentir indiferencia por elegir una u otra apuesta. A la hora de asignar probabilidades de forma subjetiva se usan preguntas de este estilo para facilitar la tarea de los expertos en el problema.

Índice

Índice

Se consideran aquí algunos conceptos básicos:

Conceptos básicos

- **experimento**: cualquier tipo de proceso para el que hay incertidumbre sobre el resultado.
- **espacio muestral**: conjunto de posibles resultados del experimento. Se asume que incluye todos los resultados que pueden ocurrir y que los resultados son exclusivos entre sí. Esto garantiza que el experimento finalice en uno y solo uno de los resultados incluidos en el espacio muestral.

Ejemplo:

- experimento: lanzamiento de dado.
- espacio muestral: $S = \{1, 2, 3, 4, 5, 6\}$

Conceptos básicos

- **evento**: cualquier subconjunto del espacio muestral. Por ejemplo, en el caso del lanzamiento de dado, el evento ***obtener un valor par*** puede definirse como:

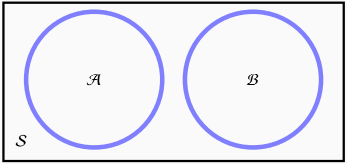
$$\mathcal{A} = \{2, 4, 6\} \subseteq \{1, 2, 3, 4, 5, 6\}$$

  En general se dice que un evento es cierto (se cumple) si el resultado del experimento contiene a cualquiera de sus elementos. Cuando un evento contiene un único elemento, se habla de **resultado**.

Conceptos básicos

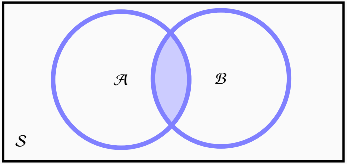
- **grado de certeza**: para medir el grado de certeza sobre un determinado experimento se asigna una probabilidad $P(\mathcal{A})$ a todos los eventos posibles $\mathcal{A} \subseteq \mathcal{S}$. Estos valores deben cumplir:
  - **Axioma 1**: $P(\mathcal{S}) = 1$
  - **Axioma 2**: para todo $\mathcal{A} \subseteq \mathcal{S}$, $P(\mathcal{A}) >= 0$
  - **Axioma 3**: si $\mathcal{A} \subseteq \mathcal{S}, \mathcal{B} \subseteq \mathcal{S}$, y $\mathcal{A} \cap \mathcal{B} = \emptyset$ entonces

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$$

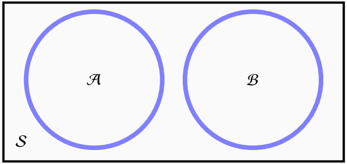


Y en el caso en se cumpla que $\mathcal{A} \subseteq \mathcal{S}, \mathcal{B} \subseteq \mathcal{S}$, y $\mathcal{A} \cap \mathcal{B} \neq \emptyset$ entonces

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



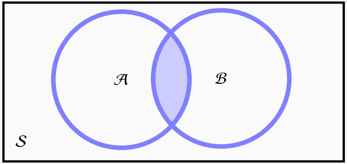
Ejemplo de intersección vacía:



Experimento de lanzamiento de dado y $\mathcal{B} = \{3\}$, $\mathcal{A} = \{2, 4, 6\}$. En este caso, se trata claramente de eventos disjuntos y la probabilidad de ocurrencia de alguno de los dos viene determinada por:

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) = \frac{1}{6} + \frac{3}{6} = \frac{4}{6}$$

Ejemplo de intersección no vacía:



Baraja francesa y extracción de carta, con $\mathcal{A} = \{pica\}$, $\mathcal{B} = \{rey\}$. En este caso, no son disjuntos (podemos sacar el rey de picas). El cálculo ahora es:

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B}) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}$$

Índice

Normalmente, cuando se da una asignación de probabilidad a un evento $\mathcal{A}$ se asume alguna información adicional, aunque no se indique de forma explícita. Por ejemplo, si decimos que al lanzar un dado la probabilidad de obtener un 6 es $1/6$, estamos asumiendo que se trata de un dado sin trucar.

En realidad, deberíamos expresarlo de la siguiente forma:

- **si el dado no está trucado**, entonces la probabilidad de obtener un 6 será $1/6$

De esta forma, la afirmación sobre el valor de probabilidad está **condicionada** a información adicional. En este tipo de situaciones se habla de probabilidades **condicionales** o **condicionadas**.

La notación que se usa es:

$$P(\mathcal{A}|\mathcal{B}) = p$$

que se interpreta como la probabilidad de ocurrencia del evento $\mathcal{A}$ dado que ha ocurrido previamente el evento $\mathcal{B}$. Debe quedar claro que:

- no debe entenderse que siempre que ocurra $\mathcal{B}$ la probabilidad de ocurrencia de $\mathcal{A}$ es p
- **sino que debe interpretarse de la siguiente forma**: si ocurre $\mathcal{B}$ y todo lo demás es irrelevante para la ocurrencia de $\mathcal{A}$, entonces la probabilidad de $\mathcal{A}$ es p

Supongamos que se dispone de la asignación de todos los valores de probabilidad del espacio muestral $\mathcal{S}$. Podemos preguntarnos si estos valores pueden usarse para calcular $P(\mathcal{A}|\mathcal{B})$. Para obtener este valor observamos que:

- dado que ha ocurrido el evento $\mathcal{B}$, todos sus posibles resultados son elementos de $\mathcal{B}$
- aquellos resultados que pertenecen también a $\mathcal{A}$ hacen que $\mathcal{A} \cap \mathcal{B}$ también sea cierto
- el objetivo es por tanto calcular la probabilidad del evento $P(\mathcal{A} \cap \mathcal{B})$ asumiendo que se ha producido el evento $\mathcal{B}$

Dados dos eventos $\mathcal{A}$ y $\mathcal{B}$, con $P(\mathcal{B}) > 0$, entonces la probabilidad condicional de $\mathcal{A}$ dado $\mathcal{B}$ es:

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

Supongamos que sabemos que el resultado del lanzamiento de un dado ha producido como resultado un valor par, $\mathcal{B} = \{2, 4, 6\}$. Podemos ahora calcular la probabilidad de que el resultado haya sido 4, $\mathcal{A} = \{4\}$.

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})} = \frac{1/6}{3/6} = \frac{1}{3}$$

Por supuesto, al trabajar con probabilidades condicionales es posible condicionar sobre más de un evento:

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

Índice

2. Fundamentos de la teoría de la probabilidad

La expresión anterior

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

puede reescribirse y se obtiene la **regla fundamental** del cálculo de probabilidades:

$$P(\mathcal{A}|\mathcal{B})P(\mathcal{B}) = P(\mathcal{A} \cap \mathcal{B})$$

Esta regla nos permite calcular el valor de la probabilidad conjunta $P(\mathcal{A} \cap \mathcal{B})$ a partir del valor de la probabilidad condicional $P(\mathcal{A}|\mathcal{B})$ y de $P(\mathcal{B})$. También puede usarse en el caso de condicionar sobre otro evento $\mathcal{C}$:

$$P(\mathcal{A}|\mathcal{B} \cap \mathcal{C})P(\mathcal{B}|\mathcal{C}) = P(\mathcal{A} \cap \mathcal{B}|\mathcal{C})$$

Ya que $P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{B} \cap \mathcal{A})$, la aplicación de la regla fundamental del cálculo lleva a:

$$P(\mathcal{A}|\mathcal{B})P(\mathcal{B}) = P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{B} \cap \mathcal{A}) = P(\mathcal{B}|\mathcal{A})P(\mathcal{A})$$

Y relacionando los extremos obtenemos la expresión del **Teorema de Bayes**, que resulta fundamental para los modelos gráficos probabilísticos.

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{B}|\mathcal{A})P(\mathcal{A})}{P(\mathcal{B})}$$

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{B}|\mathcal{A})P(\mathcal{A})}{P(\mathcal{B})}$$

Esta expresión permite actualizar nuestra creencia sobre un evento $\mathcal{A}$ a partir de información sobre otro evento $\mathcal{B}$.

- supongamos que al principio se conoce el valor de $P(\mathcal{A})$: **probabilidad a priori**
- al observar el valor de $\mathcal{B}$ actualizamos nuestra creencia sobre la ocurrencia de $\mathcal{A}$, $P(\mathcal{A}|\mathcal{B})$: **probabilidad a posteriori**
- el término $P(\mathcal{B}|\mathcal{A})$ se conoce como **verosimilitud de $\mathcal{A}$ dado $\mathcal{B}$**

Para entender el nombre del término $P(\mathcal{B}|\mathcal{A})$ (verosimilitud de $\mathcal{A}$ dado $\mathcal{B}$) consideremos el siguiente ejemplo: dos enfermedades a_1 y a_2 son susceptibles de producir un síntoma b. Sabemos $P(b|a_1) = 0.9$ y $P(b|a_2) = 0.3$. Asumamos también que $P(a_1) = P(a_2)$. Considerando el teorema de Bayes:

$$P(a_1|b) = \frac{P(b|a_1)P(a_1)}{P(b)} = 0.9\ \frac{P(a_1)}{P(b)}$$
$$P(a_2|b) = \frac{P(b|a_2)P(a_2)}{P(b)} = 0.3\ \frac{P(a_2)}{P(b)}$$

En este caso podemos afirmar que $P(a_1|b)$ es tres veces más probable que $P(a_2|b)$. El valor de esta proporción viene dado por los valores de $P(b|a_i)$. Esto explica por qué el término $P(\mathcal{B}|\mathcal{A})$ se denomina verosimilitud de $\mathcal{A}$ dado $\mathcal{B}$.

Índice

Independencia

Los eventos $\mathcal{A}$ y $\mathcal{B}$ son independientes si la ocurrencia de uno de ellos no aporta información sobre el otro:

$$P(\mathcal{A}|\mathcal{B}) = P(\mathcal{A})$$

La noción de independencia es simétrica, ya que cuando dos eventos son independientes, la regla fundamental del cálculo puede escribirse de la siguiente forma:

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

Y entonces:

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

Independencia condicional

Los eventos $\mathcal{A}$ y $\mathcal{B}$ son independientes dado el evento $\mathcal{C}$ si

$$P(\mathcal{A}|\mathcal{B} \cap \mathcal{C}) = P(\mathcal{A}|\mathcal{C})$$

En este caso, la regla fundamental del cálculo puede escribirse de la siguiente forma:

$$P(\mathcal{A} \cap \mathcal{B}|\mathcal{C}) = P(\mathcal{A}|\mathcal{C})P(\mathcal{B}|\mathcal{C})$$

Una vez observada la ocurrencia de $\mathcal{C}$ tener información sobre la ocurrencia de $\mathcal{B}$ no aporta información adicional sobre $\mathcal{A}$ y viceversa.

Índice

3. *Cálculo de probabilidad con variables*

Imaginemos un problema en que se consideran varios experimentos a la vez. Resulta conveniente en este caso hablar de **variables**. Cada variable representa a un experimento y cada resultado se corresponderá con un estado de la variable correspondiente.

El conjunto de posibles valores de una variable A se denota como $sp(A)$. De forma análoga a como se indicaba antes, los estados han de ser mutuamente exclusivos entre sí. De forma general se usan letras mayúsculas para denotar variables y letras minúsculas para estados.

Para una variable A con estados a_1 .... a_n, expresamos nuestra incertidumbre sobre su estado mediante una distribución de probabilidad $P(A)$ sobre sus estados:

$$P(A) = \{p(a_1) \ .... \ p(a_n)\}, p(a_i) \geq 0$$

$$\sum_{i=1}^{n} p(a_i) = p(a_1) + \ .... \ + p(a_n) = 1$$

donde $p(a_i)$ es la probabilidad de que la variable A tenga el estado a_i. Como notación simplificada usaremos $p(A = a_i) = p(a_i)$.

El concepto de probabilidad condicional también puede aplicarse a variables. Si la variable B tiene como estados $\{b_1 \ .... \ b_m\}$ entonces $P(A|B)$ se cuantifica mediante $n \times m$ valores (recordemos que A tiene n valores), de forma que $p(a_i|b_j)$ representa la probabilidad de ocurrencia de a_i tras haber ocurrido b_j.

Como $P(A|B)$ define una distribución de probabilidad específica para cada valor b_j, entonces, por el axioma 1, debe cumplirse que:

$$\sum_{i=1}^{n} p(A = a_i|B = b_j) = 1, \forall b_j$$

Es decir, dado un valor concreto de B, debe darse a la fuerza alguno de los valores posibles de A.

3. Cálculo de probabilidad con variables

Ejemplo:

	b_1	b_2	b_3
a_1	0.4	0.3	0.6
a_2	0.6	0.7	0.4

Esta tabla muestra la distribución $P(A|B)$ para una variable A con dos valores y otra variable B con tres valores. Imaginemos que $B = b_1$, entonces, como se ha comentado antes, $P(A|B = b_1)$ define una distribución de probabilidad que ha de sumar 1 (primera columna de la tabla):

$$p(A = a_1|B = b_1) + p(A = a_2|B = b_1) = 0.4 + 0.6 = 1$$

La probabilidad de observar resultados simultáneos para varias variables puede representarse mediante su **probabilidad conjunta**: mediante ella se asigna un valor de probabilidad a cada posible evento de la forma $(a_i, b_j, \ldots)$. En el caso de dos variables:

	b_1	b_2	b_3
a_1	0.16	0.12	0.12
a_2	0.24	0.28	0.08

En este caso, $P(A, B)$ consta de $n \times m$ parámetros. Como todos los elementos del espacio muestral son mutuamente exclusivos, debe cumplirse:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p(A = a_i, B = b_j) = 1$$

3. Cálculo de probabilidad con variables

La regla fundamental del cálculo con variables se expresa de la siguiente forma:

$$P(A, B) = P(A|B)P(B)$$
$$P(A, B|C) = P(A|B, C)P(B|C)$$

Importante: a partir de una distribución conjunta pueden calcularse las tablas de las distribuciones marginales. Por ejemplo, dada la tabla para $P(A, B)$ entonces la probabilidad $P(A)$ puede calcularse considerando que las salidas de B pueden ocurrir de forma conjunta con cada estado a_i de A. Por ejemplo, para a_i habrá que considerar los eventos $(a_i, b_1)...(a_i, b_m)$:

$$p(A = a_i) = \sum_{j=1}^{m} p(A = a_i, B = b_j)$$

Esta operación se denomina **marginalización** y se indica que la variable B se marginaliza de $P(A, B)$. La notación es:

$$P(A) = \sum_{B} P(A, B)$$

Usando la tabla anterior, tenemos:

$$p(a_1) = p(a_1|b_1) + p(a_1|b_2) + p(a_1|b_3) = 0.16 + 0.12 + 0.12 = 0.4$$
$$p(a_2) = p(a_2|b_1) + p(a_2|b_2) + p(a_2|b_3) = 0.24 + 0.28 + 0.08 = 0.6$$

De forma que $P(A) = (0.4, 0.6)$.

El teorema de Bayes para variables, se expresa ahora de la siguiente forma:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A,B)}{\sum_B P(A,B)}$$

Y también es posible condicionar sobre otra(s) variable(s):

$$P(B|A,C) = \frac{P(A|B,C)P(B|C)}{P(A|C)} = \frac{P(A,B|C)}{\sum_B P(A,B|C)}$$

Independencia condicional

Se dice que dos variables A y B son condicionalmente independientes dada otra variable C, si

$$p(a_i|b_j, c_k) = p(a_i|c_k)$$

para cada $a_i \in sp(A), b_j \in sp(B), c_k \in sp(C)$

Generalmente se usa la notación $P(A|B, C) = P(A|C)$

Esto indica que cuando se observa el valor de C entonces tener información sobre el valor de B ya no aporta información adicional sobre el valor de A. Observad que esto implica que la independencia se cumple para cada posible valor de C.

3. Cálculo de probabilidad con variables

Cuando dos variables A y B son condicionalmente independientes dado que se conoce el valor de la variable C, entonces la regla fundamental puede simplificarse:

$$P(A, B|C) = P(A|C)P(B|C)$$

En la expresión anterior, se multiplican dos probabilidades condicionales sobre dominios diferentes. Pero la forma de cálculo es sencilla si consideramos:

$$p(A = a_i, B = b_j|C = c_k) = p(a_i|c_k)p(b_j|c_k)$$

Es decir, se van recorriendo las posibles combinaciones de valores de todas las variables. Y dada una combinación, se proyectan sus coordenadas sobre las distribuciones que participan en la operación.

Ejercicio 1: realizar la multiplicación de las distribuciones $P(A|C)$ y $P(B|C)$

	c_1	c_2	c_3
a_1	0.4	0.3	0.6
a_2	0.6	0.7	0.4

	c_1	c_2	c_3
b_1	0.2	0.9	0.3
b_2	0.05	0.05	0.2
b_3	0.75	0.05	0.5

Ejercicio 2: vamos a aplicar los conceptos vistos con un ejemplo en que se parte de la distribución $P(A, B, C)$ (conjunta)

- queremos calcular como varía mi creencia sobre B al saber que $A = a_2$ y $C = c_1$. Es decir, queremos calcular $P(B|A = a_2, C = c_1)$
- también el valor de $P(B|A = a_2, C)$

La tabla con la distribución $P(A, B, C)$ es:

	b_1			b_2			b_3		
	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
a_1	0	0.05	0.05	0.05	0.05	0	0.05	0.05	0.05
a_2	0.1	0.1	0	0.1	0	0.1	0.2	0	0.05

Índice

Puede ocurrir que los cálculos probabilísticos generen distribuciones no normalizadas. Por esto se habla de **potenciales**. Un potencial ϕ es una función de valores reales definido de la siguiente forma:

$$\phi = sp(\mathcal{X}) \rightarrow \Re$$

Recordemos que $sp(\mathcal{X})$ denota el conjunto formado por todas las combinaciones de posibles valores de las variables contenidas en $\mathcal{X}$. Por otra parte, el dominio de un potencial se denota mediante dom. En el ejemplo anterior, $dom(\phi) = \{A, B, C\}$

4. *Cálculo con potenciales*

Los potenciales pueden multiplicarse. Esta operación presenta las siguientes propiedades:

1. $dom(\phi_1\phi_2) = dom(\phi_1) \cup dom(\phi_2)$
2. $\phi_1\phi_2 = \phi_2\phi_1$
3. $(\phi_1\phi_2)\phi_3 = \phi_1(\phi_2\phi_3)$
4. existencia de potencial unidad: sus valores son todos igual a 1. Se representa como ϕ_I y se cumple que actúa como el elemento neutro: $\phi_I\phi = \phi$

El operador de marginalización puede aplicarse a potenciales, de forma que $\sum_A \phi$ es un potencial definido sobre $dom(\phi) \setminus A$. La marginalización también es conmutativa:

$$\sum_A \sum_B \phi = \sum_B \sum_A \phi$$

5. Propiedad del potencial unidad: $\sum_A P(A|\mathcal{V}) = \phi_I$ (donde $\mathcal{V}$ representa un conjunto de variables)
6. si $A \notin dom(\phi_1)$, entonces $\sum_A \phi_1\phi_2 = \phi_1 \sum_A \phi_2$. La marginalización también se denomina proyección. Por ejemplo, en el caso $\sum_{A,B} \phi(A, B, C) = \phi^{\downarrow C}$
7. $(\phi^{\downarrow \mathcal{W}})^{\downarrow \mathcal{V}} = (\phi^{\downarrow \mathcal{V}})^{\downarrow \mathcal{W}}$ (propiedad conmutativa)
8. si $dom(\phi_1) \subseteq \mathcal{V}$, entonces $(\phi_1\phi_2)^{\downarrow \mathcal{V}} = \phi_1(\phi_2^{\downarrow \mathcal{V}})$

Índice

Sea $\mathcal{S}$ un espacio muestral. Una variable aleatoria es una función real definida de la siguiente forma:

$$V : \mathcal{S} \rightarrow \Re$$

Por ejemplo, considerando el lanzamiento de un dado, se obtiene una ganancia de 1 si sale un 4 o más y se pierde 1 si sale 3 o menos. Esta función toma valor -1 para los valores $\{1, 2, 3\}$ y 1 en $\{4, 5, 6\}$.

El valor medio de una variable aleatoria V definida en $\mathcal{S}$ se define como:

$$\mu(V) = \sum_{s \in \mathcal{S}} V(s) P(s)$$

Este valor también se conoce como **valor esperado**.

La varianza de V se define de la siguiente forma:

$$\sigma^2(V) = \sum_{s \in \mathcal{S}} (V(s) - \mu(V))^2 P(s)$$

En el ejemplo anterior del lanzamiento del dado:

$$\mu = -1(\frac{1}{6}) - 1(\frac{1}{6}) - 1(\frac{1}{6}) + 1(\frac{1}{6}) + 1(\frac{1}{6}) + 1(\frac{1}{6}) = 0$$
$$\sigma^2 = 3(1-0)^2(\frac{1}{6}) + 3(1-0)^2(\frac{1}{6}) = 1$$

Índice

Consideremos ahora el experimento de lanzamiento de un dardo sobre un cuadrado de 1×1 metro. Las posibles salidas son los puntos (x, y) del cuadrado unidad. A observar:

- la probabilidad de cualquier punto concreto es 0
- la distribución de probabilidad se asignará a subconjuntos de puntos
- hay que considerar que la masa de probabilidad debe repartirse sobre el espacio muestral
- por ejemplo, podemos particionar el cuadrado en $n \times n$ parcelas, cada una con superficie $1/n^2$, asignando un valor de probabilidad x a cada a parcela

En general, si $\mathcal{S}$ es un espacio muestral continuo, la función de densidad asociada es un función real no negativa f, que cumple:

$$\int_{\mathcal{A}} f(s)ds = P(\mathcal{A}), \forall \mathcal{A} \subseteq \mathcal{S}$$

En particular:

$$\int_{\mathcal{S}} f(s)ds = 1$$

Cuando $\mathcal{S}$ es un intervalo $[a, b]$ (posiblemente infinito), podemos calcular la media y la varianza de la siguiente forma:

$$\mu = \int_a^b x f(x)dx$$

$$\sigma^2 = \int_a^b (\mu - x)^2 f(x)dx$$