Universidad de Chile
Facultad De Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática
MA3403 - Probabilidades y Estadística

Conteo y Combinatoria

Alberto Vera Azócar, albvera@ing.uchile.cl

1. Nociones Generales

Queremos estudiar un fenómeno aleatorio, del que a pesar de que no podemos predecir cuál será su resultado, sí podemos decir los posibles valores. Por ejemplo si se lanza un dado corriente no se puede decir con seguridad qué número saldrá, pero sí sabemos que sale un número en el conjunto $\{1, 2, \cdots, 6\}$ y más aún nos arriesgamos a decir que si lanzamos el dado muchas veces cada número aparecerá aproximadamente $\frac{1}{6}$ de las veces. Crearemos una teoría consistente que nos permita llegar a este tipo de conclusiones.

Definición 1.1. *Decimos que el espacio muestral es el conjunto de todos los resultados posibles de nuestro experimento, lo denotamos por* Ω.

Definición 1.2. *Un evento es cualquier conjunto* $A \subseteq \Omega$.

Nuestro objetivo entonces, dado el espacio Ω, es poder calcular $|\Omega|$ (lo que se interpreta como "cuantos resultados posibles existen"), más aún, para cada $A \subseteq \Omega$ queremos obtener $|A|$, i.e. "de cuantas formas puede suceder A".

2. Teoría de conjuntos

Las siguientes propiedades se demostraron en el curso de álgebra, por lo que se dan por conocidas:

Proposición 2.1. *Sean* $(A_n)_{n\in\mathbb{N}}$, B *conjuntos. Es cierto que:*

1. $\left(\bigcup_{n\in\mathbb{N}} A_n\right)^c = \bigcap_{n\in\mathbb{N}} A_n^c$
2. $\left(\bigcap_{n\in\mathbb{N}} A_n\right)^c = \bigcup_{n\in\mathbb{N}} A_n^c$
3. $B \cup \left(\bigcap_{n\in\mathbb{N}} A_n\right) = \bigcap_{n\in\mathbb{N}}(B \cup A_n)$
4. $B \cap \left(\bigcup_{n\in\mathbb{N}} A_n\right) = \bigcup_{n\in\mathbb{N}}(B \cap A_n)$

Ahora nos centraremos en el caso de conjuntos finitos, donde es factible obtener el cardinal. Es frecuente que utilicemos particiones para los resultados, por lo que la siguiente proposición es importante.

Definición 2.1 (Partición)**.** *Sean* $A, (A_n)_{n\in\mathbb{N}}$ *conjuntos. Diremos que* (A_n) *es partición de* A *si se cumple que* $\forall\, n \neq m\ A_n \cap A_m = \varnothing$ *y* $\dot\bigcup_{n\in\mathbb{N}} A_n = A$.

Proposición 2.2. *Sea A conjunto finito y sea $(A_n)_{n\in\mathbb{N}}$ partición de A, entonces*

$$|A| = \sum_{n\in\mathbb{N}} |A_n|$$

Ejemplo 2.1. *Consideremos el experimento donde se lanza una moneda, si ésta es cara se lanza otra moneda y continúa el experimento, si sale cruz se finaliza. En otras palabras se lanza una moneda hasta observar una cruz.*

Podemos describir el conjunto Ω como tuplas ordenadas, de la siguiente forma:

$$\Omega = \{(cruz), (cara, cruz), (cara, cara, cruz), (cara, cara, cara, cruz), \cdots\}$$

Sea A el evento "se observan a lo más 4 caras", nos interesa $|A|$, para esto proponemos la siguiente partición de A: sea A_n el evento "se observan exactamente n caras", es claro que $(A_n)_{n=0}^{4}$ es partición de A y por la proposición 2.2

$$|A| = \sum_{n=0}^{4} |A_n|$$

ahora, cada A_n puede pasar de sólo una forma (salen n caras seguidas y luego una cruz), sigue que $|A_n| = 1$ y por lo tanto $|A| = 5$.

Alternativamente, podemos definir $\Omega = \mathbb{N}$, en efecto: si cada elemento de Ω representa el número de caras obtenidas antes de finalizar el experimento, entonces podemos identificar

$$\begin{aligned}
(cruz) &\rightarrow 0 \\
(cruz, cara) &\rightarrow 1 \\
(cruz, cruz, cara) &\rightarrow 2 \\
&\vdots
\end{aligned}$$

Visto de esta forma tendríamos $A_0 = \{0\}$, $A_1 = \{1\}, \cdots$ y $A = \{0, 1, 2, 3, 4\}$.

La última forma de mirar Ω es mucho más sencilla para hacer cálculos, pero quizá sea menos intuitiva.

El ejemplo anterior nos deja una moraleja: hay varias formas de representar Ω, pero si somos consistentes llegaremos siempre al mismo resultado.

3. Principios de conteo

Daremos las nociones para 2 experimentos, queda propuesto extender los resultados para n experimentos.

Definición 3.1. *Decimos que el experimento E_1 es independiente del E_2 si los resultados de E_1 no afectan a los resultados de E_2.*

La definición anterior parte de que tenemos modelado el problema, por ejemplo si E_1 es lanzar una moneda y E_2 es lanzar un dado, aceptamos que E_1 es independiente de E_2 como convención.

3.1. Principio aditivo

Si E_1 y E_2 son experimentos independientes, donde E_1 tiene n_1 resultados posibles y E_2 tiene n_2 resultados posibles, entonces la acción de elegir un resultado de algún experimento tiene n_1+n_2 formas de ser realizada.

Ejemplo 3.1. *Si en la estantería hay 3 libros de cálculo y 4 libros de computación, entonces se puede elegir un libro de 7 formas distintas. Podríamos denotar E_1 como el experimento de elegir un libro de cálculo con $n_1 = 3$ y E_2 como el experimento de elegir un libro de computación con $n_2 = 4$.*

3.2. Principio multiplicativo

Si E_1 y E_2 son experimentos independientes, con n_1 y n_2 resultados posibles respectivamente, entonces la acción de elegir un resultado de E_1 y un resultado de E_2 se puede hacer de $n_1 \cdot n_2$ formas posibles.

Ejemplo 3.2. *En el restaurante se puede elegir entre 4 entradas, 5 platos de fondo y 3 postres. Luego se pueden formar 60 menús distintos.*

4. Fórmulas combinatoriales

Ahora nos abocaremos en aprender a contar, para esto la siguiente expresión aparecerá frecuentemente, cuya interpretación daremos más adelante.

Definición 4.1. *Sean $n, k \in \mathbb{N}$, definimos*

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

4.1. Permutaciones

Tenemos n bolitas enumeradas del 1 al n, queremos saber de cuántas formas podemos ordenarlas. Por supuesto el resultado ①②③ cuenta distinto al resultado ③②①.

Esto se puede hacer de $n!$ formas distintas, pues la primera bolita se puede elegir de n formas, la segunda de $n-1$ y así sucesivamente.

4.2. Extracción con orden

Tenemos n bolitas enumeradas del 1 al n, sacamos $k \leq n$ una tras otra y nos preguntamos por el número de resultados posibles. Como sacamos una tras otra podemos distinguir el orden, i.e. es distinto extraer ①② que ②①.

La primera bolita se puede elegir de n formas, la segunda de $n-1$ y así, pero llegamos hasta la bolita k que se puede elegir de $n-(k-1)$ formas, es decir:

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

Ejemplo 4.1. *De un curso de 100 alumnos se deben elegir 50 para ir a una ceremonia donde se sentarán por orden de llegada ¿cuántos resultados posibles hay?*

Pensaremos que de los 100 "extraemos" 50 en orden y los sentamos según vayan saliendo. Como los alumnos son totalmente distinguibles (y el orden claramente importa), usamos la fórmula anterior y tenemos que se pueden obtener $\frac{100!}{50!} \approx 10^{93}$ resultados distintos.

4.3. Extracción sin orden

Tenemos n bolitas enumeradas del 1 al n, sacamos $k \leq n$ de una sola vez. Es decir, no sabemos cuál bolita salió primero y cuál después, luego extraer ①④② es exactamente lo mismo que ④①②.

También se puede pensar que se extraen de una en una, pero sin anotar el orden en que fueron saliendo.

Si consideráramos el orden, por la fórmula anterior sabemos que el resultado es $\frac{n!}{(n-k)!}$, pero esta expresión considera las permutaciones como resultados distintos, como son k bolitas sabemos que se pueden permutar de $k!$ formas.

Sigue que si denotamos x como el número que queremos obtener

$$\frac{n!}{(n-k)!} = x \cdot k! \quad \Longrightarrow \quad x = \binom{n}{k}$$

La buena interpretación para el coeficiente binomial es "de cuántas formas puedo sacar k objetos de entre n sin importar el orden".

Ejemplo 4.2. *Se lanzan 10 dados de 6 caras. ¿De cuántas formas se pueden obtener exactamente 4 Ases?.*

De cada dado se puede obtener un As de una forma y no obtener un As de 5 formas. Razonamos que se lanzan los 10 dados uno tras otro y se eligen 4 dados para que en estos salga un As y en los otros cualquier número menos un As.

Se pueden elegir 4 dados (o posiciones) de $\binom{10}{4}$ formas y cada uno de estos dados puede salir de una forma, en rigor se pueden sacar los 4 Ases de $\binom{10}{4} \cdot 1^4$ formas. Ahora, los otros 6 dados pueden sacar cualquier número menos un As, es decir, se pueden elegir 5^6 combinaciones.

Finalmente el resultado es $\binom{10}{4} \cdot 5^6$

Ejercicio: Usando los razonamientos anteriores, dé un argumento del porqué $\binom{n}{k} = \binom{n}{n-k}$

4.4. Elementos distinguibles por grupos

Ahora consideramos que tenemos n bolitas, y éstas pueden ser de M colores distintos, donde hay r_m del color m y se cumple $n = \sum_{m=1}^{M} r_m$. Extraemos las n bolitas de una a cada vez.

Si tenemos los colores blanco ⓑ y negro ⓝ, consideramos la extracción ⓑⓑⓝ distinta a la ⓝⓑⓑ. Es decir, solo distinguimos por colores, pero sabemos en que orden salieron.

Si todas las bolitas fueran distinguibles se pueden sacar de $n!$ formas distintas, pero no debemos contar las permutaciones dentro de cada color, sigue que el número de formas de realizar el experimento es:

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

éste se conoce como coeficiente multinomial.

Desafío: ¿Cómo se adapta la fórmula si solo se extraen k bolitas de entre las m ?

Ejemplo 4.3. *En un estante hay 4 libros rojos, 3 azules y 2 negros. ¿De cuántas formas se pueden ordenar?.*

Para resolver lo anterior, vamos a suponer que dentro de los libros rojos no podemos distinguirlos unos de otros (lo mismo con los otros colores). Luego podemos usar la fórmula anterior y tenemos que podemos ordenar de

$$\frac{(4+3+2)!}{4! \cdot 3! \cdot 2!} = 1260$$

5. Medida de probabilidad

Estamos realizando un cierto experimento, expresamos el conjunto Ω y ahora queremos calcular para $A \subseteq \Omega$ la probabilidad $\mathbb{P}(A)$ de ocurrencia.

Nos gusta interpretar que Ω tiene una masa y $\mathbb{P}(A)$ es la masa del conjunto A, mientras más grande, más frecuente será observar A. Para que esta interpretación sea correcta debemos tener, por ejemplo, que si dividimos Ω en dos, Ω_1 y Ω_2 con $\Omega_1 \cap \Omega_2 = \varnothing$ y $\Omega_1 \cup \Omega_2 = \Omega$ entonces deberíamos tener $\mathbb{P}(\Omega_1) + \mathbb{P}(\Omega_2) = \mathbb{P}(\Omega)$.

Podemos seguir con lo anterior, diciendo todas las buenas propiedades que nos gustaría que cumpla $\mathbb{P}(\cdot)$ y luego salir a buscar una medida que cumpla estas condiciones. Esto se llama el enfoque axiomático y fue desarrollado por el matemático ruso Kolmogorov en el siglo XIX.

Definición 5.1 (Axiomática de Kolmogorov)**.** *Sea Ω el espacio muestral, sea $\mathbb{P} : \mathcal{P}(\Omega) \to \mathbb{R}$ una función. Si cumple los siguientes axiomas diremos que $\mathbb{P}$ es una medida de probabilidad sobre Ω:*

1. $\forall\, A \subseteq \Omega, \quad 0 \leq \mathbb{P}(A) \leq 1$
2. $\mathbb{P}(\Omega) = 1$
3. $\forall\, (A_n)_{n\in\mathbb{N}} \subseteq \Omega$ *disjuntos de a pares* $\mathbb{P}\left(\dot\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$

El último axioma se conoce como "aditividad numerable" o bien "σ-aditividad", implica las propiedades fundamentales.

Proposición 5.1. *Si $\mathbb{P}$ es una medida de probabilidad sobre Ω, entonces:*

1. $\forall\, A \subseteq \Omega$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. $\forall\, A \subseteq B$ *eventos* $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$
3. *Si $(A_n)_{n\in\mathbb{N}}$ es partición del evento A, entonces* $\mathbb{P}(A) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$

Demostración

(1): Notamos que $A \,\dot\cup\, A^c = \Omega$ y por axioma 2 tenemos $\mathbb{P}(A \,\dot\cup\, A^c) = 1$.

Finalmente, por σ-aditividad $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$

(2): Definamos $C := B \setminus A$, claramente $C \cap A = \varnothing$. Además como $A \subseteq B$ tenemos que $B = C \,\dot\cup\, A$, por lo que deducimos $\mathbb{P}(B) = \mathbb{P}(C \,\dot\cup\, A)$ y aplicamos σ-aditividad a la última expresión, lo que nos da:

$$\mathbb{P}(C \,\dot\cup\, A) = \mathbb{P}(C) + \mathbb{P}(A) = \mathbb{P}(B \setminus A) + \mathbb{P}(A)$$

Igualando términos se tiene la propiedad.

(3): Directo de definición de partición y σ-aditividad

□

Ejemplo 5.1. *Se lanzan dos monedas que salen cara con probabilidad p, por cada una que salga en cara se gana un dólar. ¿cuál es la probabilidad de ganar exactamente un dólar?*

Definamos los eventos $D =$ "se gana 1 dólar", $C_1 =$ "la moneda 1 sale cara y la otra cruz" y $C_2 =$ "la moneda 2 sale cara y la otra cruz".

Observamos que C_1 y C_2 son disjuntos, más aún $C_1 \,\dot\cup\, C_2 = D$, entonces por propiedad de partición

$$\mathbb{P}(D) = \mathbb{P}(C_1) + \mathbb{P}(C_2) = p(1-p) + (1-p)p$$

6. Conjuntos equiprobables

Supongamos que realizamos un experimento donde la cantidad de resultados posibles es finita, i.e. $|\Omega| < \infty$. Es habitual modelar un fenómeno diciendo que será "totalmente al azar".

Digamos que se elegirá una persona de un grupo de 10 "totalmente al azar", este experimento tiene 10 resultados posibles e intuitivamente relacionamos el concepto "totalmente al azar" con elegir a cada uno con igual probabilidad, i.e. $\frac{1}{10}$.

Definición 6.1. *Sea Ω espacio muestral finito, decimos que Ω es equiprobable si:*

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} \quad \forall\, A \subseteq \Omega$$

La fórmula de la definición se lee como "casos favorables sobre casos totales".

Proposición 6.1. *La medida equiprobable (de la definición 6.1) es en efecto medida de probabilidad, i.e. satisface los axiomas de Kolmogorov.*

DEMOSTRACIÓN

Probaremos sólo σ-aditividad, los otros axiomas quedan propuestos.

Sean $(A_n)_{n\in\mathbb{N}}$ disjuntos y llamemos $A := \dot\cup_{n\in\mathbb{N}} A_n$. Como Ω es finito y $A \subseteq \Omega$ sigue que A es finito.

Como A es finito, se concluye que hay infinitos $A_n = \varnothing$, sino fuera así tendríamos que una unión infinita de conjuntos disjuntos da un conjunto finito, lo que es una contradicción.

Sin pérdida de generalidad pensamos entonces en los conjuntos $(A_n)_{n=0}^{N}$ pues los demás son vacíos, entonces:

$$\sum_{n=0}^{N} \mathbb{P}(A_n) = \sum_{n=0}^{N} \frac{|A_n|}{|\Omega|} = \frac{1}{|\Omega|} \sum_{n=0}^{N} |A_n| = \frac{1}{|\Omega|} \cdot |A|$$

□

Ejemplo 6.1. *Se tienen 4 amigos y se ordenan al azar en una hilera, ¿cuál es la probabilidad de que Pedro y Juan queden uno al lado del otro?*

Para resolver lo anterior identificamos Ω como el conjunto de todas las permutaciones de los 4 amigos, sabemos que $|\Omega| = 4!$.

Sea A el evento "Pedro y Juan quedan uno al lado del otro", para calcular $|A|$ razonamos que los 4 amigos quedan en las posiciones ①②③④.

Calculemos los casos donde Pedro queda a la izquierda de Juan, a esto lo multiplicamos por 2 y obtenemos el número que buscamos.

Pedro queda a la izquierda de Juan cuando ocupan las posiciones ①②, ②③ ó ③④. En cada una de éstas tenemos fijadas las posiciones de Pedro y Juan, pero los otros dos amigos se pueden ordenar libremente, es decir de $2!$ formas, sigue que como son 3 casos que Pedro queda a la izquierda de Juan tenemos $3 \cdot 2!$ formas que esto pase.

Finalmente $|A| = 2 \cdot 3 \cdot 2!$ y por lo tanto $\mathbb{P}(A) = \frac{2\cdot 3\cdot 2!}{4!} = \frac{1}{2}$

7. Probabilidad condicional

Definición 7.1. *Sea Ω el espacio de eventos con medida de probabilidad $\mathbb{P}$, sean $A, B \subseteq \Omega$ con $\mathbb{P}(B) > 0$. Definimos la probabilidad de A dado B como*

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

Proposición 7.1. *Sea $B \subseteq \Omega$ con $\mathbb{P}(B) > 0$, si definimos $\mathbb{P}_1(\cdot) = \mathbb{P}(\cdot|B)$, entonces la nueva función $\mathbb{P}_1(\cdot)$ satisface todos los axiomas de la definición 5.1, i.e. es una medida de probabilidad.*

Las consecuencias de lo anterior nos dicen que en particular la probabilidad condicional satisface todas las propiedades de la proposición 5.1.

7.1. Bayes y probabilidades totales

Proposición 7.2 (Fórmula de Bayes)**.** *Sean $A, B \subseteq \Omega$ tales que $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0$, entonces tenemos que*

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

Teorema 7.1 (Probabilidades totales)**.** *Sea $A \subseteq \Omega$ y sea (A_n) partición de A tal que $\mathbb{P}(A_n) > 0 \, \forall \, n$, entonces*

$$\mathbb{P}(A) = \sum_{n \in \mathbb{N}} \mathbb{P}(A|A_n)\mathbb{P}(A_n)$$

Demostración

Observamos que $\mathbb{P}(A|A_n)\mathbb{P}(A_n) \triangleq \frac{\mathbb{P}(A \cap A_n)}{\mathbb{P}(A_n)}\mathbb{P}(A_n) = \mathbb{P}(A \cap A_n)$.

Definamos $B_n := A \cap A_n$, tenemos que $(B_n)_{n \in \mathbb{N}}$ es partición de A, en efecto $B_n \cap B_m = \varnothing \, \forall \, n \neq m$ por cuanto $A_n \cap A_m = \varnothing$ y además

$$\bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} A \cap A_n = A \cap \bigcup_{n \in \mathbb{N}} A_n = A \cap A$$

Entonces por la propiedad de partición

$$\mathbb{P}(A) = \sum_{n \in \mathbb{N}} \mathbb{P}(B_n) = \sum_{n \in \mathbb{N}} \mathbb{P}(A \cap A_n)$$

que es lo que queríamos demostrar

$\square$