Estadística

Principios de
Estadística aplicada

Jorge Ortiz Pinilla

ediciones de la U

Estadística

Principios de
Estadística aplicada

Jorge **Ortiz** Pinilla

Profesor de la Facultad de Estadística
Universidad Santo Tomás, Bogotá

ediciones de la U

Ortiz Pinilla, Jorge

Principios de Estadística aplicada -- Bogotá : Ediciones de la U, 2013.
p.200 ; 24 cm.

1. Relaciones y conteo 2. Objetos de estudio 3. Mediciones colectivas
4. Asociación 5. Probabilidades 6. Inferencia estadística I. Tít.
519.5 cd

Área: Estadística
Primera edición: Bogotá, Colombia, enero de 2013

© Ediciones de la U - Transversal 42 No. 4 B-83 - Tel. (+57-1) 4065861
www.edicionesdelau.com - E-mail: editor@edicionesdelau.com
Bogotá, Colombia

Coordinación editorial: Adriana Gutiérrez M.
Carátula: Ediciones de la U

Prefacio

Estas son notas para un curso de 16 semanas de probabilidad y estadística. El objetivo es promover el buen uso de herramientas básicas de análisis de datos y de probabilidades en problemas de diversa índole. Durante el desarrollo del curso, el estudiante debe ir ubicando permanentemente sus conocimientos dentro de un ambiente estadístico. Para ello deberá tener como idea conductora el acceso a fuentes generadoras de información de donde extrae datos con el fin de acercarse al conocimiento de propiedades generales en las que tiene interés. El fenómeno estudiado se manifiesta a través de los datos que revelan el comportamiento de las variables observadas.

Con las técnicas de conteo se ilustran las formas clásicas más utilizadas para seleccionar elementos de un conjunto cualquiera y tomar los datos adecuados para un estudio. En este capítulo, el estudiante debe recibir y asimilar la enseñanza suficiente para que pueda establecer la analogía entre una ficha que se extrae de una urna y un elemento que se selecciona de un conjunto en una aplicación real. Por ejemplo, así como de la urna se extrae un ficha para observar su color o su forma, de una población en estudio se selecciona una persona y se examina si sufre una determinada enfermedad, o de un lote de producción se escoge una aparato y se observa si presenta un determinado defecto de fabricación.

La población y las variables que se miden en sus elementos son la fuente generadora de información que pone a disposición del observador un conjunto de datos. El principal interés está en el conocimiento de formas de presentación y de resumen de la información poblacional que faciliten tomar decisiones. Por lo general, estos resúmenes se orientan a informar sobre propiedades de los valores que toman las variables, en particular dónde se encuentran (localización), de qué magnitud son las diferencias que se presentan entre ellos (dispersión), cómo se encuentran distribuidos según agrupaciones de valores (forma distribucional) y cómo se relacionan los de unas variables con los de otras (asociación) o con los de algunos patrones específicos de observación a lo largo del tiempo o del espacio (tendencias).

La aplicación de las técnicas de conteo a una fuente de información parte de la concepción de una población o de un proceso como una urna de la que se extraen elementos para ser observados y evaluados. Dependiendo de la composición en la población, de las variables medidas y de la técnica de extracción, el resultado es un nuevo conjunto de valores cuya distribución puede llegar a ser determinada al menos dentro de niveles de aproximación suficientemente satisfactorios.

Cuando la extracción de los elementos y la disposición para su uso se desarrollan dentro de escenarios probabilísticos controlados o conocidos, los conceptos de la teoría de probabilidades entran en juego para establecer teóricamente el comportamiento de variables aleatorias útiles para conocer aspectos importantes de las poblaciones o de los fenómenos estudiados con miras a explorar relaciones entre variables, a pronosticar resultados o a examinar su coherencia con respecto a hipótesis que los investigadores pudieran haber planteado.

La inferencia estadística es una formalización del uso de variables aleatorias resultantes de ejercicios de extracción o de experimentación para buscar información sobre aspectos específicos del comportamiento de las variables estudiadas (estimación) o para tomar decisiones sobre la aceptabilidad de ciertas hipótesis de interés para el investigador (pruebas estadísticas).

Aunque la mayoría de ejercicios numéricos puede desarrollarse en hojas electrónicas como Excel o Libre Office, o con calculadoras populares como Casio, es recomendable que el estudiante aprenda a utilizar un paquete especializado para análisis estadístico. `R` es actualmente uno de los más utilizados por la comunidad estadística y SPSS es uno de los más difundidos a nivel comercial. En el capítulo final, dedicado a la inferencia estadística, se incluyen instrucciones cortas para aplicar los procedimientos con el programa `R`.

Quiero agradecer a la Universidad Santo Tomás por su apoyo para la preparación de este texto. A mis colegas de la Facultad de Estadística, en especial a los profesores Andrés Gutiérrez y Yesid Rodríguez por su interés y por sus comentarios oportunos y valiosos. Al profesor William Rincón por utilizar el material en sus cursos y a los estudiantes de las carreras de Ingeniería de la Universidad Santo Tomás. A mi esposa, Joanna, y a mis hijos, Santiago, Stefan y Kasia, que comprendieron los momentos de abandono para dedicarme a escribir.

J.O., Bogotá, noviembre de 2012.

Índice general

1. Introducción **1**

2. Relaciones y conteo **5**
2.1. Relaciones 5
2.2. Producto cartesiano y relaciones 6
2.3. Formas de extracción de objetos de una urna 8
2.4. Variables, datos y elementos indistinguibles 10
2.5. Reglas de conteo 10
2.6. Ejercicios 16

3. Objetos de estudio **23**
3.1. La población 23
3.2. Características y variables 23
3.2.1. Variables equivalentes 27
3.2.2. Clasificación de las variables y características 27
3.2.3. Tipos de variables 28
3.3. Transformaciones de variables 33
3.4. Ejercicios 34
3.5. Taller con R 36

4. Mediciones colectivas **39**
4.1. Colectividades 39
4.2. Estudio de las colectividades 41
4.3. Distribución de frecuencias 41
4.3.1. Tipos de frecuencias 43
4.3.2. Diagramas de frecuencias 45
4.3.3. Diagramas de barras 46
4.3.4. Diagramas circulares 47
4.3.5. Diagramas de Pareto 48
4.3.6. Diagramas de ramas y hojas 48
4.3.7. Histogramas 50
4.4. Resúmenes numéricos 52
4.5. Variables de agrupación 53

4.5.1. Localización 56
4.5.2. Otros promedios 60
4.5.3. Dispersión 62
4.5.4. Ejercicios de carácter teórico 64

5. Asociación **69**
5.1. Análisis de variables vectoriales 69
5.2. Covarianza y correlación lineal 70
5.2.1. Combinaciones lineales de variables 73
5.3. Aplicación a la propagación de errores 76
5.3.1. Incertidumbre de combinaciones lineales de mediciones . . . 76
5.3.2. Incertidumbre de funciones de variables 76
5.4. Regresión lineal simple 77
5.4.1. Relaciones linealizables 85
5.5. Ejercicios 89

6. La población como una urna **95**
6.1. Datos dicotómicos 96
6.1.1. Extracciones con reposición de datos dicotómicos 99
6.1.2. Extracciones sin reposición de datos dicotómicos 102
6.1.3. Extracciones con reposición hasta obtener k éxitos 105
6.2. Datos numéricos 109
6.2.1. Extracciones con reposición de datos numéricos 110
6.2.2. Teorema central del límite y distribución normal 115
6.2.3. Otras distribuciones muestrales 120
6.2.4. Extracciones sin reposición de datos numéricos 123
6.2.5. Observaciones 125
6.3. Ejercicios 125

7. Probabilidades **127**
7.1. Experimentos aleatorios 127
7.2. σ-álgebras y eventos 129
7.3. Probabilidades 130
7.4. Ejercicios 136

8. Inferencia estadística **139**
8.1. Introducción 139
8.2. Pruebas de hipótesis 140
8.2.1. Planteamiento de las hipótesis 140
8.2.2. El proyecto inferencial 142
8.3. Ejercicios 150
8.4. Estimación de parámetros 151

9. Inferencia con distribuciones conocidas **157**
9.1. Distribución binomial 157
9.1.1. Pruebas sobre p 158
9.2. Distribución hipergeométrica 162
9.3. Variables en escalas de intervalo y de razón 165
9.3.1. Pruebas T para una media 166
9.3.2. Dos muestras pareadas 169
9.3.3. Dos muestras independientes 173
9.3.4. Pruebas χ^2 para la varianza de una población 176
9.3.5. Pruebas F para varianzas con dos muestras independientes 181

Capítulo 1

Introducción

Frente al entorno que lo rodea, el ser humano desarrolla intensamente su actividad intelectual de tres formas esenciales, orientadas a:

1. Describir lo que observa

2. Explicar lo que ocurre

3. Pronosticar o especular sobre sucesos venideros o del pasado.

La observación de los fenómenos con miras a comprender el universo ha llevado a concebirlo como un sistema integrado cuyos elementos interactúan de manera múltiple y compleja y en el que todo acontecimiento es resultado de esas interacciones: "No hay efecto sin causa". Los *modelos* son herramientas construidas por el hombre para resumir sus observaciones o para esquematizar la forma como relaciona unos fenómenos con otros con el propósito de buscar explicaciones, elaborar pronósticos o incluso intervenirlos para modificar su comportamiento.

Las formas de actividad intelectual presentadas tienen lugar en dos escenarios básicos: el primero, donde un evento ya ha sido observado y se buscan las causas que lo generaron o los factores que se le pueden asociar y el segundo, donde se conocen unos factores asociables en algún grado con uno o varios eventos aún no observados.

La capacidad de actuar deliberadamente sobre la naturaleza ofrece diversas opciones para buscar explicaciones de los eventos resultantes, desde limitarse a establecer las condiciones de observación de los factores asociables con dichos eventos, a la manera de un espectador atento pero pasivo frente al desarrollo del fenómeno estudiado (estudios observacionales), hasta asumir el control manipulando rigurosamente las condiciones de las variables que se consideran explicativas del comportamiento de las que se toman como resultantes (estudios experimentales).

Generalmente el hombre está limitado a efectuar unas cuantas observaciones de lo que ocurre y le es inalcanzable o inconveniente tomar en consideración, una a una, todas las formas posibles del comportamiento de un fenómeno. Es preciso

entonces diseñar estrategias optimizadas de recolección y de análisis de los datos que ayuden a describir adecuadamente en su integralidad lo que se quiere conocer.

En cada disciplina se establecen pautas que orientan las metodologías de investigación según los intereses y las limitaciones que le son propios. Enseguida mencionamos sólo algunos de los aspectos que es necesario considerar para organizar un análisis estadístico:

1. El tema. Se trata de la delimitación conceptual de lo que se quiere estudiar. Puede ser algo tan simple como las dimensiones físicas de un objeto construido, o tan complejo como la satisfacción de los empleados de una empresa. La siguiente es una lista muy corta de algunos ejemplos en diferentes áreas de aplicación:

   a) Composiciones poblacionales por grupos de edad, sexo, región, morbilidad, mortalidad, perfiles poblacionales por ocupación, tenencia de vivienda, educación.

   b) Accidentalidad de cualquier tipo como laboral o de tránsito durante un período de tiempo y sus consecuencias, como días de incapacidad, costos hospitalarios, indemnizaciones.

   c) Pruebas diagnósticas de afecciones psicológicas y su evaluación con instrumentos de medida, habilidades, aptitudes, actitudes.

   d) Asistencia a espectáculos, eventos o fenómenos masivos, consultas de sitios o páginas web según tema consultado.

   e) Imagen de personalidades o de instituciones, favoritismo de candidatos, pronósticos de votaciones.

   f) Consumo y calidad de servicios (educación, salud, energía, agua potable), o de bienes (inmuebles, muebles, electrodomésticos).

   g) Estudios de calidad, errores en la recolección, registro o digitación de datos, fallas en un sistema, causas de las fallas, tiempo de vida de equipos, calidad de productos, deterioro de materiales, efectividad de procesos para mejorar la calidad o reducir costos de fabricación, horas/obrero para realizar una tarea.

   h) Medición de objetos y sustancias, peso, densidad, resistencia, opacidad, pH, diámetro, largo, ancho, comparación de medidas frente a estándares: peso real *versus* peso nominal, ingredientes de productos, aportes nutricionales, dosis de medicamentos.

   i) Fenómenos naturales, precipitación, vientos, temperatura, caudales, inundaciones, sequías, riquezas minerales, agrícolas, biodiversidad, contaminación vehicular e industrial, producción de basuras y residuos.

   j) Economía, industria, exportaciones, importaciones, divisas, costos de producción, endeudamiento, vivienda, costo de vida, inflación.

   Como veremos más adelante, será necesario esquematizar y formalizar cómo se realizarán las observaciones mediante el uso de variables que servirán para precisar la forma como el investigador aborda su tema.

2. La población. Es el conjunto de elementos donde el tema se manifiesta como de interés para el investigador. Así como en el punto anterior se habló del tema para concretar *lo que se quiere estudiar*, ahora se trata de delimitar *en dónde se quiere estudiar*. Como ejemplos:

   a) Los habitantes de una región, los trabajadores de una comunidad o de una empresa, los aspirantes a un cargo laboral, los asistentes a un espectáculo, los electores de un cargo gubernamental, los usuarios de un servicio, los compradores de un producto o los clientes de una empresa comercial o de servicios.

   b) Los registros de una base de datos, lotes de producción, conjuntos de sitios de un ecosistema, regiones geográficas, conjuntos de empresas, instituciones o negocios o conjuntos de tiempos o períodos de funcionamiento de un sistema.

3. El contexto. Son las circunstancias, temporales, espaciales y otras bajo las que se lleva a cabo el estudio para concretar su significado.

   a) Períodos de crisis, de negociación o de creación de empresas, de cambios organizacionales o políticos, de conflictos internos o externos.

   b) Limitaciones espaciales o temporales por fenómenos naturales.

4. El método de observación. Una primera opción es el *método observacional*, donde el investigador es un espectador dedicado a recoger datos bajo condiciones determinadas, sin intervenir en el desarrollo de los fenómenos que estudia. Por ejemplo, para examinar el impacto de una fuente contaminante del medio ambiente, recoge sus datos en el entorno de influencia sin introducir elementos adicionales que modifiquen las condiciones del lugar.

   El *método experimental* es otra opción en la cual el investigador es un participante activo que crea o modifica las condiciones de las que luego observa el impacto sobre los fenómenos estudiados. Por ejemplo, condiciona artificialmente diferentes formas de riego para examinar sus efectos sobre los rendimientos de algunos tipos de abono aplicados a un cultivo.

5. La estrategia de muestreo. Las fuentes de datos que se toman para estudiar los fenómenos son, por lo general, inasequibles en su integralidad y el investigador sólo puede realizar observaciones parciales. Las muestras se definen como subconjuntos de una población determinada y su papel es análogo al de una o varias fotografías que se toman para mostrar una realidad. Sin ser la realidad misma, luego de un proceso que se desarrolla en varios pasos, deben ofrecer una imagen que permita describirla de manera adecuada en aquellas características para las que fueron tomadas. Esa imagen la ofrecen mediante el uso de estimadores de las propiedades que se busca conocer (Gutiérrez, 2010)

6. La estrategia de análisis. Es una propuesta justificada de cada uno de los métodos estadísticos que se intentan aplicar para apoyar el logro de los objetivos que el investigador ha fijado para su estudio.

Capítulo 2

Relaciones y conteo

Como se sabe, la naturaleza puede concebirse como un sistema dinámico de elementos de diferente índole y relacionados en formas tan complejas que el hombre sólo logra observar proporciones muy pequeñas de lo que ocurre. De manera muy esquemática, puede verse como una urna de donde se extraen objetos, elementos o fenómenos, de los que se observan unas pocas características para estudiarlos. En los próximos capítulos revisaremos las bases para formalizar los conceptos de *relación* y de *función* y las técnicas más elementales para contar las formas diferentes de obtener resultados de este ejercicio de extracción y de observación.

2.1. Relaciones

Relacionar dos conjuntos A y B es establecer vínculos entre los elementos del primer conjunto y los del segundo.

Ejemplo: En la figura 2.1, el área encerrada es un lago. A partir del origen O es posible llegar a cualquiera de los destinos D1, D2, D3 o D4 por algún medio de transporte: tierrra (T), aire (A), navegación (N) o algunas combinaciones de ellos, como se indica en la siguiente tabla:

	Formas de transporte desde O					
Destino	T	A	N	TN=T+N	AN=A+N	AT=A+T
D1	✓	✓				
D2				✓	✓	
D3	✓					✓
D4	✓					

La tabla presenta una relación entre los destinos a los que se puede llegar desde O y los medios de transporte. La información que contiene permite saber cuántas y cuáles formas de transporte se tienen para llegar a cada destino y cuántos y cuáles destinos son asequibles desde O por cada medio de transporte. Por ejemplo, hay

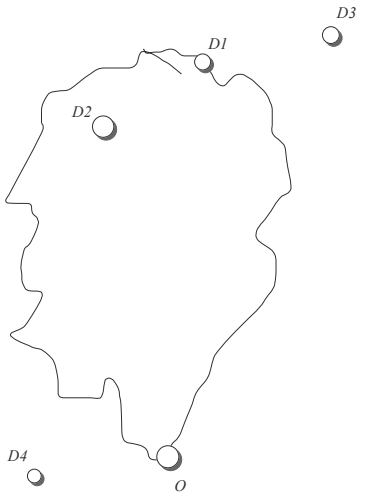


Figura 2.1: Formas de transporte desde O hasta D1, D2, D3 y D4. La regi´ on delimitada por la curva cerrada es un lago y D2 es un punto interior.

una sola forma de llegar a D4 y es por tierra. Para llegar a D1 hay dos formas: una por tierra y otra por aire, a D3 se puede llegar por tierra o parcialmente por tierra y parcialmente por aire, etc. Adem´ as, solo por tierra se puede llegar a tres destinos (D1, D2 y D4), mientras que por combinaci´ on aire-tierra se llega s´olo a D3, etc.

Formalmente las relaciones se representan con conjuntos de parejas formadas con los elementos relacionados. El orden en cada pareja es importante: el primer elemento pertenece al primer conjunto (D) y el segundo al segundo conjunto (T). Por ejemplo, la relaci´on anterior se escribe:

$$R_{D,T} = \{(D1,T),(D3,T),(D4,T),(D1,A),(D2,TN),(D2,AN),(D3,AT)\} \quad (2.1)$$

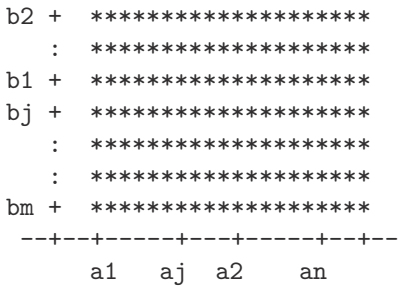
donde una pareja muestra una asociaci´on entre un destino y una forma de transporte que lo hace asequible desde el origen O.

Ejemplo 2.1.1. Mediante l´ineas continuas (—) para el transporte a´ ereo, onduladas ( ) para el transporte por navegaci´on y trazos discontinuos (––) para el transporte terrestre, unir los puntos O, D1, D2, D3 y D4 para que se tenga la relaci´on de acceso descrita.

2.2. Producto cartesiano y relaciones

Para introducir el tema del conteo, utilizaremos s´olo conjuntos con cantidades finitas de elementos.

Dados dos conjuntos finitos $A = \{a_1, a_2, \ldots, a_n\}$ y $B = \{b_1, b_2, \ldots, b_m\}$, el producto cartesiano $A \times B = \{(a_i, b_j), a_i \in A, b_j \in B\}$ contiene todas las $n \times m$ formas posibles de asociar cada elemento de A con cada uno de los de B.

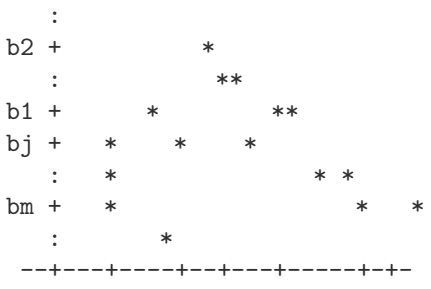


La pareja (a_j, b_j) con $a_j \in A,\ b_j \in B$ se utiliza para indicar que los elementos $a_j \in A$ y $b_j \in B$ se encuentran asociados.

Esta definición se puede extender a más de dos conjuntos, digamos $A \times B \times C \times D = \{(a_i, b_j, c_k, d_l) \mid a_i \in A, b_j \in B, c_k \in C, d_l \in D\}$ y el número de elementos de este producto es igual a $n_A \times n_B \times n_C \times n_D$, donde n_A representa el número de elementos de A, etc. El arreglo (a_i, b_j, c_k, d_l) se utiliza para indicar que los elementos $a_i \in A, b_j \in B, c_k \in C, d_l \in D$ se encuentran asociados.

El producto cartesiano contiene todas las asociaciones posibles entre elementos de los conjuntos y entonces puede considerarse como el universo de posibilidades. Sin embargo, en muchas ocasiones, lo que se quiere es conocer o establecer un conjunto más reducido de elementos relacionados. Por ejemplo, si A es un conjunto de pacientes de una misma enfermedad y B uno de tratamientos pertinentes, el producto cartesiano incluye todas las opciones de aplicación de cualquiera de los tratamientos a cada paciente. Un conjunto de mayor interés es el de los que pueden aplicarse a cada paciente teniendo en cuenta su estado de salud y sus condiciones particulares. Muy seguramente ya no todos los tratamientos serán convenientes para todos los pacientes y esto hace necesario considerar un subconjunto del producto cartesiano que corresponda a esta situación.

Los subconjuntos de $A \times B$ se conocen como *relaciones* de A en B, como la que se esquematiza en la siguiente figura:



En una relación, algunos elementos de A se asocian con algunos de los de B y a_i se asocia con una determinada cantidad $\mathfrak{n}(a_i)$ de elementos de B. Entonces el número total de relaciones de elementos de A con elementos de B es igual a:

$$\sum_{i=1}^{n} \mathfrak{n}(a_i) \tag{2.2}$$

Cuando cada elemento de A se relaciona con todos los de B, $\mathfrak{n}(a_i) = m$ y entonces, el número total de relaciones entre elementos de A y B es igual a $\sum_{i=1}^{n} \mathfrak{n}(a_i) = \sum_{i=1}^{n} m = n \times m$, que es el mismo número de elementos del producto cartesiano.

Más adelante se verá que estas dos formas de contar, mediante productos cuando cada elemento de A se relaciona con todos los de B, o mediante sumas cuando la cantidad de relaciones depende de cada elemento de A, dan origen a lo que se conoce como las reglas de la multiplicación y de la suma, respectivamente.

En muchas ocasiones se utiliza el conjunto A como base de conocimiento para acercarse al conjunto B. Por ejemplo, si A es un conjunto de síntomas y B uno de enfermedades, se tiene interés en conocer la forma como se asocian unos con otros para utilizar los síntomas como predictores de las enfermedades.

Potencias cartesianas de un conjunto

De manera especial, el producto cartesiano de A consigo mismo k veces, $A^k = A \times A \times \cdots \times A$, es el conjunto de k-uplas $(a_{i_1}, a_{i_2}, \ldots, a_{i_k})$, donde $a_{i_j} \in A$, y el número total de elementos (k-uplas) es igual a

$$\mathfrak{n}(A^k) = n^k \tag{2.3}$$

Cuando $k = 2$, si cada elemento de A se relaciona con los demás pero no consigo mismo, $\mathfrak{n}(a_i) = n - 1$ y el número de parejas de elementos diferentes que pueden formarse es:

$$\sum_{i=1}^{n} \mathfrak{n}(a_i) = n(n-1) \tag{2.4}$$

En el caso de k-uplas de elementos *diferentes* de un mismo conjunto A, las posibilidades son:

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!} \tag{2.5}$$

2.3. Formas de extracción de objetos de una urna

Para exponer las ideas relacionadas con las técnicas de conteo utilizaremos el concepto de *urna* Ω, como un repositorio donde se encuentran n elementos, en

esencia diferentes. De ellos se *extraen* k siguiendo ciertas reglas y el propósito del *conteo* es enumerar o contar los posibles resultados de estas operaciones.

Existen dos formas básicas de extracción de un elemento de una urna:

1. **Con reposición**. Cuando el elemento seleccionado se utiliza para tomar algunos datos y se lo regresa a la urna, se dice que se ha realizado una *extracción con reposición*. La extracción no genera cambios en el contenido de la urna y las condiciones para las siguientes operaciones permanecen iguales. Además, como cada elemento seleccionado regresa a la urna, es posible que en las siguientes extracciones vuelva a ser seleccionado. Por ello, este procedimiento se conoce también como *extracción con repetición*.

   Si $r = (r_1, r_2, \ldots, r_k)$ representa el resultado de k extracciones con reposición de elementos de Ω, entonces $r \in \Omega^k$ y se puede aplicar (2.3) para encontrar que hay n^k resultados posibles.

2. **Sin reposición.** Esta forma se presenta cuando el elemento seleccionado no regresa a la urna, generando condiciones diferentes para el siguiente paso, pues la composición de la urna cambia. Al efectuar k extracciones de elementos de Ω sin reposición, no es posible que un elemento se repita. El procedimiento se conoce también como *extracciones sin repetición*.

   Si $r = (r_1, r_2, \ldots, r_k)$ representa el resultado de k extracciones sin reposición, entonces $r \in \Omega^k$, pero con componentes diferentes, así que, aplicando (2.5), se encuentran $n(n-1)(n-2)\cdots(n-k+1) = \dfrac{n!}{(n-k)!}$ resultados posibles.

La extracción de k elementos a la vez equivale a k extracciones sin reposición, cada una de uno de los elementos que aún permanecen en la urna. En este caso, ningún elemento puede salir repetido, mientras que en las extracciones con reposición sí.

Arreglos o permutaciones. Es posible que el uso que se dé a los elementos extraídos dependa de algún ordenamiento que se establezca entre ellos. En este caso, un resultado como $e_1\, e_2\, e_3$ se considera diferente de otro como $e_2\, e_3\, e_1$ aunque los elementos escogidos sean los mismos. Un resultado de este proceso donde se obtienen k elementos y luego se establece un orden entre ellos se llama un *arreglo* o una *permutación* de k elementos de Ω. Dos arreglos difieren si contienen elementos diferentes o si se presentan en orden diferente.

Por ejemplo, si tenemos veinte libros diferentes y queremos obsequiar uno a Jaime, otro a María y otro a Laura, debemos escoger tres de los veinte libros y luego decidir cuál obsequiamos a cada uno, así que con un mismo conjunto de tres libros extraídos obtendremos obsequios diferentes dependiendo de cómo los asignamos a cada persona.

Combinaciones. Cuando entre los k elementos extraídos de Ω, el orden en que se presentan no tiene ninguna importancia y cualquier otro ordenamiento