Metodología Cuantitativa I

Dra. Llery Ponce

Profesora asistente
Instituto de Estudios Avanzados en Educación
Universidad de Chile

Dr. Patricio Rodríguez

Profesor asistente
Instituto de Estudios Avanzados en Educación
Universidad de Chile

Aspectos generales de la investigación cuantitativa

Magíster en Investigación en Educación
Diplomado de postítulo en Metodologías de Investigación en Educación

UNIVERSIDAD DE CHILE
INSTITUTO DE ESTUDIOS AVANZADOS EN EDUCACIÓN

Santiago, Chile
Marzo 27, 2024

Contenido

1

Estadística inferencial

1

Estadística inferencial

Conceptos básicos

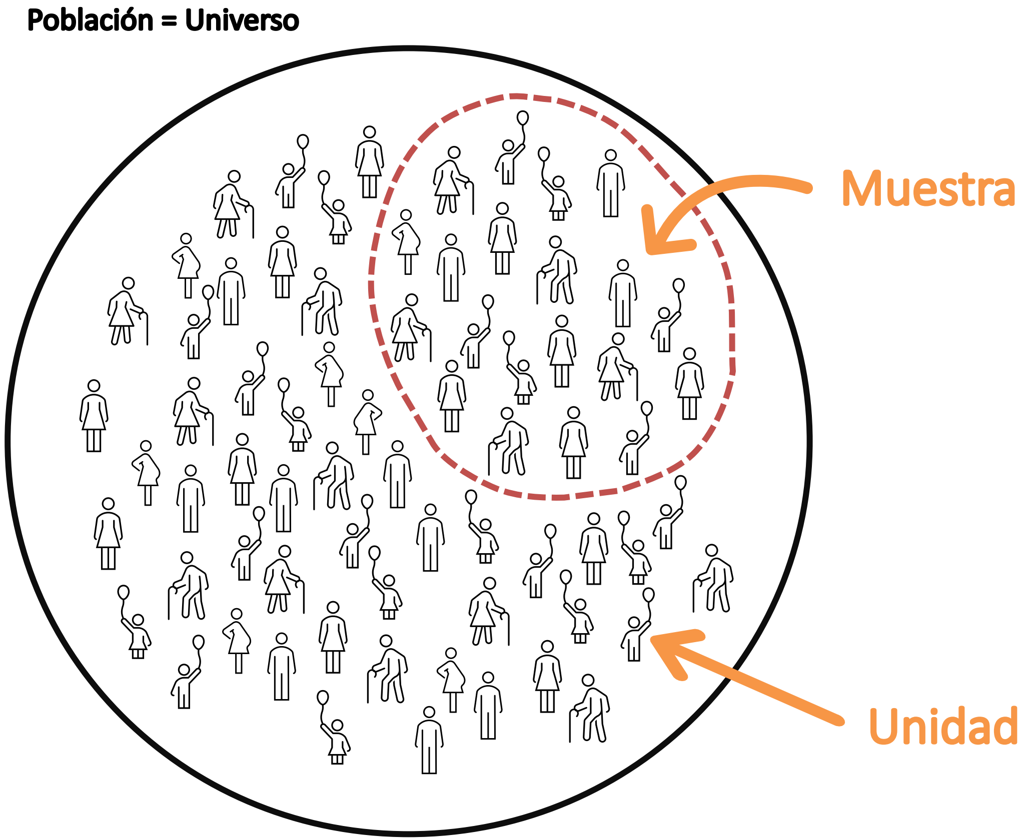
Se analiza la información de una **muestra** que se extrae de una **población** para **inferir** lo que ocurre en dicha población a partir de los resultados de la muestra:

- **Unidad de análisis:** cualquier elemento que porte información sobre el fenómeno que se estudia (personas, días, aulas, sucesos, etc.)
- **Población**: conjunto de todas las unidades de análisis (personas, objetos, aulas, etc.) que porten información sobre el fenómeno que se estudia.
- **Muestra**: subconjunto que se selecciona de la población.

1

Estadística inferencial

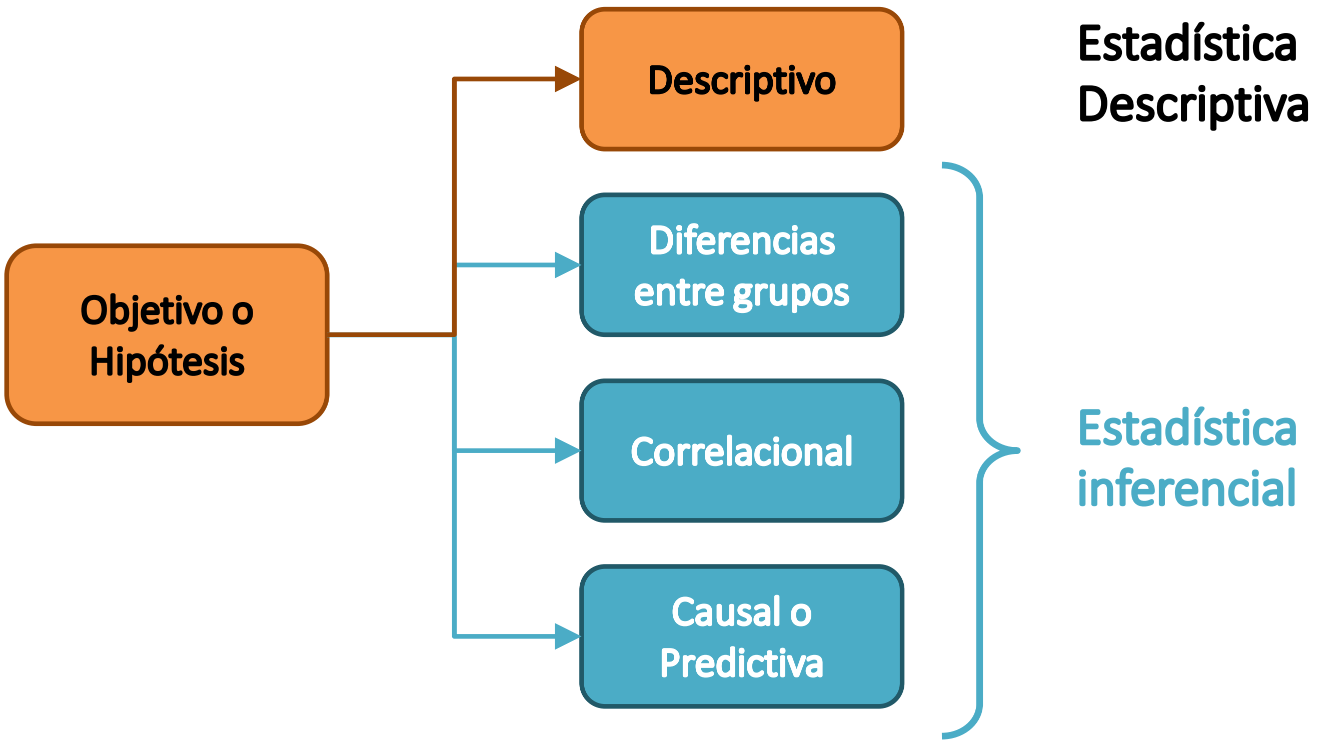
Unidad de análisis: Muestra: Población



1

Estadística inferencial

Unidad de análisis: Muestra: Población



2

Probabilidad

2 Probabilidad

- La **probabilidad** es la tasa o frecuencia con la que puede ocurrir el evento o suceso que estamos estudiando.
- Depende de la cantidad de casos posibles totales que existan y del evento particular en que estemos interesados.
- Se calcula de la siguiente forma:

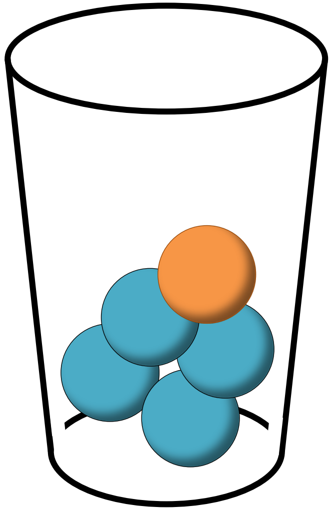
$$\mathbb{P}(caso\ favorable) = \frac{Nº\ casos\ favorables}{Nº\ de\ casos\ totales}$$

2

Probabilidad

Ejemplo

¿Cuál es la probabilidad de sacar una bola naranja?

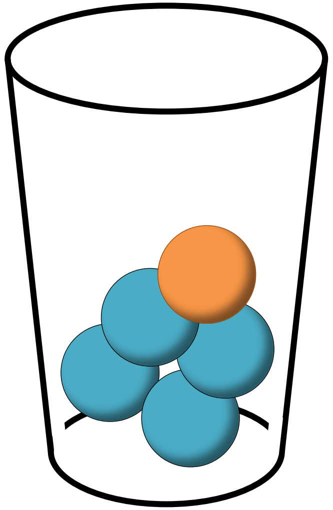


$$\mathbb{P}(bola\ \boldsymbol{naranja}) = \frac{Nº\ bolas\ \boldsymbol{naranjas}}{Nº\ de\ casos\ totales} = \frac{\mathbf{1}}{(\mathbf{1}+\mathbf{4})} = \frac{\mathbf{1}}{\mathbf{5}} = \mathbf{20\%}$$

Probabilidad

Ejemplo

¿Cuál es la probabilidad de sacar una bola celeste?



$$\mathbb{P}(bola\ \boldsymbol{celeste}) = \frac{N^{o}\ bolas\ \boldsymbol{celestes}}{N^{o}\ de\ casos\ totales} = \frac{\mathbf{4}}{(\mathbf{1}+\mathbf{4})} = \frac{\mathbf{4}}{\mathbf{5}} = \mathbf{80\%}$$

¿Qué es más probable?: ¿Sacar una bola celeste o una naranja?

¿Por qué?

Probabilidad

Ejemplo

- Aquí solamente tenemos **2** eventos posibles: sacar una bola **naranja** o una **celeste**
- Hay más posibilidades de sacar bolas celestes porque hay más de ellas.
- Dado la existencia de dos eventos disjuntos, que no pueden ocurrir simultáneamente, las **probabilidades individuales** de cada evento suman **1**.

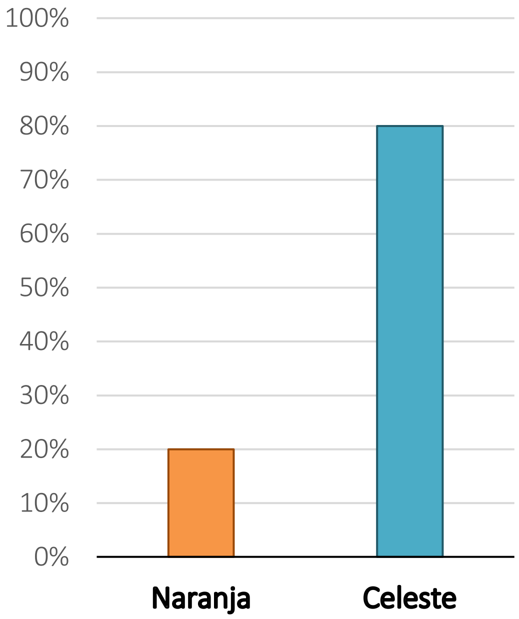
$$\mathbb{P}(naranja) + \mathbb{P}(celeste) = \mathbf{1}$$

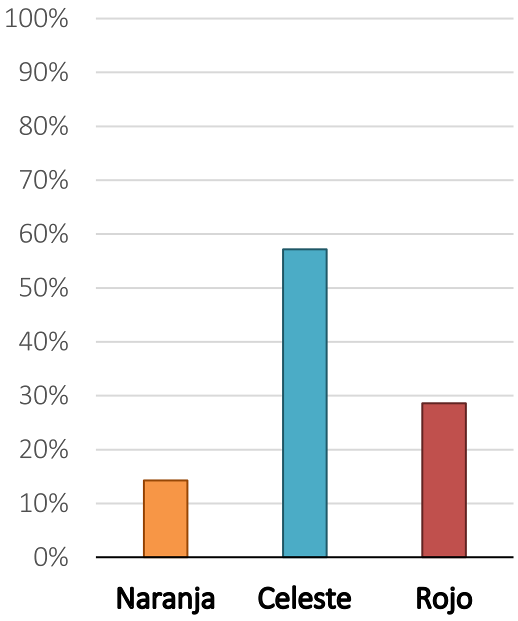
$$\mathbb{P}(naranja) = \mathbf{1} - \mathbb{P}(bola\ celeste)$$

$$\mathbb{P}(celeste) = \mathbf{1} - \mathbb{P}(naranja)$$

Probabilidad

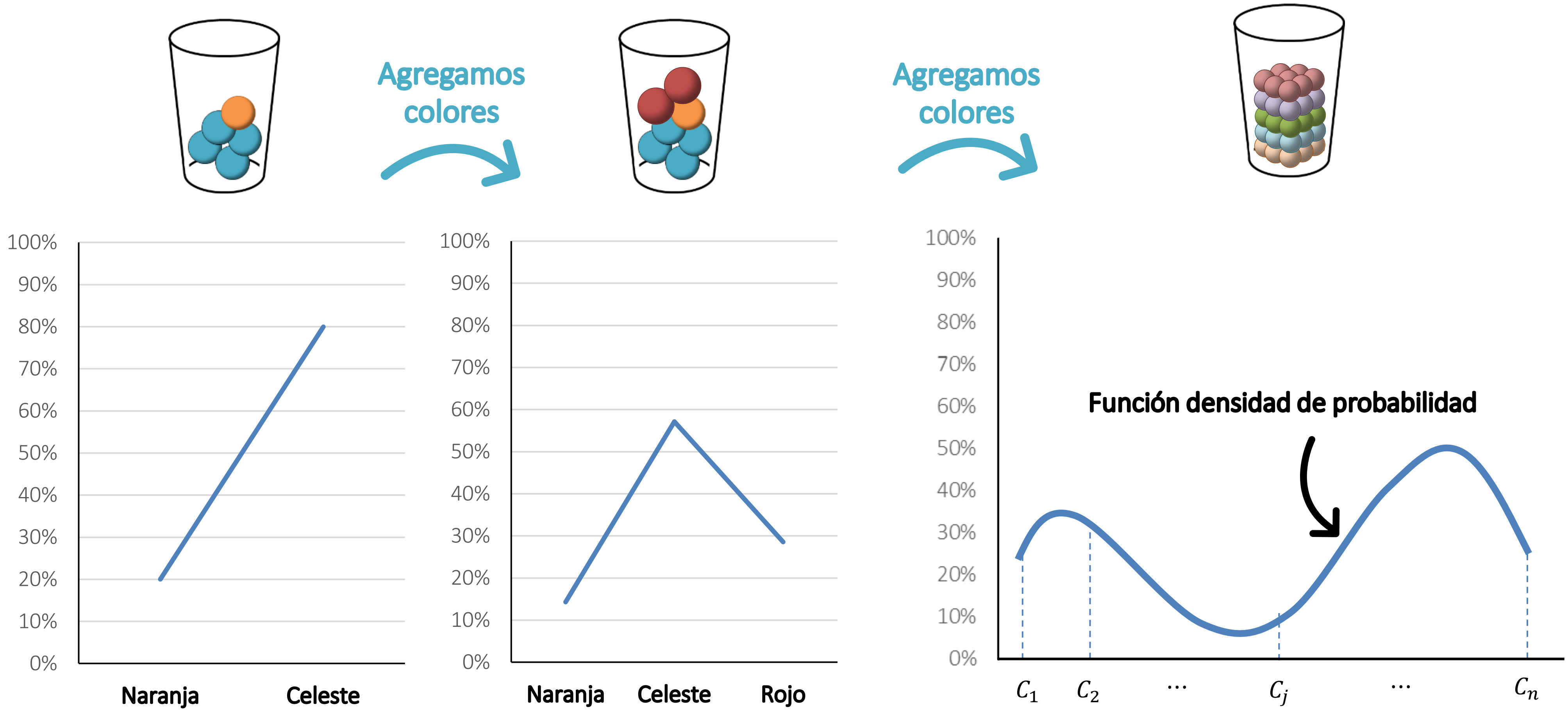
Frecuencias





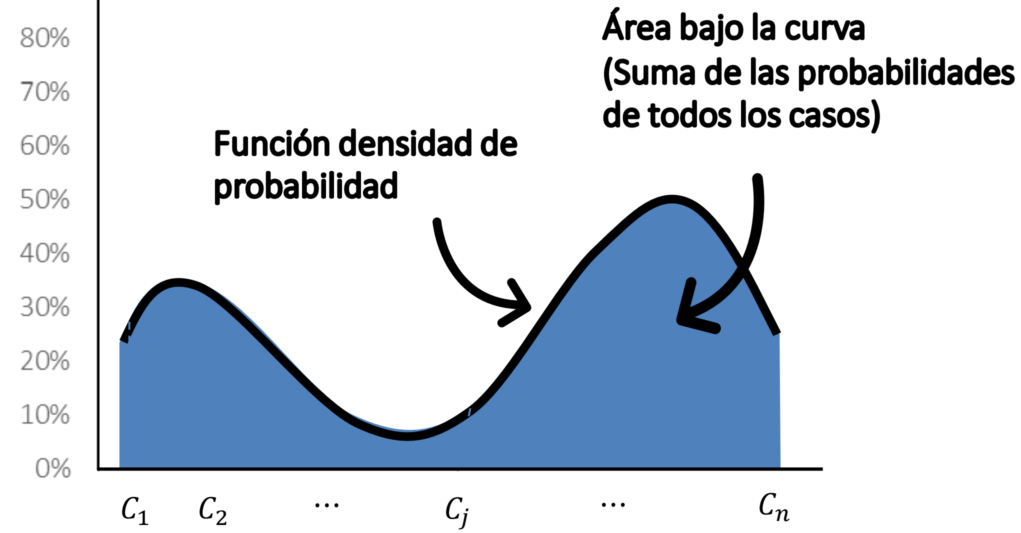
2 Probabilidad

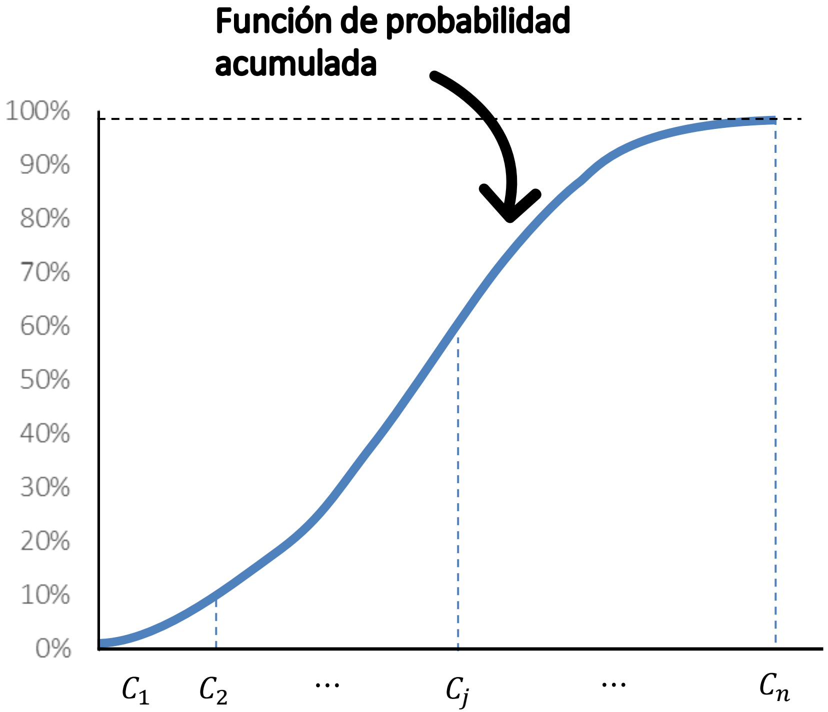
Frecuencias



Probabilidad

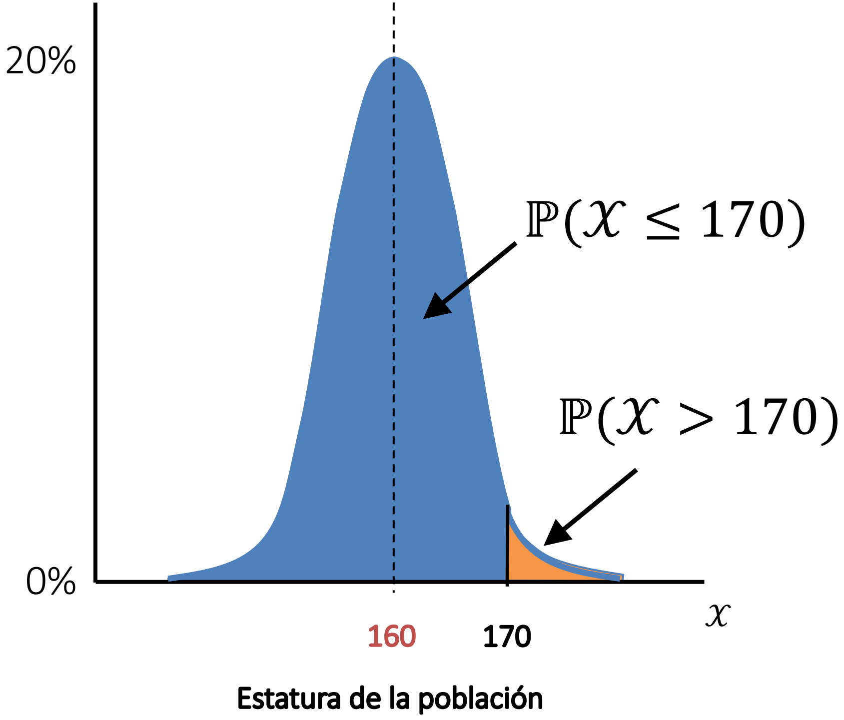
Funciones de densidad y acumulada

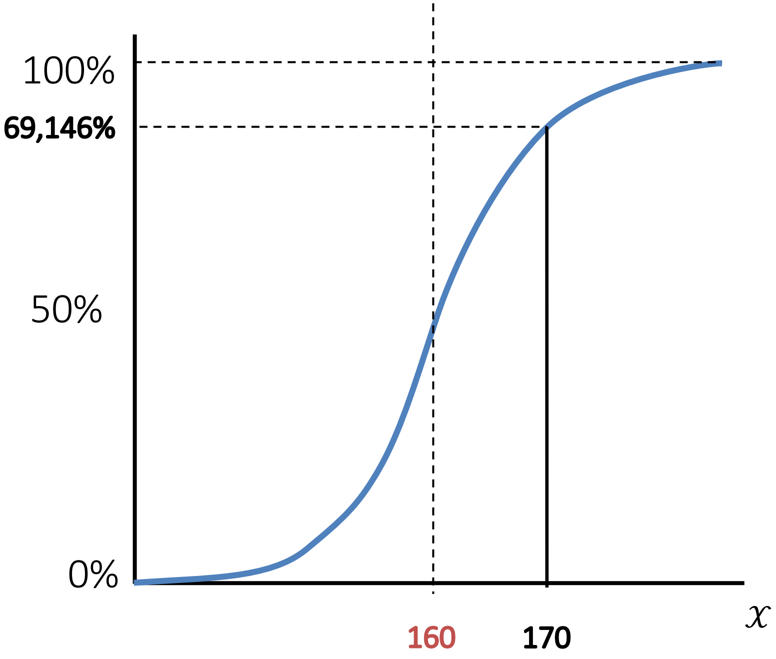




Probabilidad

Funciones de densidad y acumulada: Ejemplo





Creada a partir de: https://homepage.divms.uiowa.edu/~mbognar/applets/normal.html

3

Teorema del límite central

3

Teorema del límite central

Definiciones básicas

- **Experimento**: proceso que entrega uno o más resultados.
- **Resultados elementales**: Todos los posibles resultados más simples de un experimento.
- **Variable aleatoria**: Función que asigna un valor numérico único a cada resultado de un experimento. Pueden ser:
  - **Discretas**: Cuando puede tomar sólo un número finito de distintos valores.
  - **Continuas**: Puede asumir un número infinito de posibles valores (dentro de un rango).
- **Espacio muestral**: Conjunto de todos los resultados elementales

3

Teorema del límite central

Distribución muestral de parámetros

- Tal como se puede estimar la distribución de una variable aleatoria, se puede estimar la distribución muestral de un parámetro.
- **Definición**: es una distribución de probabilidad de un estadístico muestral calculado a partir de todas las muestras posibles de tamaño n, elegidas al azar en una población determinada.
- **Resume** la probabilidad de todos los resultados posibles de un experimento aleatorio.

3

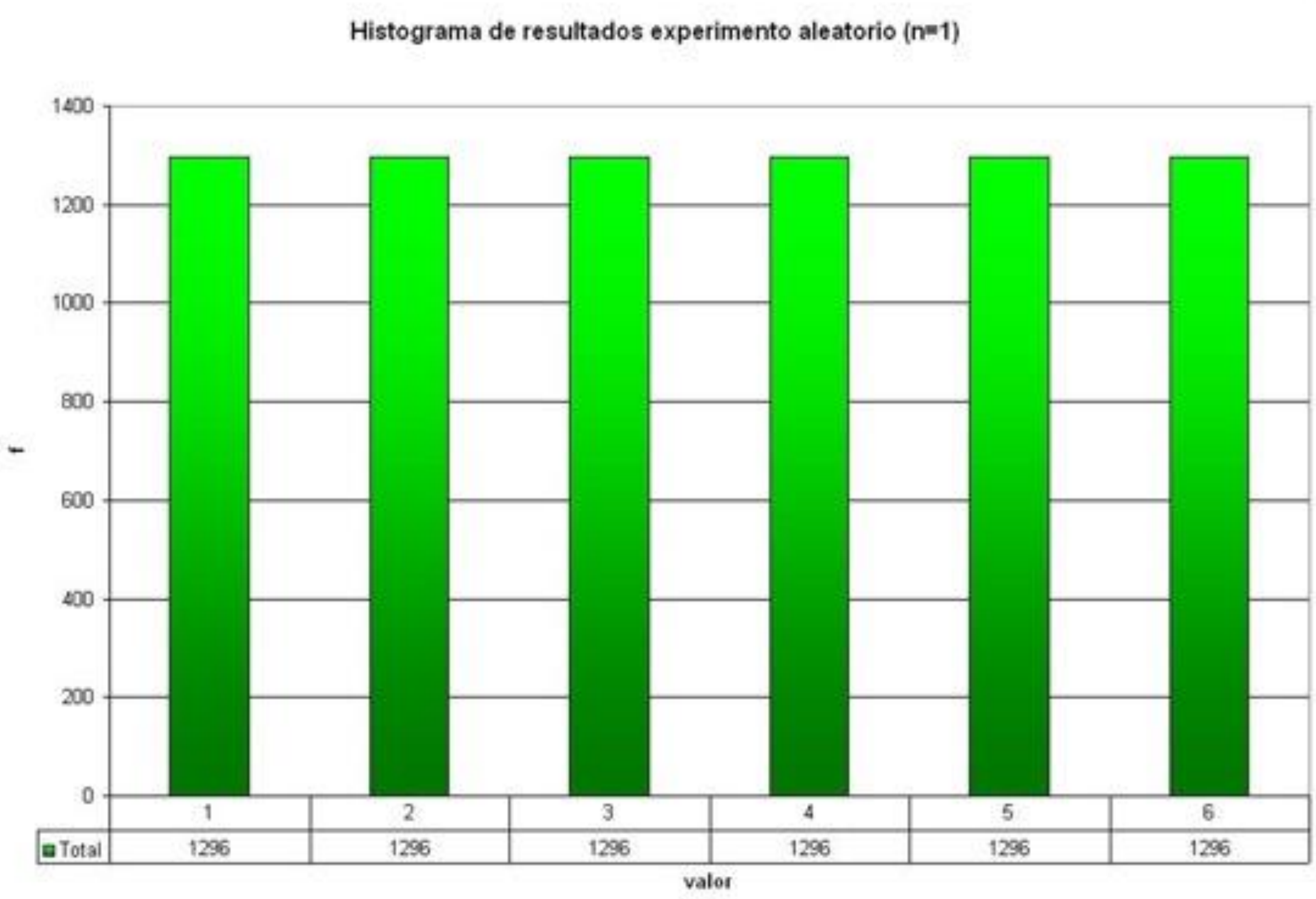
Teorema del límite central
Distribución muestral de parámetros

Imaginemos un experimento simple: **lanzar un dado**. Y veamos que pasa en las siguientes condiciones:

- Anotamos todos los valores obtenidos al lanzar miles de veces un dado (una población de lanzamientos de dados) y anotamos el promedio y la varianza de los resultados obtenidos
- Luego tratamos de ver qué pasaría si en vez de trabajar con la población tomamos muestras de ellas y
  - Generamos distribuciones para los resultados posibles usando dos, tres, cuatro y cinco dados (el promedio de los lanzamientos).
  - Anotamos el promedio y la varianza en cada caso.
  - Buscamos propiedades (patrones) en lo que encontramos.

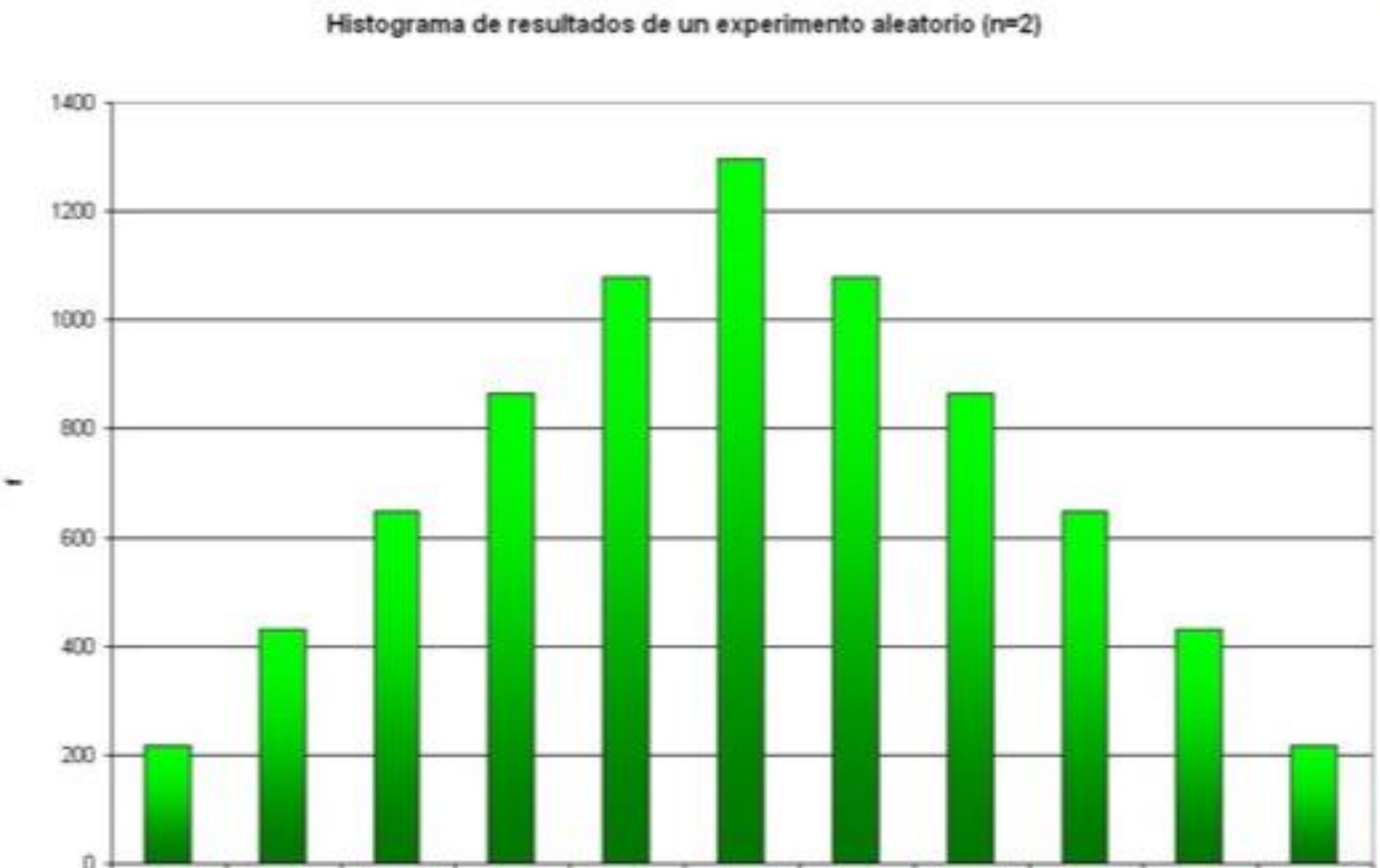
Teorema del límite central

Trabajando con la población



3 Teorema del límite central

Con muestras de 2 dados

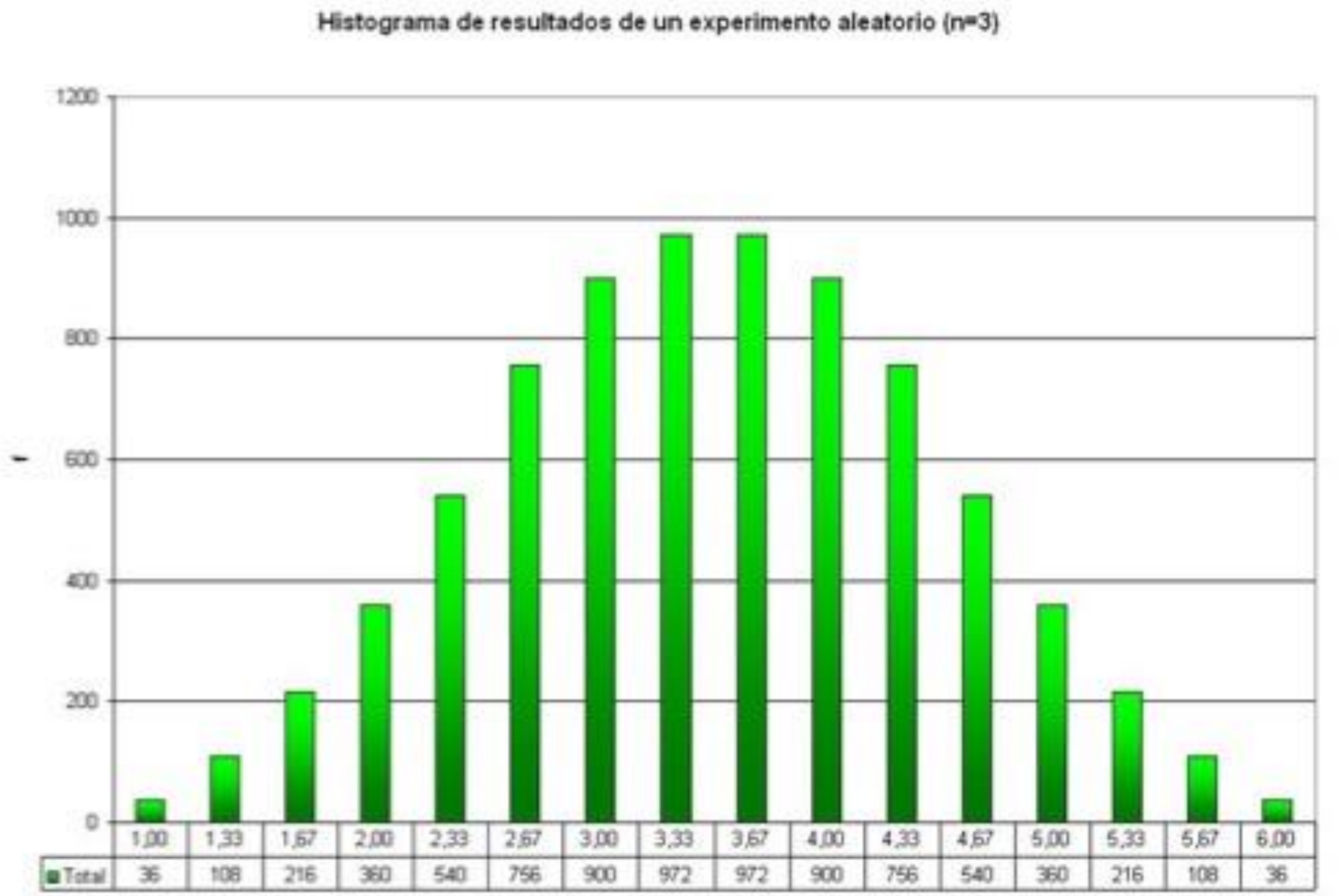


	1	1,5	2	2,5	3	3,5	4	4,5	5	5,5	6
Total	216	432	648	864	1080	1296	1080	864	648	432	216

Teorema del límite central

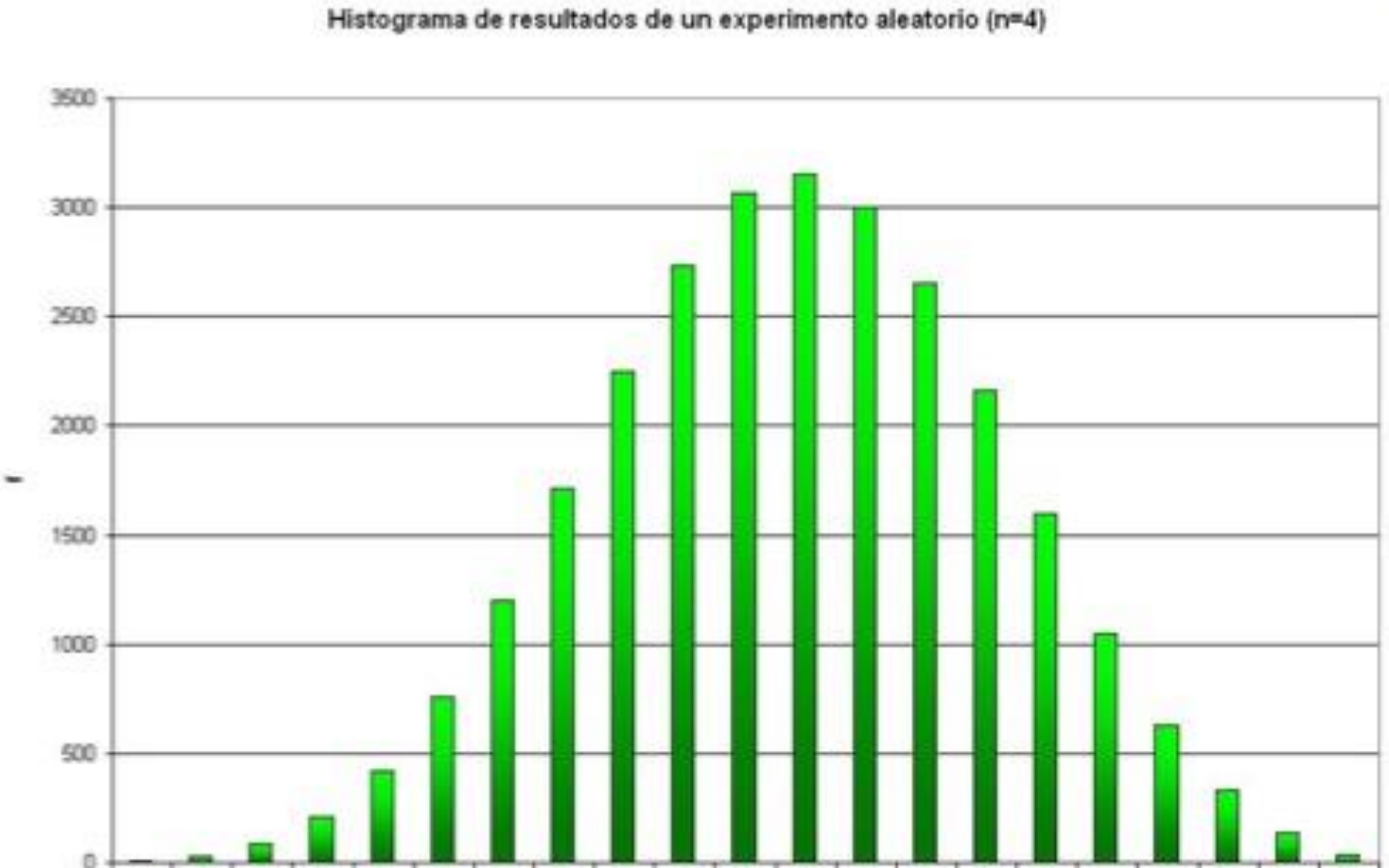
3

Muestras de 3 dados



3 Teorema del límite central

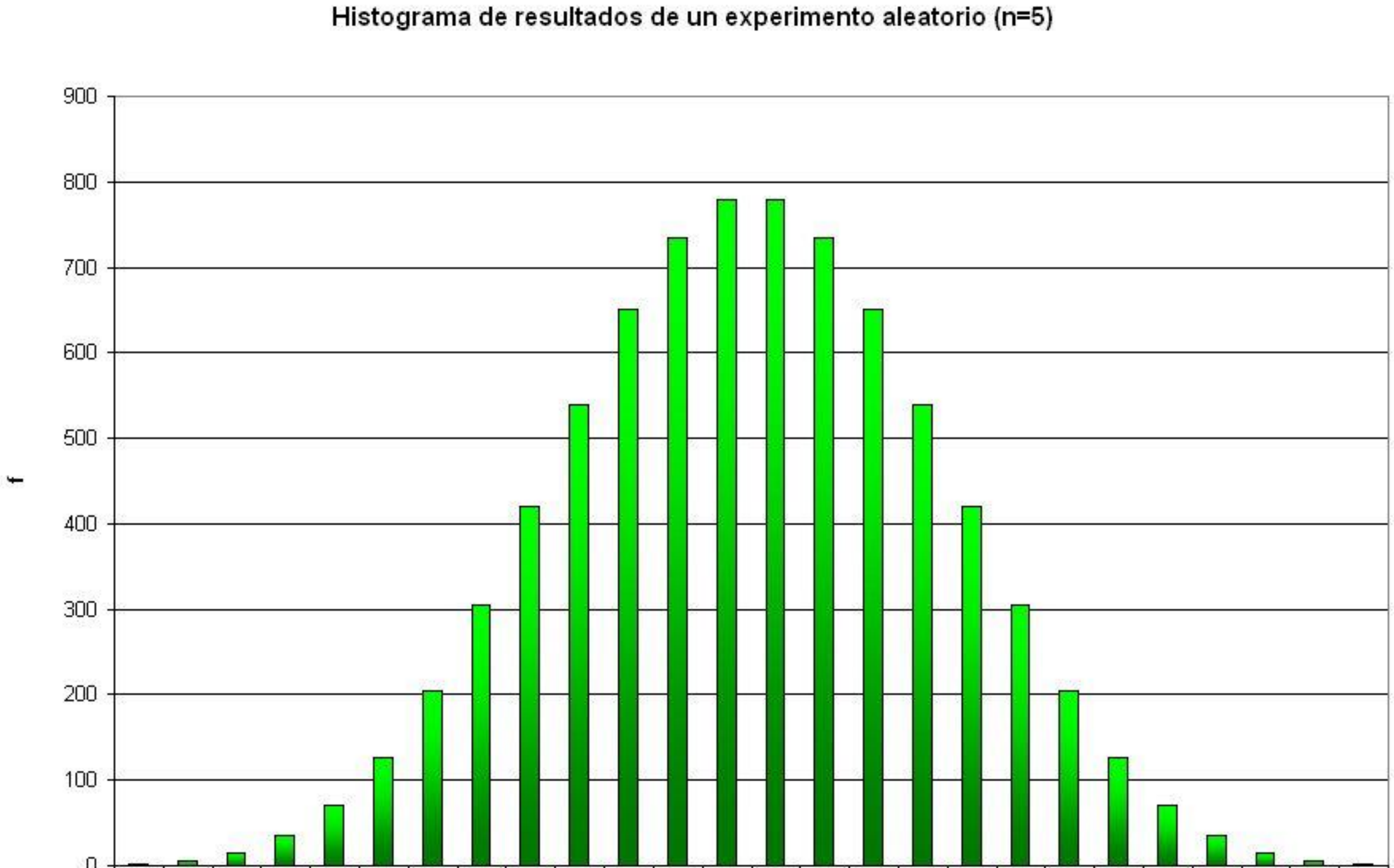
Muestras de 4 dados



	1	1,25	1,5	1,75	2	2,25	2,5	2,75	3	3,25	3,5	3,75	4	4,25	4,5	4,75	5	5,25	5,5	5,75	6
Total	6	30	90	210	420	756	1200	1716	2250	2730	3066	3150	3000	2652	2160	1596	1050	630	330	138	36

3 Teorema del límite central

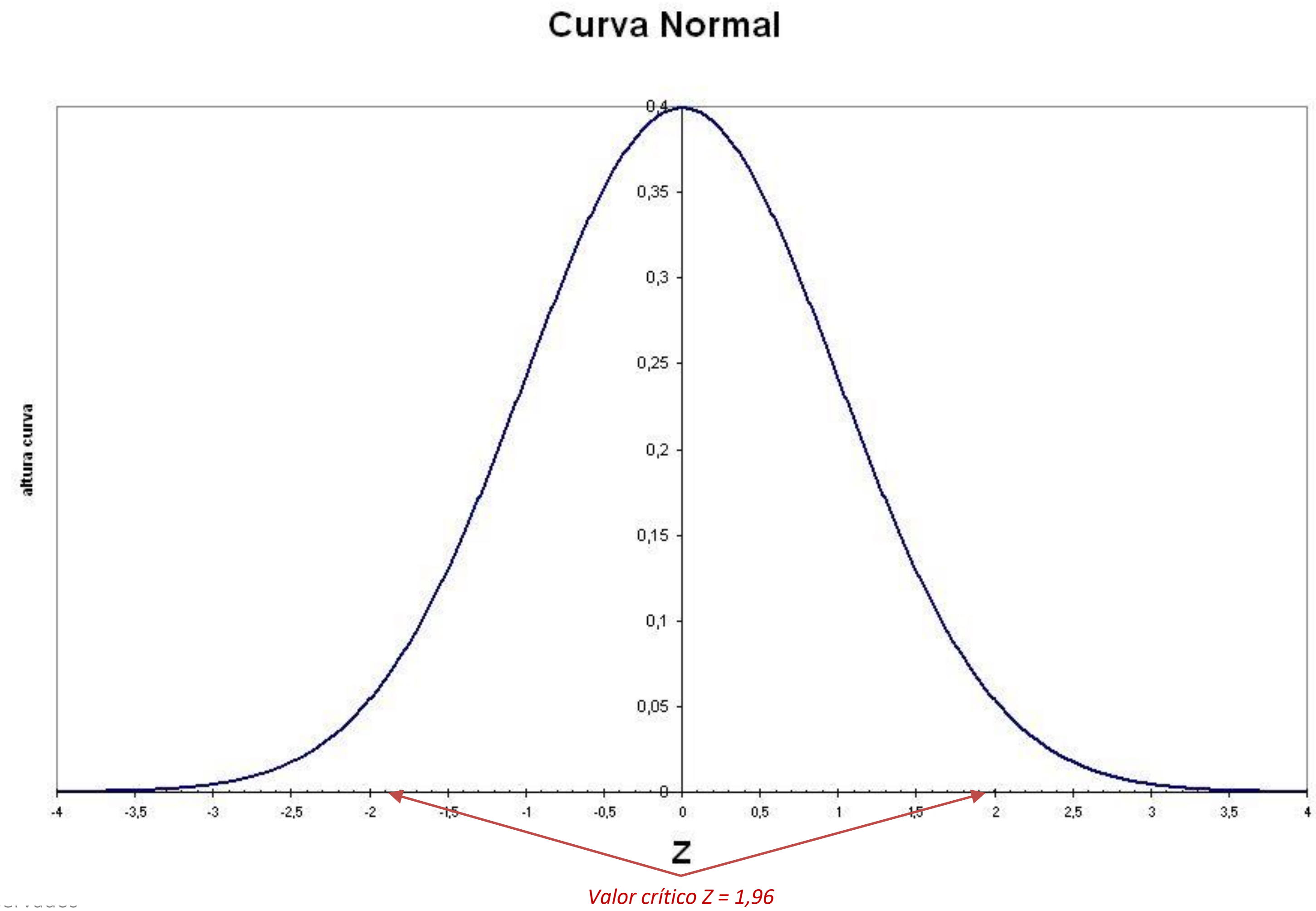
Muestras de 5 dados



	1	1,2	1,4	1,6	1,8	2	2,2	2,4	2,6	2,8	3	3,2	3,4	3,6	3,8	4	4,2	4,4	4,6	4,8	5	5,2	5,4	5,6	5,8	6
Total	1	5	15	35	70	126	205	305	420	540	651	735	780	780	735	651	540	420	305	205	126	70	35	15	5	1

Teorema del límite central

Distribución normal



3

Teorema del límite central

Resumen de promedios y desviación estándar según el n muestral

Número de dados	Media	Desv. Est.
1	3,5	2,92
2	3,5	1,46
3	3,5	0,97
4	3,5	0,73
5	3,5	0,58

3

Teorema del límite central

Propiedades

- Llamaremos **distribución muestral**:
  - Al conjunto de todos los resultados obtenidos al extraer todas las muestras aleatorias posibles, que tengan un tamaño n específico, en una población.
  - En rigor, es una distribución de probabilidad de un estadístico muestral calculado a partir de todas las muestras posibles de tamaño n, elegidas al azar en una población determinada.
- Si observamos los gráficos veremos que las distribuciones van tomando una forma aproximada a la **curva normal**.
- Los promedios se mantienen.
- Las varianzas de reducen en forma proporcional al n de la muestra.

3

Teorema del límite central

Propiedades

Características de los estimadores:

- **Insesgados:** valor esperado del estadígrafo se aproxima al valor del parámetro en la población.
- **Consistentes:** a medida que el tamaño de la muestra aumenta, el valor del estadígrafo se acerca cada vez más al valor del parámetro en la población.
- **Eficientes:** precisión del estadígrafo. Esto implica una baja variabilidad en la distribución muestral del estadígrafo.

3

Teorema del límite central

Enunciado

Si se extraen muestras aleatorias de tamaño n de una población con media μ y varianza (σ^2), la distribución muestral de los promedios ($\overline{X}$)

- tendrá una distribución aproximadamente normal,
- con media ($\mu_{\overline{X}}$) igual a μ,
- y varianza ($\sigma^2_{\overline{X}}$) igual a $\sigma^2\,/n$,

siempre que n sea lo **suficientemente grande**