


“Introducción a la Estadística, Probabilidad e Inferencia”

Maestría en Estadística Aplicada
Facultad de Ciencias Económicas y Estadística
UNR

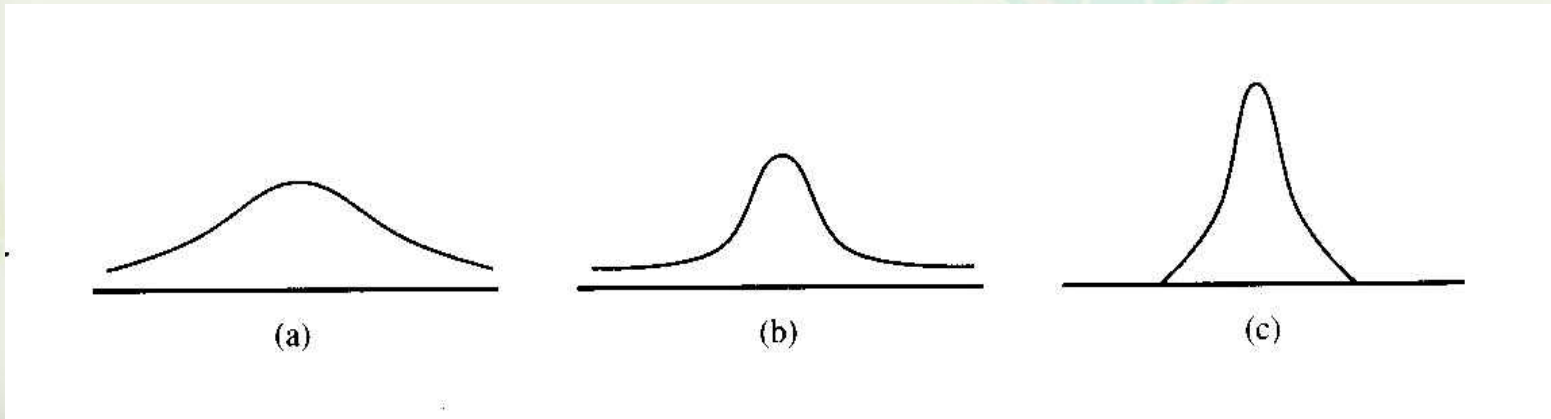
The background of the slide is a light gray gradient. It is decorated with numerous circular patterns that resemble stylized wheels or mandalas. These patterns are composed of concentric circles and radial lines, with segments in various colors including red, orange, yellow, green, and blue. The patterns are scattered across the slide, with some being larger and more prominent than others.

Material complementario de la Unidad 2 y de la Unidad 4

Forma de los histogramas

Una forma específica, la curva normal, aparece muy frecuentemente. Muchos histogramas pueden ser aproximados por una curva normal.

Por ejemplo: presión sanguínea, altura de personas adultas, resultados de un test de inteligencia.



Forma de los histogramas

Una curva normal no solo es simétrica sino que tiene una forma de campana.

Sin embargo, no todas las curvas con forma de campana son "normales".

Desde la parte superior de la campana, la curva decrece hacia ambas direcciones a una tasa definida por una función matemática.

- Una curva con colas que no disminuyen tan rápidamente como las de una distribución normal se dice que corresponde a una distribución con "*colas pesadas*".
- Similarmente, si las colas disminuyen más rápidamente que las de una distribución normal, se trata de una distribución con "*colas livianas*".

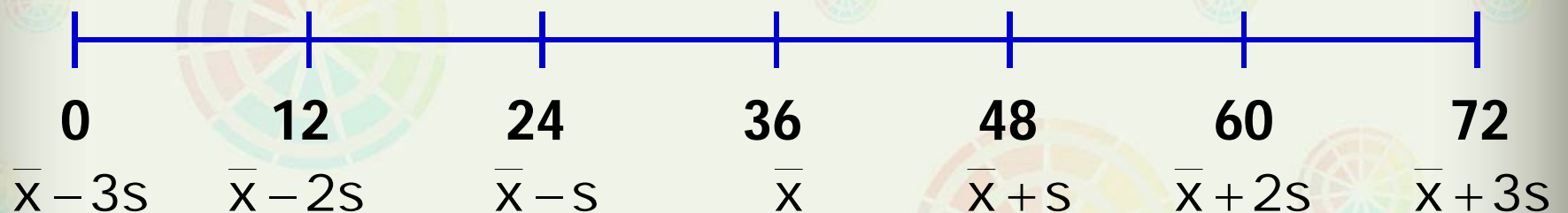
Regla de Chebyshev

Considerar cualquier número k , con $k \geq 1$. El porcentaje de observaciones que se encuentran dentro de k desvíos estándares de la media es al menos $100(1 - 1/k^2)\%$.

k	$1 - 1/k^2$	% dentro de k desvíos de la media
2	$1 - 1/4 = 0.75$	al menos 75%
3	$1 - 1/9 = 0.89$	al menos 89%
4	$1 - 1/16 = 0.94$	al menos 94%
4.472	$1 - 1/20 = 0.95$	al menos 95%
5	$1 - 1/25 = 0.96$	al menos 96%
10	$1 - 1/100 = 0.99$	al menos 99%

Ejemplo

Para una muestra de familias se reportó que los tiempos dedicados al cuidado de niños tenían una media de 36 horas por semana y un desvío estándar de 12.



Al menos 75% de las observaciones muestrales deben estar entre 12 y 60 horas.

Como al menos 89% de las observaciones deben estar entre 0 y 72, a lo sumo 11% estarán afuera de este intervalo y como el tiempo no puede ser negativo, se puede concluir que a lo sumo el 11% de las observaciones excede 72.

Regla Chebyshev

Dado que la regla de Chebyshev es aplicable a cualquier conjunto de datos, ya sea simétrico o asimétrico, se debe tener cuidado cuando se describe la proporción de observaciones ubicadas por encima o por debajo de determinado valor.

Si bien la regla establece que al menos 75% de las observaciones se encuentran dentro de 2 desvíos de la media, en muchos conjuntos de datos una proporción mayor de las observaciones se encontrará dentro de 2 desvíos de la media.

El hecho de que afirmaciones basadas en la regla de Chebyshev sean frecuentemente muy conservadoras sugiere que se debería buscar otra regla más precisa.

Regla empírica

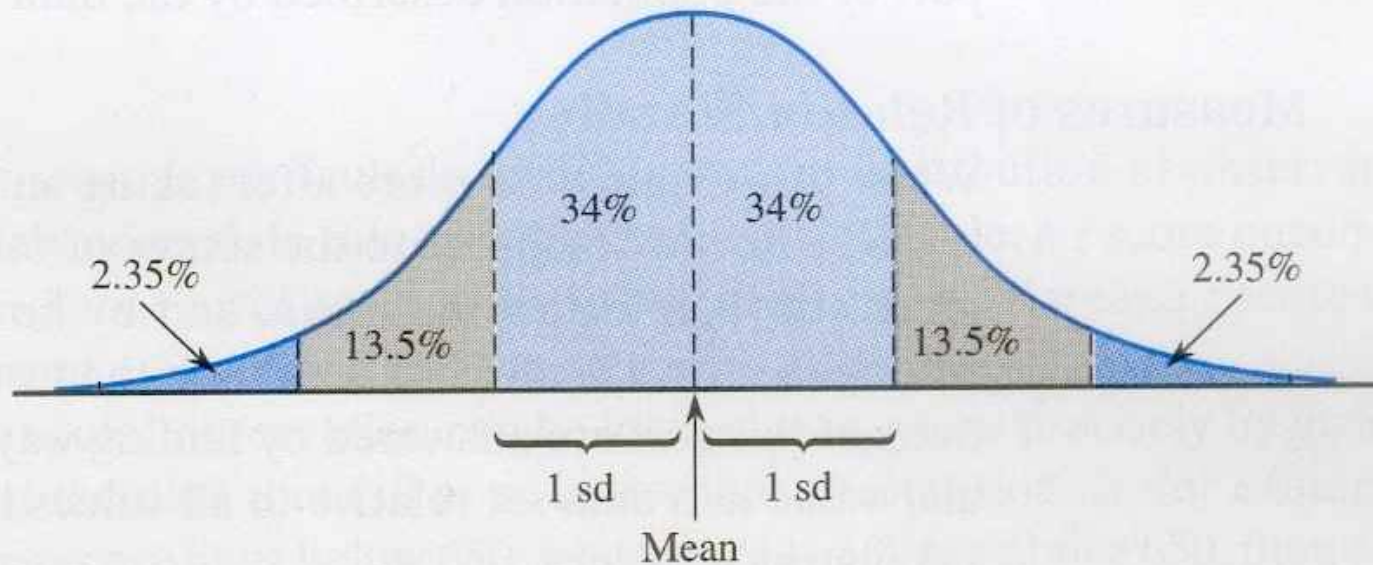
Una regla muy útil es la llamada **regla empírica** que puede ser aplicada siempre que la distribución de los datos pueda ser descripta por la distribución normal.

Si el histograma de los valores de un conjunto de datos puede ser razonablemente aproximado por una curva normal, entonces:

- Aproximadamente 68% de las observaciones están dentro de 1 desvío estándar de la media.
- Aproximadamente 95% de las observaciones están dentro de 2 desvíos estándares de la media.
- Aproximadamente 99.7% de las observaciones están dentro de 3 desvíos estándares de la media.

Regla empírica

La regla empírica utiliza "**aproximadamente**" en lugar de "**al menos**" y los porcentajes para $k=1, 2$ y 3 desvíos estándares son mucho mayores que los utilizados en la regla de Chebyshev.



Ejemplo

A partir de un conjunto de datos correspondiente a 1052 mujeres se obtuvo que la altura promedio es 158,7 cm y el desvío estándar 6,1 cm.

La curva normal resultó un buen ajuste para los datos.

Desvíos	Intervalo	Real	Regla Empírica	Regla de Chebyshev
1	152.6 – 164.8	72.1%	68.0%	al menos 0%
2	146.6 – 170.9	96.2%	95.0%	al menos 75%
3	140.5 – 176.9	99.2%	99.7%	al menos 89%