

Actividad 3: Estandarización de distribuciones normales

PROPÓSITO

Los estudiantes reconocen el término “normal” desde una mirada estadística y admiten que sirve para analizar conjuntos de datos y juzgar el comportamiento de uno o de un grupo de ellos, en comparación con el resto. Se espera que comprendan el uso de la desviación estándar y cómo tomar decisiones a partir de ella, que utilicen la tabla probabilística y eviten cálculos tediosos. Igual que en la actividad anterior, se espera una actitud de respeto hacia la privacidad de los compañeros, sectores y culturas.

Objetivos de Aprendizaje

OA 2: Fundamentar decisiones en situaciones de incerteza, a partir del análisis crítico de datos estadísticos y con base en los modelos binomial y normal.

OA c. Tomar decisiones fundamentadas en evidencia estadística y/o en la evaluación de resultados obtenidos a partir de un modelo probabilístico.

OA f. Evaluar modelos para estudiar un fenómeno, analizando críticamente las simplificaciones requeridas y realizando conexiones entre variables para predecir posibles escenarios de solución a un problema, y tomar decisiones fundamentadas.

Actitudes

- Actuar de acuerdo con los principios de la ética en el uso de la información y de la tecnología, respetando la propiedad intelectual y la privacidad de las personas.

Duración: 9 horas pedagógicas

DESARROLLO

ESTATURA DE RECIÉN NACIDOS

- En un Centro de Salud Familiar, se tiene registro de todos los niños a los que se ha atendido durante el primer semestre. Los siguientes datos corresponden al registro de la longitud (en centímetros) de todos los bebés de 6 meses controlados en ese periodo.

Conexión
interdisciplinaria:
**Ciencias para la
Ciudadanía**
OA c, 3° y 4° medio

59cm ; 60cm ; 62cm ; 63cm ; 65cm ; 65cm ; 66cm ; 67cm ; 67cm ; 68 cm ; 70cm ; 72cm

- a. Completa la tabla y calcula el promedio y la desviación estándar de las 12 estaturas.

Tabla 1: Registro de estaturas y detalle del cálculo de la desviación estándar

n	x_n	$x_n - \bar{x}$	$(x_n - \bar{x})^2$
1	59		
2	60		
3	62		
4			
5			
6			
7			
8			
9			
10			
11			
12			
Promedio			

$$S_x \approx \underline{\hspace{2cm}}$$

- b. ¿Cómo se interpreta el promedio y la desviación estándar en este contexto?
- c. Puedes usar la herramienta “Análisis de datos” de una planilla electrónica para calcular más rápido.
2. Respecto de la Figura 1 a continuación, responde:
- ¿Qué representan los puntos azules?
 - ¿Qué representa la línea vertical roja?
 - ¿Qué representa la línea vertical naranja?
 - ¿Qué representa las líneas horizontales verdes?

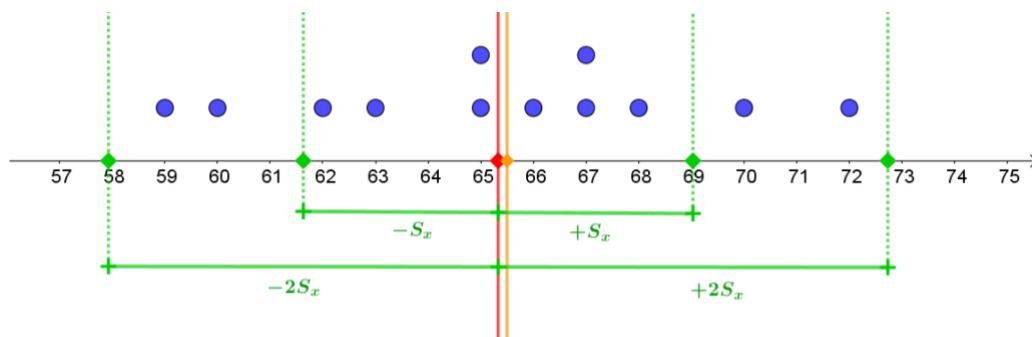


Fig. 1: Representación de las estaturas mediante diagrama de puntos.

3. Para interpretar el gráfico anterior, responde:
- ¿Qué información se puede obtener calculando $65,33 + 3,9 = 69,23$ y $65,33 - 3,9 = 61,43$?
 - ¿Qué datos se encuentran en el intervalo $(61,43; 69,23)$? ¿Qué porcentaje del total de datos representan?
 - ¿Es correcto afirmar que 8 de los 12 niños controlados tienen una estatura entre 61,43 cm y 69,23 cm?
 - ¿Qué puedes afirmar sobre la diferencia entre la estatura máxima y la mínima?
 - ¿Qué puedes afirmar sobre la cantidad de datos que se encuentran a dos desviaciones estándar o menos del promedio?
 - ¿Qué distribución crees que responde a los datos graficados? Conjetura una respuesta, aunque los datos sean pocos, y justifica.

¿CÓMO DESCRIBIMOS LA NORMALIDAD?

- Visita la página web DataChile (<https://www.curriculumnacional.cl/link/https://es.datachile.io/>) y extrae datos de algún tema que te interese: economía, educación, vivienda, demografía, salud, educación cívica.
 - Haz un histograma para observar la distribución de los datos.
 - Marca el polígono de frecuencias en el mismo gráfico. Describe la forma de la curva.
- Determina el promedio, la mediana y la desviación estándar para esos datos, utilizando una planilla electrónica.
 - Interpreta cada estadístico en relación con el contexto estudiado.
 - Compara los valores de la media y la mediana y señala una interpretación posible.
 - Marca en el histograma, con distintos colores, el promedio, la mediana y los intervalos $(\bar{x} - \sigma, \bar{x} + \sigma)$ y $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$
 - Argumenta si se puede asegurar o no que la distribución de estos datos corresponde a una distribución normal o es asimétrica.
- Determina la cantidad de datos que se encuentran en el intervalo $(\bar{x} - \sigma, \bar{x} + \sigma)$ y el porcentaje de dichos datos respecto del total.
 - ¿Cómo se relaciona con el porcentaje de datos que se encuentran a una desviación estándar de la media que se describe en una distribución de datos normal?
 - Determina la cantidad de datos que se encuentran en el intervalo $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$ y el porcentaje de dichos datos respecto del total.
 - ¿Cómo se relaciona con el porcentaje de datos que se encuentran a una desviación estándar de la media que se describe en una distribución de datos normal?

4. Sobre los aportes de la desviación estándar:
 - a. Señala al menos una interpretación que se puede dar de σ sobre un conjunto de datos.
 - b. Indica cómo se usa para determinar si un dato específico está dentro de los márgenes de normalidad respecto de los demás datos de un mismo estudio (o experimento).
 - c. ¿Qué se considera normal en el contexto seleccionado?

UTILIZACIÓN DE LA TABLA PROBABILÍSTICA PARA Z

1. Observa la Figura 2. ¿Qué tipo de distribución tienen los datos de esta población?
 - a. ¿Cuál es la media? ¿Cuál es la desviación estándar?
 - b. ¿Puedes inferir el valor de la mediana y la moda? ¿Cuáles deberían ser?
 - c. ¿Cuál es el área total bajo la curva? ¿Cómo se relaciona con las probabilidades de los datos?
 - d. Porcentualmente, ¿cuántos datos se encuentran a 1, 2 o 3 desviaciones estándar de la media?
 - e. Relaciona la respuesta anterior con la probabilidad de que un dato de esta población tomado al azar se encuentre entre $(\bar{x} - \sigma, \bar{x} + \sigma)$, entre $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$ y entre $(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$.

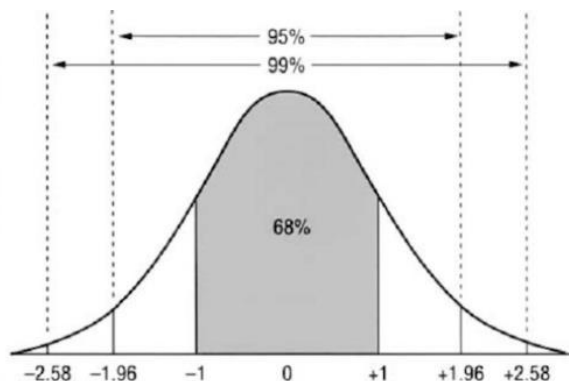


Fig. 2: Distribución normal estándar.

2. En internet⁴ puedes encontrar una lista de las probabilidades asociadas con los intervalos desde la media ($z = 0,00$) hasta un valor específico de z . Se puede hallar las probabilidades de otros intervalos usando las entradas de la tabla y las operaciones de adición y sustracción, de acuerdo a las propiedades de las probabilidades.

Por ejemplo:

$$P(0 < z < 1,52) = 0,4357$$

Esto quiere decir que la probabilidad de tomar al azar un dato de la población distribuida normalmente que se encuentre entre 0 y 1,52, es 0,4357.

⁴ Por ejemplo: <https://www.curriculumnacional.cl/link/https://es.slideshare.net/josejoaquinmunoz/tablas-de-distribucion-normal-con-la-probabilidad-definitiva-con-todos-los-valores-de-z>

La Figura 3 muestra cómo usar la tabla de probabilidades:

Z	0	0.01	0.02
0.0	0.0000	0.0040	0.0080
0.1	0.0398	0.0438	0.0478
0.2	0.0793	0.0832	0.0871
0.3	0.1179	0.1217	0.1255
0.4	0.1554	0.1591	0.1628
0.5	0.1915	0.1950	0.1985
0.6	0.2257	0.2291	0.2324
0.7	0.2580	0.2611	0.2642
0.8	0.2881	0.2910	0.2939
0.9	0.3159	0.3186	0.3212
1.0	0.3413	0.3438	0.3461
1.1	0.3643	0.3665	0.3686
1.2	0.3849	0.3869	0.3888
1.3	0.4032	0.4049	0.4066
1.4	0.4192	0.4207	0.4222
1.5	0.4332	0.4345	0.4357
1.6	0.4452	0.4463	0.4474

Fig. 3: Parte de la tabla de probabilidades de la distribución normal estándar.

- Determina $P(0 < z < 0,91)$. ¿Qué interpretación tiene el resultado?
 - Determina $P(-1 < z < 0,91)$. ¿Cómo se debe usar la tabla en este caso? ¿Qué propiedades de las probabilidades te permiten llegar a la respuesta?
 - Determina $P(-2 < z < -1,1)$. ¿Qué interpretación tiene el resultado?
 - Determina $P(z < -1,1)$. ¿Qué interpretación tiene el resultado?
 - Determina $P(-2 < z)$. ¿Qué interpretación tiene el resultado?
- Volviendo al caso de las estaturas de los recién nacidos.
 - Recuerda cuál es el valor de \bar{x} y σ .
 - Recuerda también que se podía obtener ciertas probabilidades aproximadas de valores x . Por ejemplo:

$$P(47,8 < x < 51,2) \text{ y } P(46 < x < 53)$$

ORIENTACIONES PARA EL DOCENTE

1. Se espera que analicen los aportes de la media y la desviación estándar para caracterizar una población con distribución normal. Ya han estudiado antes estas medidas, de tendencia central y de dispersión respectivamente, en distribuciones de datos no necesariamente normales.
2. Se propone un análisis con una cantidad muy limitada de datos, 12 estaturas, para que luego analicen 650 estaturas. Con esto, se espera que perciban que, a mayor cantidad de datos, más evidente es la aproximación a la curva normal y que, por ser datos empíricos, hay pequeñas variantes respecto de la definición. Por ejemplo: en este caso la media se acerca mucho a la mediana, pero no son iguales. Además, la cantidad de datos alejados una desviación estándar de la media (cantidad de datos en el intervalo $(\bar{x} - \sigma, \bar{x} + \sigma)$) no es exactamente 68%, sino una cantidad próxima. Lo mismo ocurre con los 12 datos en el intervalo a dos desviaciones estándar de la media; en ese caso, el 100% de los datos pertenece a ese intervalo, a diferencia del 95% esperado.
3. Se sugiere que examinen qué representa una distribución normal estándar, junto con el valor de su media, moda, mediana, desviación estándar, y los intervalos a una o dos desviaciones estándar de la media. Se profundiza en el cálculo de probabilidades de un dato z cualquiera, que cumpla con una condición dada –por ejemplo: $P(a < z < b)$ –, usando el puntaje z y la tabla de probabilidades.
4. Se sugiere los siguientes indicadores para evaluar formativamente los aprendizajes:
 - Utilizan la tabla de probabilidades para determinar la probabilidad de tomar, en forma aleatoria, un dato de una población distribuida normalmente.
 - Evalúan la pertinencia de usar modelos binomial o normal para interpretar situaciones de incerteza.
 - Evalúan los alcances y límites de un argumento estadístico o probabilístico antes de tomar una decisión.

RECURSOS Y SITIOS WEB

Sitios sugeridos para estudiantes y profesores:

- ¿Qué es la desviación estándar?
<https://www.curriculumnacional.cl/link/https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/data-concepts/what-is-the-standard-deviation/>
- Tabla de distribución normal Z
<https://www.curriculumnacional.cl/link/http://matepedia-estadistica.blogspot.com/2016/01/tabla-de-distribucion-normal.html>
- Distribución normal: ¿Es tan frecuente como parece?
https://www.curriculumnacional.cl/link/http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0034-98872012000400021
- Normal, simetría o asimetría
<https://www.curriculumnacional.cl/link/http://facilestadistica.wixsite.com/estadisticafacil/single-post/2014/08/24/Estad%C3%ADstica-Descriptiva-III-Medidas-de-Forma-y-Medidas-de-Concentraci%C3%B3n>
- Datos públicos de Chile
<https://www.curriculumnacional.cl/link/https://es.datachile.io/>