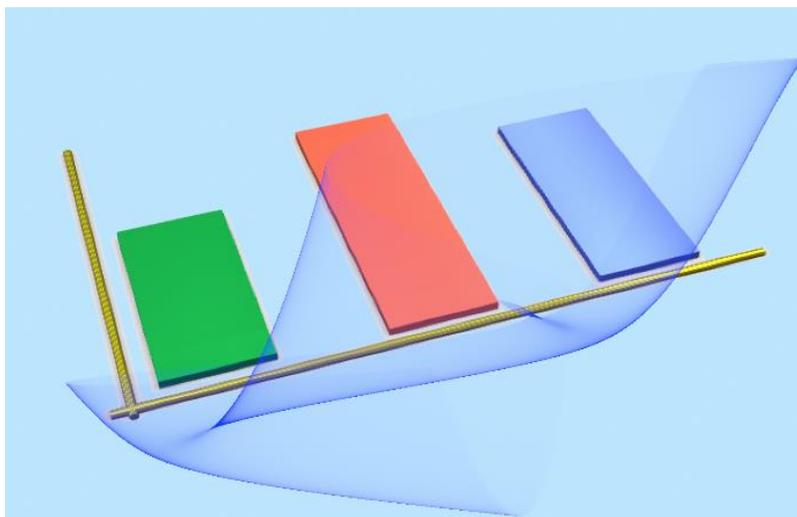


REFORMA DE LA EDUCACIÓN MATEMÁTICA EN COSTA RICA



Colección *Preparación Matemáticas Bachillerato*



ESTADÍSTICA Y PROBABILIDAD: **Tablas y gráficos estadísticos**

Material complementario

Costa Rica

2017

Contenido

I.	Introducción	3
1.	Representaciones estadísticas con datos cualitativos: valores absolutos y porcentuales	3
	Problema 1. Deserción y selección de área científica	3
	Problema 2. Selección de área científica según sexo del estudiante	6
2.	Representaciones estadísticas con datos cuantitativos	8
	Problema 3. Disminución de matrícula escolar	8
	Problema 4. Producción de piezas electrónicas	9
	Modelos simétricos y asimétricos	10
	Problema 5. Vida útil de los bombillos	11
	Problema 6. Selección de adolescentes	12
	Problema 7. Enfermos por picadura de mosquito	14
	Problema 8. Consumo promedio anual de electricidad por abonado residencial	15
	Problema 9. Enfermos por picadura de mosquito	16
	Problema 10. Años de experiencia en labores docentes	17
3.	Representación de la distribución de datos mediante un diagrama de cajas	18
	Diagrama de cajas	18
	Problema 11. Vida útil de los bombillos	19
	Problema 12. Defunciones por accidentes de tránsito según provincia	21
II.	Bibliografía	23
III.	Créditos	23

Conceptos:

[Asimetría](#) (p. 10)

[Diagrama de cajas](#) (p. 18)

I. Introducción

El presente documento ha sido elaborado por el Proyecto Reforma de la Educación Matemática en Costa Rica. Su propósito básico consiste en apoyar a estudiantes para la preparación de las Pruebas Nacionales de Matemáticas en el análisis de las medidas estadísticas.

Este documento complementa el Mini MOOC denominado *Tablas y gráficos estadísticos* de la colección *Preparación Matemáticas Bachillerato* (<http://minimoocs.reformamatematica.net>) diseñado por este mismo proyecto.

Se analizan diferentes conocimientos vinculados con el uso de las medidas estadísticas de posición y variabilidad para la resolución de problemas de acuerdo con las temáticas incluídas en los Programas de Estudios de Matemáticas para la Educación Diversificada. Se incluyen además

Esperamos que el material sea de provecho para todos aquellos lectores que lo utilicen.

1. Representaciones estadísticas con datos cualitativos: valores absolutos y porcentuales

[Volver a Contenido](#)

Problema 1. Deserción y selección de área científica

Suponga que usted estudia en el Liceo El Progreso, que es una institución académica en la cual en el año 2015 se matricularon 192 estudiantes en undécimo año, pero 12 de ellos se retiraron antes de concluir el año académico, los demás aprobaron el curso lectivo.

De ellos, 18 se inscribieron para presentar Física como área científica para las pruebas de bachillerato, mientras 46 optaron por realizar la prueba de Química y el resto se inscribieron para presentar Biología. El director del colegio desea resumir esta información de manera que sea comprensible para entregar un informe a los padres de familia.



<http://www.freepik.es/>

Ayude al señor Director a resumir esta información para que la incluya en dicho informe y resulte comprensible para los padres de familia.

Solución:

En este caso la unidad estadística es el estudiante de noveno año del Liceo el Progreso, las variables de interés del estudio son:

- *Condición de desertor*, cuyas respuestas pueden ser: (1) desertor o (2) no desertor
- *Especialidad científica* que va a presentar en bachillerato: (1) Física, (2) Química y (3) Biología.

Los datos que generan estas variables no son numéricos, debido a que para cada una se anota lo siguiente. Observe que se pueden utilizar números para representar las categorías o respuestas vinculadas con cada estudiante, pero es meramente simbólico, cualquier número que se puede emplear se utiliza para simplificar la escritura.

En primer lugar, cuando se analiza la condición de desertor solamente caben dos posibilidades para cada uno de los estudiantes matriculados (desertor o no desertor), entonces la descripción del comportamiento es sencilla, basta con determinar el número de estudiantes que abandonaron la institución, en este caso fueron 12, por lo que 180 no se retiraron y aprobaron el curso. Sin embargo, estos datos por sí solos no son suficientemente ilustrativos desde el punto de vista estadístico. Se requiere determinar un valor relativo o porcentual que permita hacer comparaciones con otras instituciones o con otros años. Entonces el Director podría utilizar porcentajes en el informe:

- *En el año 2015, en el Liceo El Progreso desertaron 12 estudiantes de un total de 192 que matricularon undécimo año, por ello el porcentaje de deserción fue $\frac{12}{192} \cdot 100\% = 6,25\%$. Esto es equivalente a decir que de los 192 estudiantes que se matricularon en undécimo año en esta institución, el 93,75% logró aprobar el año.*

En segundo lugar, en cuanto al área de ciencias que escogieron los estudiantes, de los 180 jóvenes con derecho a realizar las pruebas de bachillerato, 18 seleccionaron Física, 46 Química y por lo tanto $180 - 18 - 46 = 116$ escogieron Biología. Para dar una mejor interpretación a estos números se pueden determinar los porcentajes correspondientes. Se recomienda al Director indicar en el informe lo siguiente:

- *En el Liceo El Progreso para el año 2015, de los 180 estudiantes que debían realizar las pruebas de bachillerato, 18 estudiantes seleccionaron el área de Física es decir el $\frac{18}{180} \cdot 100\% = 10,0\%$, por su parte 46 estudiantes seleccionaron el área de Química que corresponde al $\frac{46}{180} \cdot 100\% = 25,6\%$, (aproximadamente), y finalmente, 116 estudiantes seleccionaron el área de Biología, que representa el $\frac{116}{180} \cdot 100\% = 64,4\%$ (aproximadamente)*

Otra forma de presentar esta información corresponde a un pequeño cuadro:

Liceo El Progreso: área científica seleccionada por los 180 estudiantes que realizaron la prueba de Bachillerato, año 2015

Área científica seleccionada	Número de estudiantes	Porcentaje de estudiantes
Física	18	10,0
Química	46	25,6
Biología	116	64,4
Total	180	100,0

Información ficticia utilizada con fines didácticos

En el problema anterior, aunque los datos originales eran cualitativos (no numéricos), se debió realizar un conteo de cada categoría lo que a la postre generó la suma total en cada caso y los porcentajes correspondientes, de modo que la variabilidad de los datos fuera comprensible para un lector. En estos casos los porcentajes constituyen una medida estadística adecuada para resumir la información. Esta alternativa se vuelve fundamental cuando se requiere realizar comparaciones entre grupos con diferentes variables.

[Volver a Contenido](#)

Problema 2. Selección de área científica según sexo del estudiante

Con base en el problema anterior, suponga ahora que el Director del Liceo El Progreso desea presentar la información sobre el área científica seleccionada por los estudiantes; pero desea realizar una comparación por sexo (hombres y mujeres)



Decide presentar en el informe el siguiente cuadro:

Liceo El Progreso: área científica seleccionada por los 180 estudiantes que realizaron la prueba de Bachillerato según el sexo, año 2015

Área científica seleccionada	Número de estudiantes	
	Hombres	Mujeres
Física	9	9
Química	16	30
Biología	37	79
Total	62	118

Información ficticia utilizada con fines didácticos

¿Qué críticas haría usted al señor Director en relación con el cuadro anterior, si lo que se pretende es realizar una comparación por sexo en la escogencia del área científica, de modo que sea de fácil comprensión para los padres de familia?

Solución:

Cuando se desean realizar comparaciones entre grupos, no es conveniente utilizar los números absolutos, debido a que normalmente ocurre que los números totales entre los grupos son diferentes (en el ejemplo anterior hay 62 varones contra 118 mujeres), entonces los valores parciales de cada categoría no son comparables. Por ejemplo, sería un error indicar que la preferencia por el área de Física fue igual entre hombres y mujeres basados únicamente en el hecho que nueve hombres y nueve mujeres seleccionaron esta área. Pero también es error señalar que las mujeres presentaron más preferencia al área de Química que los hombres.

Para realizar una lectura adecuada de los datos del cuadro se requiere hacer una comparación porcentual, tal como se indica en el cuadro siguiente:

Liceo El Progreso: área científica seleccionada por los estudiantes que realizaron la prueba de Bachillerato según el sexo, año 2015

Área científica seleccionada	Hombres		Mujeres	
	Total	Porcentaje	Total	Porcentaje
Física	9	14,5	9	7,6
Química	16	25,8	29	24,6
Biología	37	59,7	80	67,8
Total	62	100,0	118	100,0

Información ficticia utilizada con fines didácticos

Al comparar los porcentajes se puede notar que hubo mayor preferencia de los hombres por el área de Física y de las mujeres por el área de Biología; pero en el área de Química fue similar la preferencia. Entonces se debe recomendar al Director que sustituya el cuadro que quería incluir por este otro que resulta más ilustrativo.

Los problemas anteriores realizan un análisis de variables cualitativas (no numéricas) por medio de porcentajes. Sin embargo, si las variables son cuantitativas (numéricas) entonces se requiere recurrir a otras técnicas porque los porcentajes no aplican para ellas.

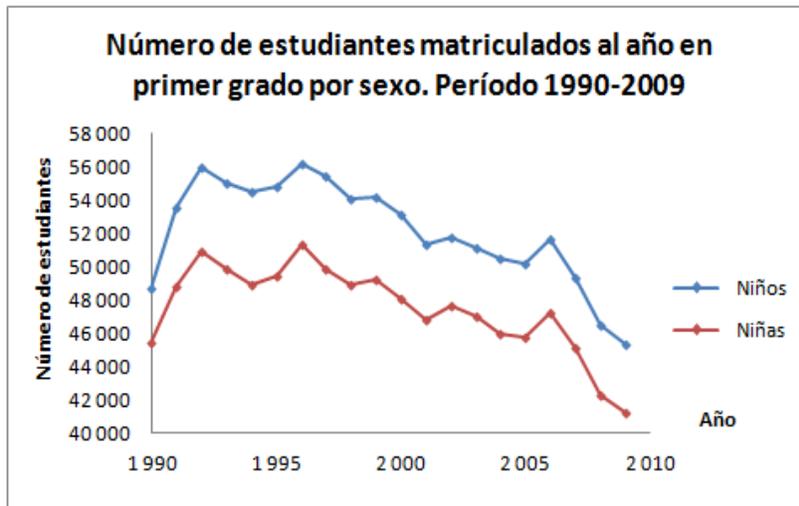
[Volver a Contenido](#)

2. Representaciones estadísticas con datos cuantitativos

[Volver a Contenido](#)

Problema 3. Disminución de matrícula escolar

La matrícula anual de estudiantes en primer grado ha disminuido en los últimos años, la siguiente gráfica refleja ese comportamiento.



Fuente: <http://www.estadonacion.or.cr/index.php/estadisticas/costa-rica/compendio-estadistico/estadisticas-sociales>

Además, se observa una importante diferencia entre la matrícula de hombres y de mujeres, la cual se mantiene durante todo el período. Sin embargo, el director de una pequeña escuela rural indica que la matrícula en primer grado ha venido aumentando en los últimos años, y normalmente se matriculan más niñas que niños. Considere las siguientes proposiciones que se elaboraron de acuerdo con la información que suministra del gráfico. Identifique aquellas que son verdaderas:

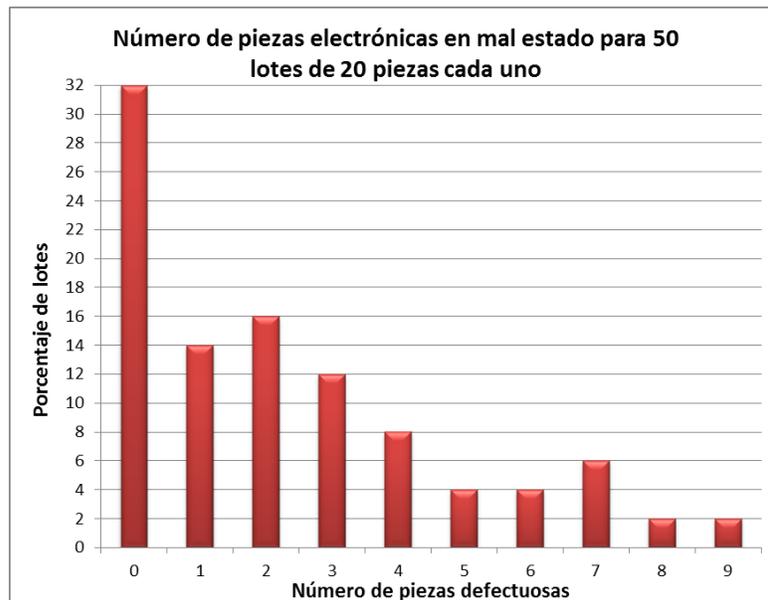
- I. La información del director es falsa, pues es de esperar que los resultados generales de la gráfica deben reflejarse en cada una de las escuelas.
- II. Los resultados proporcionados por el director reflejan que la información de la gráfica debe estar incorrecta.
- III. Los resultados proporcionados por el director de la escuela no contradicen la información de la gráfica.

Solución

La representación gráfica resume el comportamiento general del número de estudiantes matriculados en primer grado de primaria en el país según sexo, lo cual corresponde al comportamiento de la población total del país. Una pequeña escuela rural representa un caso específico o una muestra particular, por lo que puede no cumplir con el patrón representado en la población. De acuerdo con esto, los resultados proporcionados por el director de la escuela no contradicen la información de la gráfica. Por ello, la proposición III es la única correcta.

Problema 4. Producción de piezas electrónicas

Una fábrica empaqueta en lotes de 200 unidades las piezas electrónicas que produce. Se establece un plan de inspección por muestreo que consiste en examinar 20 piezas de cada lote elegidas al azar. Se rechaza el lote completo si de las 20 piezas aparecen tres o más defectuosas. Se revisan 50 lotes y la distribución del número de piezas defectuosas por lote se muestra en la siguiente gráfica:



Determine el número de lotes que deberían ser rechazados.

Solución:

En la representación gráfica se muestra el comportamiento de 50 lotes de piezas electrónicas. En cada uno de ellos se han seleccionado 20 piezas aleatoriamente y se determina si cada una es defectuosa o no lo es. Debido a que se rechazan aquellos lotes para los cuales hay tres o más piezas defectuosas, del gráfico se extrae la siguiente información:

Número de piezas defectuosas	Porcentaje de lotes
0	32
1	14
2	16
3	12
4	8
5	4
6	4
7	6
8	2
9	2
Total	100%

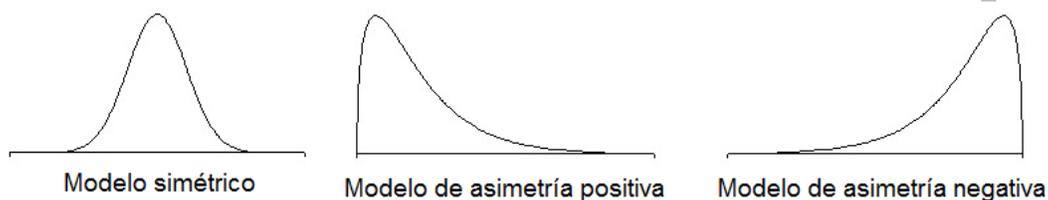
Un recuadro gris resalta los valores de porcentaje para 3, 4, 5, 6, 7, 8 y 9 piezas defectuosas, con una llave azul que indica un total de 38%.

Los valores en azul representan los porcentajes de lotes que deberían ser rechazados, esto quiere decir que el 38% de los lotes tienen tres o más piezas defectuosas. Entonces de los 50 lotes, el 38% debería ser rechazado, esto equivale a 19 lotes en total.

[Volver a Contenido](#)

Modelos simétricos y asimétricos

En el análisis de los problemas anteriores, interpretar la forma o patrón de variación de los datos fue clave para dar respuesta a los cuestionamientos planteados. La distribución de los datos puede tomar diversas formas, pero en términos generales existen tres patrones básicos (aunque pueden existir otros):

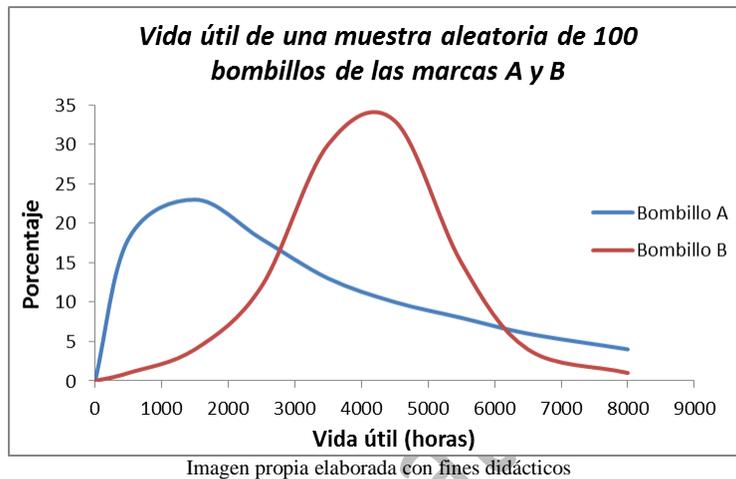


Cuando la forma de la distribución de los datos es simétrica o aproximadamente simétrica, significa que la mayor cantidad de datos se ubica en el centro, con un comportamiento similar a ambos lados. Mientras que en las distribuciones asimétricas las mayores frecuencias de datos tienden a concentrarse hacia uno de los extremos. En ellas se observa la presencia de pocos valores en la dirección contraria, pero que tienen un importante efecto en el mensaje de los datos, tal como veremos más adelante.

[Volver a Contenido](#)

Problema 5. Vida útil de los bombillos

Suponga que le envían a comprar unos bombillos a una ferretería, la persona que le atiende indica que de acuerdo con las características del bombillo que usted está solicitando, tiene dos marcas con el mismo precio: A y B. Le pregunta que cuál de ellas desea. Usted recuerda que en un artículo de una revista que leyó recientemente salió publicado la siguiente representación gráfica sobre la vida útil de estos bombillos:



Utilice esta información para decidir qué marca de bombillos compraría y argumente esta decisión.

Solución:

Cuando se compra un bombillo, como con cualquier otro producto electrónico no tenemos certeza de cuánto tiempo va a funcionar, esta duración o vida útil varía de un bombillo a otro. Para resumir la variabilidad generada entre los bombillos, se pueden utilizar diferentes herramientas estadísticas. La representación gráfica es una de las herramientas que permite modelar la variabilidad en la vida útil de los bombillos y que es muy utilizada en diferentes estudios.

La representación gráfica anterior corresponde a una aproximación de este modelo y se basa en muestras aleatorias de 100 bombillos de los tipos A y B. En el eje "X" se ha representado la vida útil (duración) y en el eje "Y" el porcentaje de bombillos. El área que está encerrada bajo las curvas representa el porcentaje de bombillos que tienen la vida útil correspondiente, entonces en los intervalos donde hay más área significa que hay un mayor porcentaje de bombillos que tiene una vida útil en ese rango.

Por ejemplo, se presentó un mayor porcentaje de bombillos de la marca A que tienen una vida útil menor de 3000 horas, con respecto a los bombillos de la marca B. Asimismo un mayor porcentaje de bombillos de la marca B tuvo una vida útil entre 3000 y 6000 horas respecto a la marca A. También se observa que hay mayor porcentaje de bombillos de la marca A que duran más de 6000 horas.

De acuerdo con lo anterior, debido a que no podemos estar seguros de cuánto nos va a durar de un bombillo que vamos a comprar, entonces deberíamos adquirir bombillos de la marca B, debido a que según la representación gráfica anterior hay más posibilidad de que estos bombillos nos duren más.

Hay que tener claro que el análisis anterior se fundamenta en el hecho de que la información de estas muestras debe ser una buena representación de lo que ocurre en la producción total de bombillos.

[Volver a Contenido](#)

Problema 6. Selección de adolescentes

Suponga que un asistente médico requiere hacer un estudio con adolescentes (jóvenes entre 12 y 17 años) en una clínica de salud. Debe escoger entre dos clínicas A y B, pero no sabe en cuál de ellas tiene más opciones de seleccionar adolescentes. Solicita colaboración a la administración de las clínicas para que le indiquen la edad (en años cumplidos) de los menores que asistieron a la clínica en la última semana. Debido a que las personas encargadas de proporcionar la información les gustan las representaciones estadísticas, en vez de darle los datos en forma simple, se pusieron de acuerdo para proporcionar la siguiente información al asistente médico.

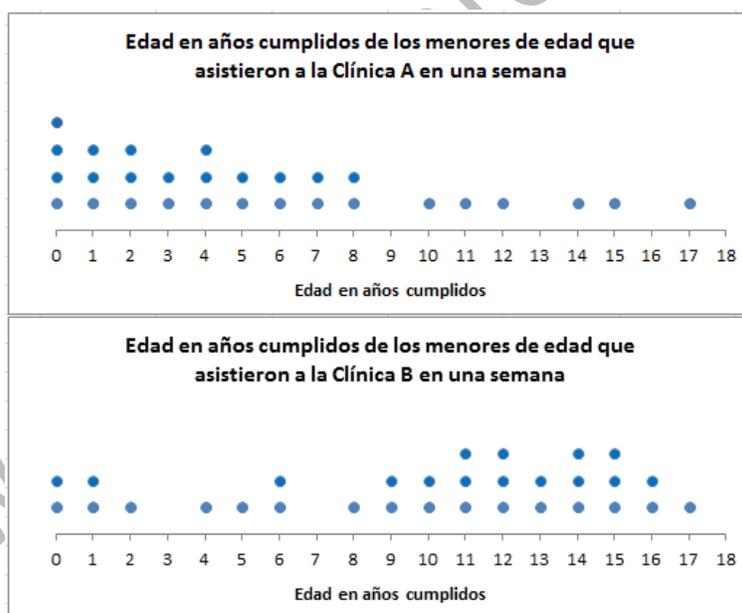


Imagen propia elaborada con fines didácticos

Ayude al asistente médico a tomar una decisión respecto a la clínica en la que debería realizar el estudio.

Solución:

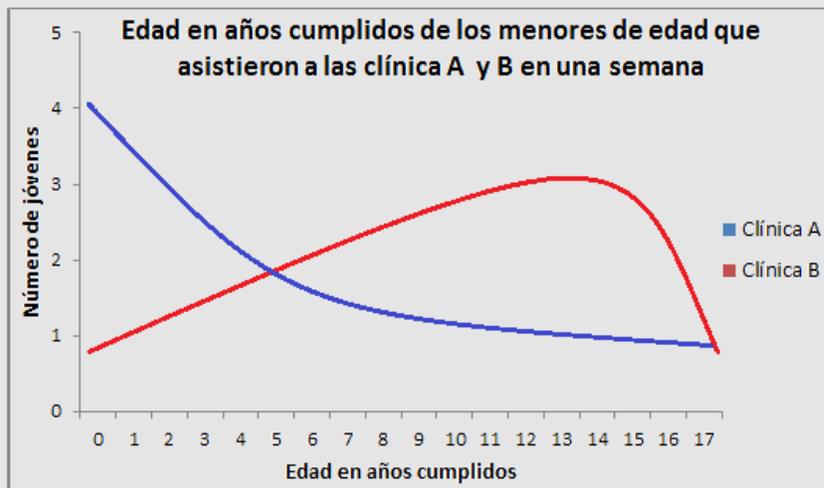
Esta representación se conoce como *diagrama de puntos*: los puntos representan la edad de los jóvenes y el número de puntos en cada edad (frecuencia) representa el número de jóvenes que tienen dicha edad.

De acuerdo con la información recabada en esa semana, el número de adolescentes fue mucho mayor en la clínica B; por su parte, en la clínica A se tuvo una mayor presencia de niños menores de ocho años. Por lo tanto, se debe indicar al asistente médico que, si lo acontecido en esa semana es un reflejo de lo que ocurre regularmente en las clínicas, el estudio debería realizarlo en la clínica B.

Nota: En vez de un diagrama de puntos, la información se pudo haber resumido mediante gráficas de barras, tal como se indica a continuación.



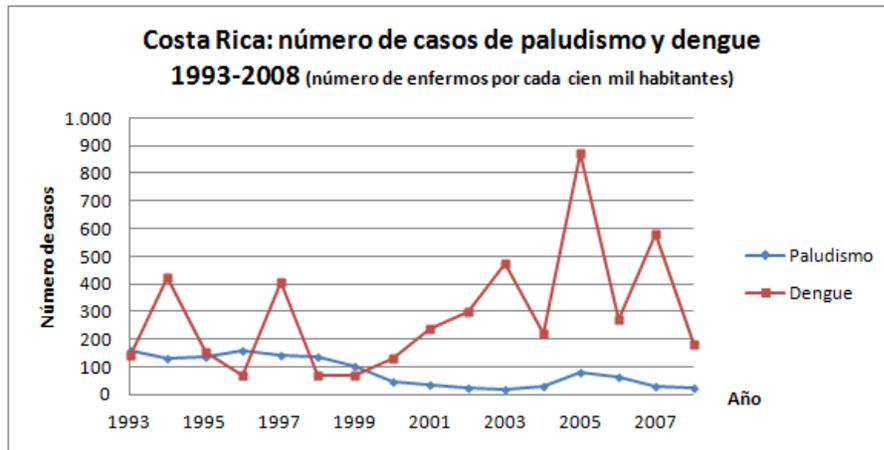
Debido a que el número de jóvenes menores de 18 analizados en las dos clínicas es diferente (29 en la clínica A y 31 en la B), entonces no se deben comparar las barras individuales, para responder el problema debe observarse el patrón de la variabilidad en los datos.



La visualización de estos patrones permite identificar que en esa semana se presentaron más adolescentes en la clínica B que en la clínica A.

Problema 7. Enfermos por picadura de mosquito

Existen diversas enfermedades que se adquieren por la picadura de mosquitos, dos de las más conocidas son el paludismo y el dengue. La siguiente gráfica muestra el número de casos de estas enfermedades que se presentó en el país entre 1993 y 2008.



Fuente: <http://www.estadonacion.or.cr/index.php/estadisticas/costa-rica/compendio-estadistico/estadisticas-sociales>

Considere las siguientes proposiciones que se relacionan con la lectura realizada al gráfico:

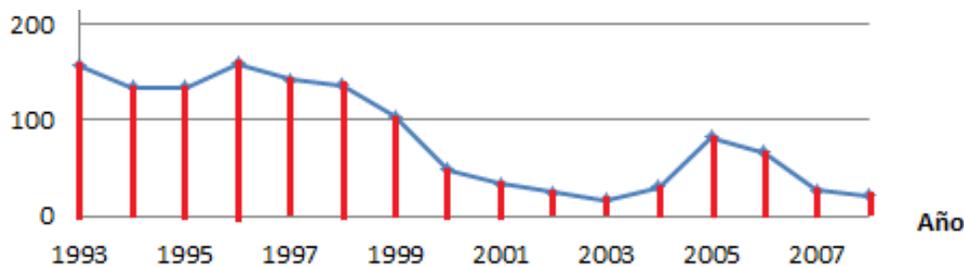
- I. Para el período 1993-2008, el número promedio anual de enfermos por dengue es superior al número promedio anual de enfermos por paludismo.
- II. Para el período 1993-2008, el número promedio anual de enfermos por paludismo fue menor de 100 casos por cada cien mil habitantes.

Determine si las proposiciones anteriores son verdaderas o falsas.

Solución:

Hay que destacar que la representación gráfica anterior muestra un índice de salud que corresponde al número de enfermos por cada cien mil habitantes para un período de 18 años (de 1993 hasta el 2008). Según esta representación en la mayoría de los años analizados el número de enfermos por dengue superó al de enfermos por paludismo (esto último se puede observar ya que hay una mayor cantidad de años en el período en donde la gráfica del dengue está por encima que la gráfica del paludismo), estas diferencias fueron más grandes después de 1999. Entonces para este período del estudio, al sumar el número de enfermos por dengue se obtiene un valor mayor que la suma del número de enfermos por paludismo. Por ello es claro que el número promedio anual de enfermos por dengue superó al número promedio anual de enfermos por paludismo.

Por otro lado, si observamos solamente el comportamiento de la curva correspondiente a los casos de paludismo, vemos que entre 1993 y 1999 se presentaron más de cien casos por cada cien mil habitantes, pero después de ese año el número de casos fue menor a cien como se muestra en la siguiente imagen.



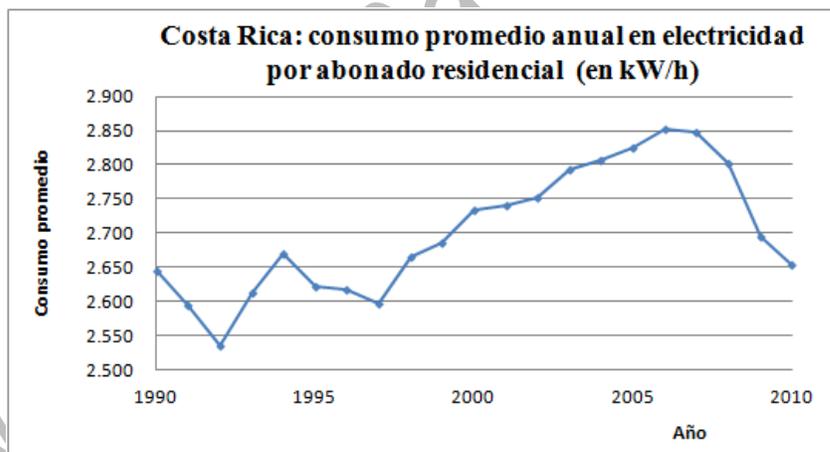
Bajo el supuesto que el gráfico está construido con una escala adecuada, al extraer el promedio de las magnitudes de los segmentos que se representan con rojo que corresponden al número anual de casos de paludismo por cada cien mil habitantes, se puede observar visualmente que este promedio debería ser menor de 100.

Por lo anterior se puede decir que las dos proposiciones son verdaderas.

[Volver a Contenido](#)

Problema 8. Consumo promedio anual de electricidad por abonado residencial

Observe la siguiente representación gráfica.



Fuente: <http://www.estadonacion.or.cr/index.php/estadisticas/costa-rica/compendio-estadistico/estad-ambientales>

Analice las siguientes afirmaciones vinculadas con la información que presenta el gráfico anterior:

- I. El consumo total de electricidad entre los abonados residenciales fue aproximadamente igual en 1990 y en el 2010.
- II. El máximo consumo de electricidad presentado en una vivienda fue de 2850 kW/h.

Determine cuál o cuáles de estas afirmaciones se puede asegurar con certeza que sean verdaderas:

Solución:

Es evidente que en 1990 había mucho menos abonados residenciales que en el 2010. Por esta razón, no es suficiente saber que el consumo promedio anual de electricidad (en kW/h) fue el mismo para 1990 y 2010 (como lo presenta la gráfica) para concluir que el consumo total también haya sido aproximadamente igual en esos años, puesto que ello depende del número total de abonados residenciales en cada uno de dichos años. Hay que recordar que el promedio de cada año se determina por la fórmula:

$$\text{Promedio} = \frac{\text{Consumo total en electricidad de los abonados residenciales}}{\text{Total abonados residenciales}}$$

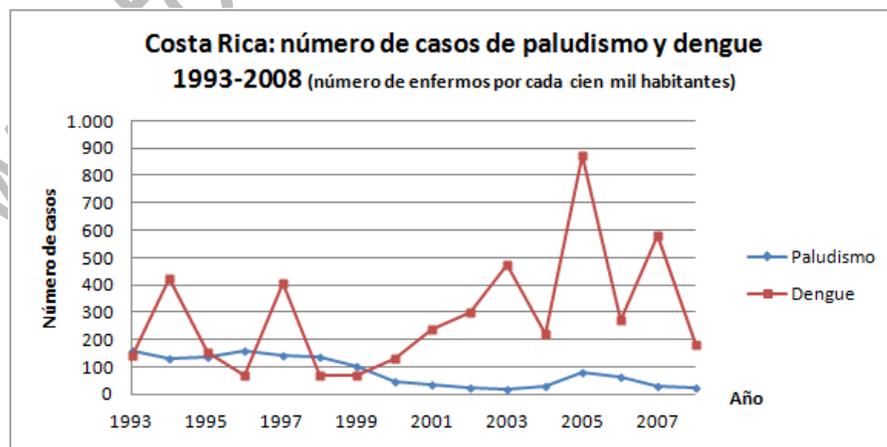
Entonces si los consumos promedio por abonado residencial de 1990 y 2010 son iguales y el número total de abonados residenciales son diferentes, también el consumo total de electricidad de los abonados residenciales en ambos años son diferentes.

Por otro lado, la gráfica ilustra el consumo promedio anual de electricidad y ello no permite asegurar el valor del máximo consumo de electricidad en una vivienda particular, basta con analizar la fórmula de cálculo del promedio, para identificar que con la información del gráfico no es posible determinar el consumo total de electricidad por vivienda ni tampoco el acumulado para todas las viviendas.

Por lo anterior, ambas proposiciones son falsas.

Problema 9. Enfermos por picadura de mosquito

Volvamos al problema 7, en donde se analiza dos enfermedades que se vinculan con la picadura de mosquitos: el paludismo y el dengue. La siguiente gráfica muestra el número de casos de estas enfermedades en el país entre 1993 y 2008.



Fuente: <http://www.estadonacion.or.cr/index.php/estadisticas/costa-rica/compendio-estadistico/estadisticas-sociales>

De acuerdo con esta información, ¿en cuál de las variables se presenta mayor variabilidad: en el número anual de enfermos por paludismo, o en el número anual de enfermos por dengue?

Solución:

A simple vista puede observarse que el número de enfermos por paludismo fue menos variable en este período e incluso presentó una importante disminución. Por su parte, el número de enfermos por dengue fue muy inestable con grandes variaciones entre años y con una tendencia al aumento. En general se puede concluir que el número anual de enfermos por dengue fue mucho más variable durante el periodo 1993-2008.

Problema 10. Años de experiencia en labores docentes

Considere la información del cuadro y las proposiciones dadas.

En el cuadro siguiente se muestran los años de experiencia que tienen 196 docentes de educación primaria que fueron encuestados en el año 2009.

Distribución de los docentes encuestados según número de años de experiencia en labores de docencia. Año 2009

Años de experiencia	Cantidad de docentes	Porcentaje de docentes
De 1 a menos de 5	28	14,3
De 5 a menos de 10	31	15,8
De 10 a menos de 15	37	18,9
De 15 a menos de 20	40	20,4
De 20 a menos de 25	48	24,5
De 25 a menos de 30	12	6,1
Total	196	100

Con base en los datos del cuadro anterior analice las siguientes proposiciones:

- I. La mayoría (más del 50%) de los docentes encuestados tenía menos de 15 años de experiencia en labores docentes.
- II. El valor de la mediana de la experiencia de los docentes encuestados se encuentra entre 15 y 20 años.

Determine cuál o cuáles de las proposiciones anteriores son verdaderas.

Solución:

Los datos sobre los años de experiencia de los docentes fueron resumidos en el cuadro que se presenta, por lo que no se sabe cuáles fueron los datos individuales. Sin embargo, al observar el ordenamiento de los datos y las frecuencias absolutas y porcentuales, si vamos sumando las frecuencias porcentuales tal como se muestra en la representación de abajo, se puede observar que un 49% de los docentes encuestados tenía menos de 15 años de experiencia, que es menor al 50%, entonces no corresponde a la mayoría de los encuestados, por lo que la primera proposición es falsa.

Por otro lado, de acuerdo con la definición de la mediana, esta medida acumula el 50% de los datos, si analizamos los porcentajes acumulados, puede notarse que hasta la clase de 15 a 20 años de experiencia, se ha acumulado este 50%, lo cual significa que la mediana en los años de experiencia está en esta clase, entonces la segunda proposición es verdadera.

Años de experiencia	Cantidad de docentes	Porcentaje de docentes	Porcentaje acumulado
De 1 a menos de 5	28	14,3	14,3
De 5 a menos de 10	31	15,8	30,1
De 10 a menos de 15	37	18,9	49,0
De 15 a menos de 20	40	20,4	69,4
De 20 a menos de 25	48	24,5	93,9
De 25 a menos de 30	12	6,1	100,0
Total	196	100	

3. Representación de la distribución de datos mediante un diagrama de cajas

[Volver a Contenido](#)

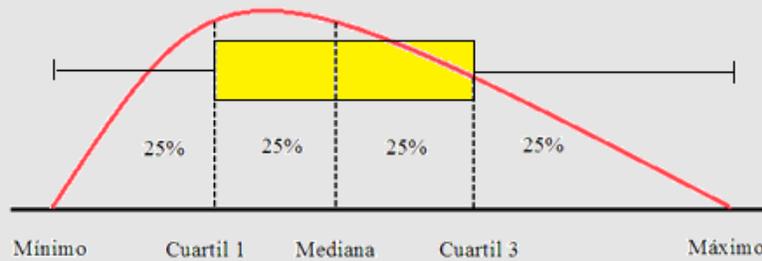
Diagrama de cajas

Como se mencionó anteriormente, las medidas estadísticas de orden tienen una interpretación gráfica que favorece una visualización práctica sobre el patrón de la distribución de los datos. En este sentido los diagramas de cajas se convierten en una importante herramienta estadística para llevar a cabo esta visualización. Un diagrama de cajas, también llamado gráfico de bigotes, consiste en una representación estadística en la cual se incluyen las principales medidas de orden que hemos analizado previamente. Cada caja debe estar en escala según el valor de la medida estadística en el eje x. Entre el primer cuartil y el tercer cuartil se dibuja una caja tal como se muestra. Esta caja se une con el mínimo y el máximo por medio de segmentos. Finalmente la mediana se representa con una línea que divide la caja.



De acuerdo con lo anterior, un diagrama de cajas proporciona información de estas cinco medidas de orden, pero además, son un indicador de la forma aproximada que tiene la distribución de datos. Sin embargo, no se puede perder de vista que estos diagramas son simplemente un resumen de los datos, por medio de cinco medidas. Es importante mencionar que hay que ser muy cuidadosos al interpretar la información que comunican.

Nota: Observe que por medio de diagrama de cajas se puede visualizar en forma aproximada la distribución general de los datos, tal como se hace con un polígono de frecuencias.



Observe que la distribución de los datos representada por la curva muestra una ligera asimetría positiva, esto se evidencia también en diagrama de cajas, debido a que la caja está sesgada hacia la izquierda.

Seguidamente el problema 5, será analizado de acuerdo con el comportamiento de un diagrama de cajas.

[Volver a Contenido](#)

Problema 11. Vida útil de los bombillos

Suponga que le envían a comprar unos bombillos a una ferretería, la persona que le atiende indica que de acuerdo con las características del bombillo que usted está solicitando tiene dos marcas con el mismo precio: A y B. Le pregunta que cuál de ellas desea. Usted recuerda que en un artículo de una revista que leyó recientemente salió publicado la siguiente representación gráfica sobre la vida útil de estos bombillos:

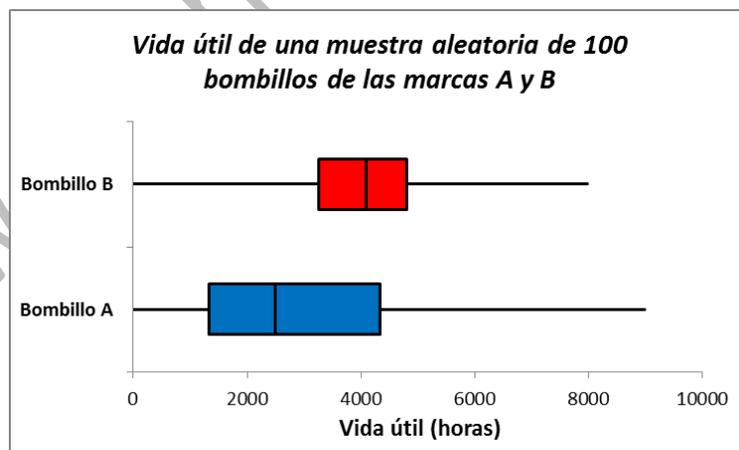
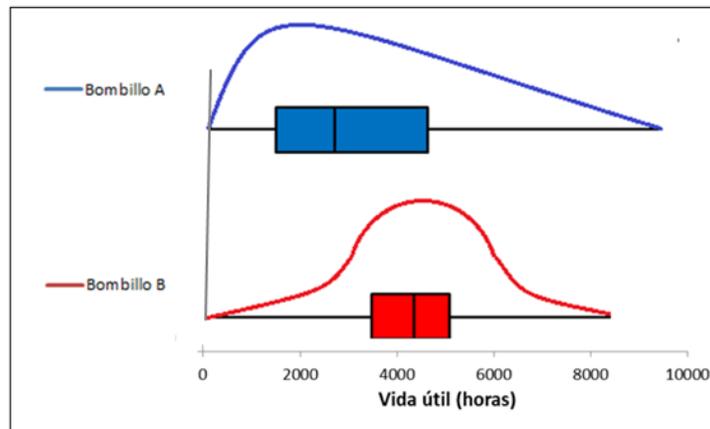


Imagen propia elaborada con fines didácticos

Utilice esta información para decidir qué marca de bombillos compraría y argumente esta decisión.

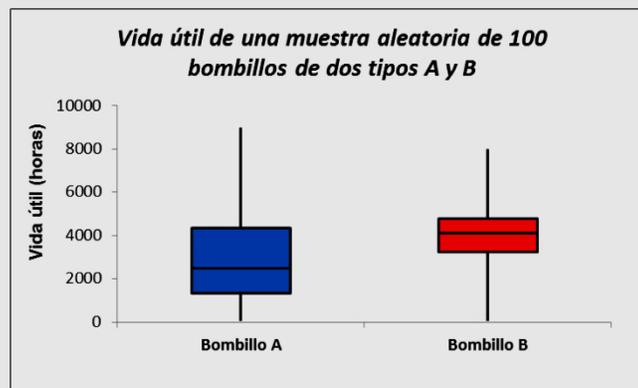
Solución

Note que en la caja correspondiente, a cada tipo de bombillo se puede visualizar el tipo de asimetría que presenta la distribución de los datos de acuerdo con la representación gráfica que apareció en el problema 3.



La información que proporciona el diagrama de cajas es evidencia suficiente para decidir adquirir los bombillos B debido a que existe mayor posibilidad de que tenga una más duración.

Nota: Los diagramas de cajas pueden expresarse en forma vertical tal como se muestra a continuación:

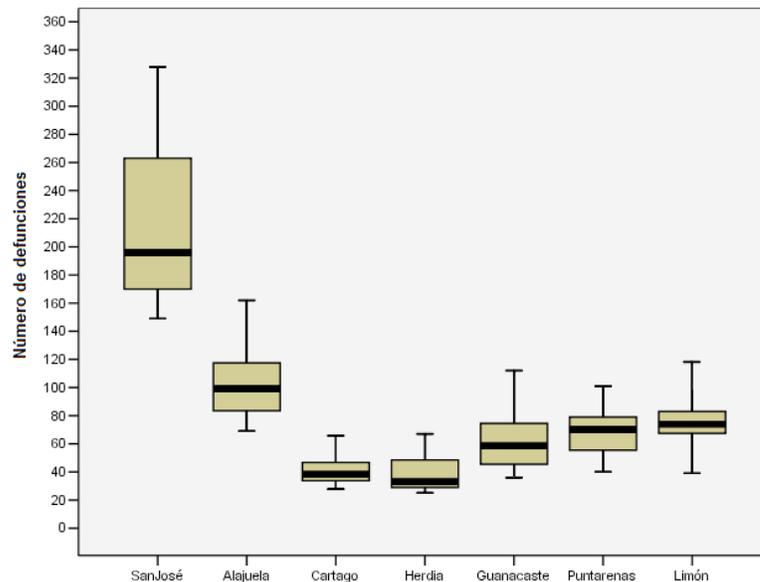


[Volver a Contenido](#)

Problema 12. Defunciones por accidentes de tránsito según provincia

Considere el siguiente diagrama de cajas.

Costa Rica: Número anual de defunciones por accidentes de tránsito
Período 1991-2010



Fuente: <http://www.estadonacion.or.cr/index.php/estadisticas/costa-rica/compendio-estadistico/estadisticas-sociales>

De la información proporcionada por el diagrama, ¿cuál de las siguientes proposiciones son verdaderas:

- I. Limón y Cartago son las provincias en las que se presentó la menor cantidad de muertes por años debidas a accidentes de tránsito en este período.
- II. Para San José, en más de la mitad de los años del período 1991-2010 se presentó más de 220 defunciones anuales por accidentes de tránsito.
- III. Heredia fue la provincia en que se presentó la menor cantidad de defunciones por accidentes de tránsito por año en este período.

Solución

Al observar las cajas que representan al número anual de defunciones por accidentes de tránsito en nuestro país en el período 1991-2010, se identifica que es en Heredia donde se presentó la menor cantidad de defunciones por accidentes de tránsito en dicho período: tiene menor mediana que todas las demás provincias y el tercer cuartil y valor máximo son similares a Cartago que sería la provincia que le sigue en este análisis de valores absolutos.

En este sentido se tiene que la proposición I sería falsa, debido a que las provincias con menor cantidad de defunciones son Heredia y Cartago. En San José la mediana es menor de 200 defunciones anuales, por ello no es cierto que “*en más de la mitad de los años del período 1991-2010, presentó más de 200 defunciones anuales por accidentes de tránsito*”, debido a

que el porcentaje de años en los que hubo más de 220 defunciones es mucho menor que la mitad.

Por lo tanto la proposición III es la única correcta.

[Volver a Contenido](#)

www.reformamatematica.net

II. Bibliografía

Ministerio de Educación Pública (2015). Curso bimodal para el II Ciclo: Estadística mucho más que procedimientos y técnicas. Unidad didáctica Estadística. San José, Costa Rica: autor.

Ministerio de Educación Pública (2012). Programas de Matemáticas para la Educación General Básica y el Ciclo Diversificado. San José, Costa Rica: autor.

III. Créditos

Material complementario Estadística y Probabilidad: Tablas y gráficos estadísticos, es parte del MiniMOOC Medidas estadísticas, una actividad del Proyecto Reforma de la Educación Matemática en Costa Rica.

Este proyecto del Ministerio de Educación Pública es apoyado por Asociación Empresarial para el Desarrollo y por la Fundación Costa Rica - Estados Unidos de América para la Cooperación.

Autor del presente documento

Edwin Chaves Esquivel

Revisores de este documento

Ángel Ruiz, Edison De Faria, Johanna Mena, Keibel Ramírez, Luis Hernández, Xinia Zúñiga.

Director general del proyecto Reforma de la Educación Matemática en Costa Rica

Ángel Ruiz

Para referenciar este documento

Ministerio de Educación Pública, Proyecto Reforma de la Educación Matemática en Costa Rica (2017). *Material complementario Estadística y Probabilidad: Tablas y gráficos estadísticos*, San José, Costa Rica: autor.



Material complementario Estadística y Probabilidad: Tablas y gráficos estadísticos por Ministerio de Educación Pública de Costa Rica, Proyecto Reforma de la Educación Matemática en Costa Rica, se encuentra bajo una Licencia Creative Commons Atribución-CompartirIgual 3.0 Unported.

[Volver a Contenido](#)