

Bienvenidos a RTPSTAT- Módulo Descriptiva!

"RTPSTAT - Módulo Descriptiva" es una interfaz para la resolución automatizada y tutorizada de problemas de Estadística Descriptiva, que integra un editor de problemas y un generador de presentaciones PDF para su resolución. Es la primera piedra del "Proyecto RTPSTAT" para la resolución de problemas de Estadística.

"RTPSTAT - Módulo Descriptiva" ha sido desarrollado dentro de la acción de innovación docente de la Universidad de Granada "Resolución automatizada y tutorizada de problemas de Estadística. Aplicaciones a la Estadística Descriptiva", coordinado por el profesor Juan José Serrano Pérez del Departamento de Estadística e Investigación Operativa.

Ésta es la versión 1.0.0 de la interfaz, que está en fase de depuración y en la que algunas opciones de los menús aún no están disponibles. Si usted detecta errores le agradeceríamos nos lo indique a la dirección de correo jjserra@ugr.es

Dada la especial naturaleza de los problemas a resolver, nuestro generador requiere disponer en nuestra plataforma de Latex, con la clase Beamer y PdfLatex, y del entorno de programación estadística R.

Latex es un editor de documentos científicos que aporta numerosas ventajas frente a los tradicionales editores, como Microsoft Word, sobre todo si los documentos contienen gran cantidad de fórmulas. Los documentos en Latex no son más que un fichero "sólo texto" con extensión "tex" que contiene el texto propiamente dicho que queremos escribir junto con comandos que explican a un intérprete de Latex cómo debe tratar cada parte del texto.

Latex es más que un procesador de textos, en realidad es un potente lenguaje de programación para la edición de textos de gran calidad. Además, fue creado como software libre, y crece cada vez más con paquetes de comandos que expertos o aficionados desarrollan con fines variadísimos. Y cada año se ofrece una nueva versión gratuita a través de Internet.

Una vez compilado el fichero por Latex, éste genera un fichero de salida con extensión dvi, que significa "device independent", es decir, independiente de la plataforma y sistema; y que puede ser exportado fácilmente a otros formatos, como por ejemplo

al universalmente difundido PDF (Portable document format). Además, Latex es capaz de generar, mediante PDFLatex, presentaciones en formato PDF similares a las tradicionales presentaciones de PowerPoint, pero con claras ventajas, a saber, ocupan menos, son independientes del dispositivo de salida y resolución, y los visores PDF, tales como Acrobat Reader, se distribuyen libremente.

Aunque Latex dispone de múltiples paquetes (pdfscreen, pdfslide, etc.) para generar presentaciones PDF, nuestro generador trabaja con la clase Beamer. Las características más notables de la clase Beamer son:

- Los comandos estándar de Latex siguen funcionando.
- Es fácil crear overlays y efectos dinámicos.
- Podemos elegir uno de los diseños disponibles y la apariencia de la presentación queda completamente definida.
- La composición, los colores y las fuentes empleadas en la presentación pueden cambiarse fácilmente de modo global y en el más mínimo detalle.
- Es fácil crear anotaciones fuera de las presentaciones.

R es un lenguaje estadístico en modo comando muy potente para el análisis de datos y gráficos. Se trata de un lenguaje de alto nivel, aunque con una sintaxis muy simple, que trabaja directamente con objetos (matrices, vectores, gráficos, etc.) e integra una colección muy amplia de procedimientos matemáticos y estadísticos. Permite programar con un mínimo esfuerzo, de modo que una única línea de programación en R puede equivaler a decenas de líneas de programación con otros lenguajes. Los programas en R son funciones que se editan como un archivo de texto con órdenes del lenguaje que pueden editarse y ejecutarse llamándolas dentro del propio entorno, o fuera de él. Además, R es también software libre diseñado por expertos, que puede bajarse de Internet.

"RTPSTAT - Módulo Descriptiva" se distribuye en un cd que incluye un directorio de ejemplos de problemas, y las distribuciones libres de MikTex 2.4.1705 y de R 2.5.0. Usted puede obtener más información sobre estos programas y descargar versiones más actualizadas en las siguientes direcciones:

- <http://www.r-project.org>
- <http://www.miktex.org>

"RTPSTAT - Módulo Descriptiva" permite editar un problema y generar de forma automática un archivo Latex con su resolución convenientemente editada con ayuda de la clase Beamer, y lo compila también de forma automática con PDFLatex produciendo una presentación animada en formato PDF.

Las presentaciones generadas mediante RTPSTAT presentan dos ventajas fundamentales respecto a las tradicionales:

- Son dinámicas.
En efecto, pueden actualizarse sobre la marcha de modo que muestren la resolución de nuevas cuestiones sobre un mismo problema o la resolución derivada de cambios en los datos o en el enunciado del problema.
- Aprovechan la potencia de cálculo y gráfica del ordenador.
No se trata de hacer uso del ordenador como proyector, sino de aprovechar su capacidad para generar las tablas de cálculos y gráficos estadísticos necesarios para la presentación.

Todo el tratamiento estadístico se realiza tal y como se estudia teóricamente, y difiere en algunos casos del tratamiento computacional que se hace con el software estadístico habitualmente utilizado, y, a diferencia de éste, no muestra sólo los resultados finales, sino todas las etapas del método.

La interfaz implementada simplifica extraordinariamente la labor docente, y aumenta su efectividad:

- Los alumnos podrán satisfacer por completo su demanda individual de problemas resueltos de acuerdo con su capacitación y habilidades personales. Además cuando resuelvan problemas propuestos, podrán comprobar si sus soluciones son correctas o no, y si no lo son, determinar si se trata de un error de cálculo o de estrategia estadística. De esta manera aprenderán a estructurar adecuadamente los métodos estadísticos, así como los cálculos necesarios para su aplicación.
- El profesor podrá presentar cuantos problemas desee, pudiendo centrarse más en una selección adecuada de los mismos, dado que su resolución pormenorizada le viene dada automáticamente. Además, podrá valorar mejor, y sin esfuerzo, el tiempo y la dificultad que implica la resolución completa de un problema antes de proponerlo a los alumnos.

- Dado que las presentaciones generadas muestran cómo se ha aplicado el método estadístico, los cálculos necesarios, e incluyen referencias teóricas, también constituyen una herramienta muy destacable en innovación tutorial automatizada.
- El atractivo de las presentaciones, junto con su valor didáctico, estimula el trabajo individual de los alumnos y facilita la asimilación de los procedimientos estadísticos.

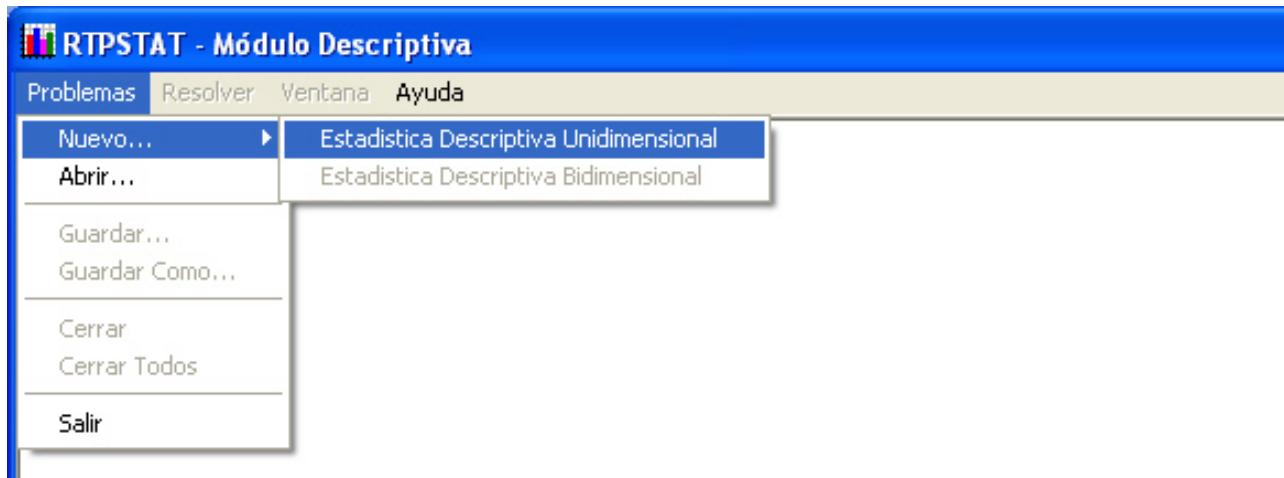
Al ejecutar "RTPSTAT - Módulo Descriptiva", éste analiza si están disponibles los recursos que necesita (R y Latex con Beamer y PDFLatex), y si no es así, nos indica que el generador de presentaciones no estará disponible, y únicamente podremos crear y editar problemas, pero no resolverlos.



El menú principal de la aplicación presenta las opciones:

- **Problemas**, para crear, editar y guardar problemas.

- **Resolver**, disponible sólo si el problema activo no presenta errores y tenemos instalados en nuestra plataforma los recursos necesarios.
- **Ventana**, para organizar las ventanas de los distintos problemas en edición.
- **Ayuda**, que da acceso a este archivo y a la información sobre la autoría de la aplicación.



La interfaz permite trabajar "simultáneamente" con varios problemas. Un problema de Estadística Descriptiva Unidimensional queda perfectamente definido cuando especificamos:

- Una descripción general del **asunto** tratado en el problema (**opcional**).
- El **planteamiento** del problema.
- La **variable en estudio** (su descripción es opcional) y su **tipo**.
- Los **datos observados** y la **organización** de los mismos.
- El **enunciado** de cada **apartado**, si tiene o no subapartados, y, si no tiene, los **cálculos** necesarios para resolverlo.
- El **enunciado** de cada **subapartado** y los **cálculos** necesarios para su resolución.

C:\RTPSTAT\Descriptiva1D\Ejemplo Continua En Serie.rtp

Asunto: Ejemplo de funcionamiento para datos en serie de una variable continua

Planteamiento del Problema:

Los siguientes datos son observaciones de la resistencia al corte, X, en libras, de soldaduras de punto ultrasónicas hechas en un cierto tipo de láminas de duraluminio:

Datos

Variable en estudio (X): Resistencia al corte (en libras) de una soldadura de punto ultrasónica

Tipo: Discreta Continua Organización: En serie En una tabla de frecuencias Editar

Apartados

Apartado: 1) Determina la distribución de frecuencias de la resistencia al corte de las soldaduras de punto ultrasónicas y promédiala en magnitud con diferentes

Subapartados: Sí No Añadir Eliminar

Cálculos Tab(X), m1(X), H(X), Q(X), G(X) Ordenar

Subapartados

Subapartado: Nuevo (Enunciado) Añadir Eliminar Ordenar

Cálculos (Lista de cálculos)

Es posible **añadir**, **eliminar** y **ordenar apartados y subapartados** haciendo uso de los correspondientes botones, que sólo estarán activos cuando estas acciones tengan sentido.

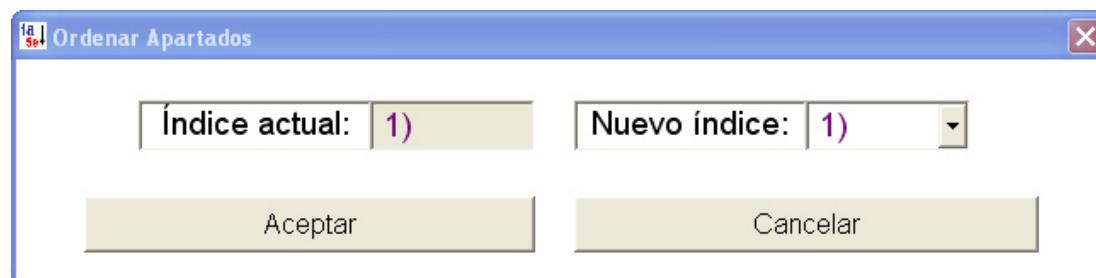
Para añadir un apartado (o un subapartado) debemos seleccionar Nuevo, en la lista desplegable de apartados (o subapartados).

Si un apartado no tiene subapartados hay que definir antes su enunciado y los cálculos a realizar; y, si tiene subapartados, hay que definir antes su enunciado y añadir al menos un subapartado. No obstante, el apartado no se añadirá hasta hacer clic en el botón Añadir.

Para añadir un subapartado basta especificar simplemente su enunciado y los cálculos necesarios, y hacer clic en el botón Añadir.

El número de apartados y subapartados por apartado se ha limitado a 5 con el fin de que el menú de navegación del diseño empleado en la presentación PDF generada deje suficiente espacio libre de pantalla para presentar los cálculos, gráficos y explicaciones necesarios para resolverlos.

Podemos reordenar apartados y subapartados haciendo clic en los botones Ordenar correspondientes.



Para **editar los datos** de la variable en estudio hemos diseñado un editor de datos accesible a través del botón Editar, que adopta distinto formato en función del tipo de variable y de la organización de los datos.

Distribución de frecuencias

Opciones Amplitud

Datos		Intervalos de Clase	
	Valores		
1	5434.0000		
2	4948.0000		
3	4521.0000		
4	4570.0000		
5	4990.0000		
6	5702.0000		
7	5241.0000		
8	5112.0000		
9	5015.0000		

Fila 1 —
Valor: 5434

Distribución de frecuencias

Opciones Amplitud

Datos		Intervalos de Clase	
	Extremos inferiores de clase	Extremos superiores de clase	
1	4000.0000	4200.0000	
2	4200.0000	4400.0000	
3	4400.0000	4600.0000	
4	4600.0000	4800.0000	
5	4800.0000	5000.0000	
6	5000.0000	5200.0000	
7	5200.0000	5400.0000	
8	5400.0000	5600.0000	

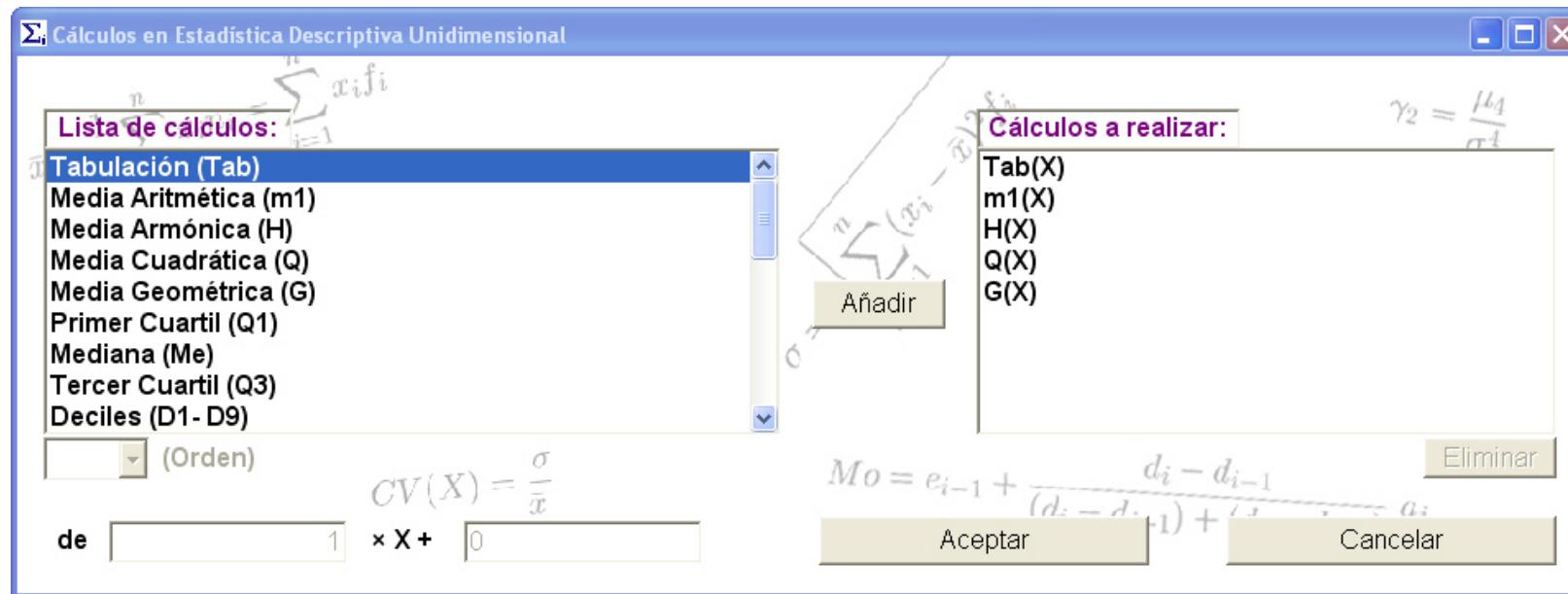
Primer extremo de clase: 4000
Amplitud de clase: 200

Podemos especificar datos organizados en una tabla de frecuencias, tanto para variables discretas como continuas; y datos en serie, junto con los intervalos de agrupación en el caso de una variable continua.

La edición es inteligente, es decir, detecta errores y/o incoherencias en los datos.

Para indicar los **cálculos a realizar** en cada apartado o subapartado disponemos del botón Cálculos, que nos permite seleccionarlos de una lista de cálculos con la mayoría de medidas de resumen de una variable estadística unidimensional, a saber, medidas de posición (promedios y cuantiles), medidas de dispersión (absolutas y relativas) y medidas de forma (asimetría y curtosis). Además, para datos en serie, hemos añadido también su tabulación como otro cálculo más.

La lista de cálculos a realizar puede reordenarse, y también podemos eliminar cálculos. Los formularios para especificar los cálculos en apartados y subapartados son similares.



La opción Resolver permite **crear una presentación PDF** con la resolución pormenorizada del problema en edición activo o ver una presentación generada en otro momento. Y podemos especificar la precisión que emplearemos en los cálculos, lo que posibilita evaluar la sensibilidad de la solución en este sentido.

RTPSTAT - Módulo Descriptiva

Problemas Resolver Ventana Ayuda

C:\RTT Asun Serie.rtp

Aumentar
Disminuir
4

namiento para datos en serie de una variable continua

Planteamiento del Problema:

Los siguientes datos son observaciones de la resistencia al corte, X, en libras, de soldaduras de punto ultrasónicas hechas en un cierto tipo de láminas de duraluminio:

Datos

Variable en estudio (X): Resistencia al corte (en libras) de una soldadura de punto ultrasónica

Tipo: Discreta Continua Organización: En serie En una tabla de frecuencias Editar

Apartados

Apartado: 1) Determina la distribución de frecuencias de la resistencia al corte de las soldaduras de punto ultrasónicas y promédiala en magnitud con diferentes

Subapartados: Sí No

Cálculos Tab(X), m1(X), H(X), Q(X), G(X)

Subapartados

Subapartado: Nuevo (Enunciado)

Cálculos

(Lista de cálculos)

Si hemos realizado cambios en el problema y éstos no han sido guardados, la presentación PDF generada se presentará como un archivo temporal Temp.pdf

Para generar una presentación PDF basta hacer clic en la opción Crear Presentación y esperar un poco mientras se genera automáticamente y se abre. El tiempo de espera depende obviamente de la cantidad de cálculos a realizar y de la potencia del ordenador, pero para que nos hagamos una idea podemos indicar que en un Pentium(R) 4, con CPU a 2.53 GHz y 512 MB de RAM, que es un ordenador bastante limitado hoy en día, podemos generar una presentación con 126 diapositivas en apenas 1 minuto.

La presentación generada es inteligente en el sentido de que no repite cálculos, es decir, si un nuevo cálculo requiere algo previamente calculado, se hace referencia a dicho cálculo mediante un hipervínculo, pero no se repite el cálculo. Y los cálculos están perfectamente estructurados, de modo que si el cálculo de una medida requiere el cálculo de otras, se indican primero todos los cálculos necesarios y luego éstos se presentan uno a uno enlazados con hipervínculos.

Cada diapositiva de la presentación PDF generada presenta:

- Un **menú de navegación**, en la parte superior, en el que se destaca el apartado o subapartado en el que nos encontramos.
- Una **barra de información**, debajo del menú de navegación, que muestra información de interés sobre lo que se muestra más abajo (enunciados, datos o cálculos), o hipervínculos a los cálculos realizados en cada apartado o subapartado, según el caso.
- El **enunciado** o cada uno de los **cálculos** de un apartado o subapartado, según el caso. Los cálculos incluyen tablas, gráficos y fórmulas, con explicaciones paso a paso de cada cálculo.
- Una serie de **botones de navegación**, en la esquina inferior derecha, para recorrer las diapositivas.
- Una **barra de estado**, que informa del día y hora en que fue generada la presentación, y una dirección de contacto.

Planteamiento

- Apartado 1
- Apartado 2
- Apartado 3
- Apartado 4
- Apartado 5

Enunciado

- Datos
- Tabulación

Ejemplo de funcionamiento para datos en serie de una variable continua

Los siguientes datos son observaciones de la resistencia al corte, X, en libras, de soldaduras de punto ultrasónicas hechas en un cierto tipo de láminas de duraluminio:

Planteamiento

Apartado 1
Apartado 2
Apartado 3
Apartado 4
Apartado 5

Enunciado
Datos
Tabulación

X ≡ “Resistencia al corte (en libras) de una soldadura de punto ultrasónica”

5434, 4948, 4521, 4570, 4990, 5702, 5241, 5112, 5015, 4659, 4806, 4637, 5670, 4381, 4820, 5043, 4886, 4599, 5288, 5299, 4848, 5378, 5260, 5055, 5828, 5218, 4859, 4780, 5027, 5008, 4609, 4772, 5133, 5095, 4618, 4848, 5089, 5518, 5333, 5164, 5342, 5069, 4755, 4925, 5001, 4803, 4951, 5679, 5256, 5207, 5621, 4918, 5138, 4786, 4500, 5461, 5049, 4974, 4592, 4173, 5296, 4965, 5170, 4740, 5173, 4568, 5653, 5078, 4900, 4968, 5248, 5245, 4723, 5275, 5419, 5205, 4452, 5227, 5555, 5388, 5498, 4681, 5076, 4774, 4931, 4493, 5309, 5582, 4308, 4823, 4417, 5364, 5640, 5069, 5188, 5764, 5273, 5042, 5189, 4986



Planteamiento

Apartado 1
Apartado 2
Apartado 3
Apartado 4
Apartado 5

Enunciado
Datos
Tabulación

Tabular los datos es representar su distribución de frecuencias en una tabla

5434, 4948, 4521, 4570, 4990, 5702, 5241, 5112, 5015, 4659, 4806, 4637,
5670, 4381, 4820, 5043, 4886, 4599, 5288, 5299, 4848, 5378, 5260, 5055,
5828, 5218, 4859, 4780, 5027, 5008, 4609, 4772, 5133, 5095, 4618, 4848,
5089, 5518, 5333, 5164, 5342, 5069, 4755, 4925, 5001, 4803, 4951, 5679,
5256, 5207, 5621, 4918, 5138, 4786, 4500, 5461, 5049, 4974, 4592, 4173,
5296, 4965, 5170, 4740, 5173, 4568, 5653, 5078, 4900, 4968, 5248, 5245,
4723, 5275, 5419, 5205, 4452, 5227, 5555, 5388, 5498, 4681, 5076, 4774,
4931, 4493, 5309, 5582, 4308, 4823, 4417, 5364, 5640, 5069, 5188, 5764,
5273, 5042, 5189, 4986

$I_i = (e_{i-1}, e_i]$	n_i
(4000, 4200]	1
(4200, 4400]	2
(4400, 4600]	9
(4600, 4800]	12
(4800, 5000]	19
(5000, 5200]	22
(5200, 5400]	20
(5400, 5600]	7
(5600, 5800]	7
(5800, 6000]	1

Calcula las frecuencias absolutas (n_i) de los intervalos de clase contando el número de valores que pertenecen a cada intervalo. Por ejemplo, en el intervalo $I_6 = (5000, 5200]$ se encuentran $n_6 = 22$ valores

Planteamiento

Apartado 1
Apartado 2
Apartado 3
Apartado 4
Apartado 5

SubApartado 1
SubApartado 2

► P30(X) ► P56(X)

Enunciado

La resistencia al corte máxima del 30 % de las soldaduras de punto ultrasónicas que presentan menor resistencia al corte, y la resistencia al corte mínima del 44 % de las soldaduras de punto ultrasónicas que presentan mayor resistencia al corte

P_{56} es un valor tal que el 56 % de los individuos presentan valores menores o iguales a él, y el 44 % de los individuos presentan valores mayores o iguales a él

► P30(X) P56(X)

$I_i = (e_{i-1}, e_i]$	n_i	N_i
(4000, 4200]	1	1
(4200, 4400]	2	3
(4400, 4600]	9	12
(4600, 4800]	12	24
(4800, 5000]	19	43
(5000, 5200]	22	65
(5200, 5400]	20	85
(5400, 5600]	7	92
(5600, 5800]	7	99
(5800, 6000]	1	100

- ① Los intervalos de la variable están ordenados por sus extremos de menor a mayor
- ② Determina la columna de frecuencias absolutas acumuladas (N_i) como sigue:

$$\begin{aligned} N_1 &= n_1 = 1 \\ N_2 &= N_1 + n_2 = 1 + 2 = 3 \\ &\vdots \\ N_{10} &= N_9 + n_{10} = 99 + 1 = 100 = n \end{aligned}$$

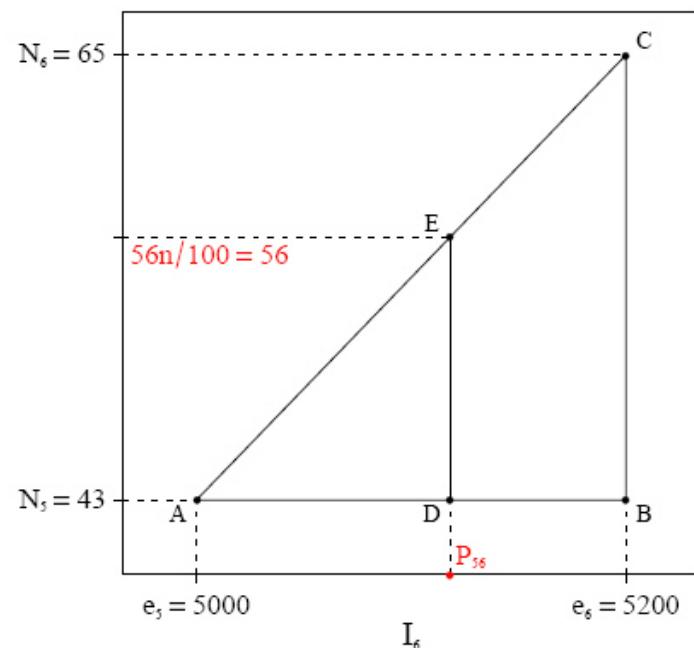
- ③ $56 \frac{n}{100} = 56$
- ④ Ninguna frecuencia absoluta coincide con 56, sino que

$$N_5 = 43 < 56 < 65 = N_6$$

Entonces $P_{56} \in I_6 = (5000, 5200]$ y $P_{56} = 5118.1818$

P_{56} es la solución de la ecuación $F(x) = 56/100 = 0.56$ con F la función de distribución de la variable, esto es, $n F(P_{56}) = 56 n / 100 = 56 \times 100 / 100 = 56$

► P30(X) P56(X)



- ① Determina la gráfica de la función $nF(x) = 100F(x)$ sobre el intervalo I_6 suponiéndola lineal
 - ② P_{56} es la abscisa que corresponde a la ordenada 56 de la función $nF(x)$, y suele calcularse teniendo en cuenta la semejanza de los triángulos \widehat{ABC} y \widehat{ADE} , de donde

$$\frac{AD}{AB} = \frac{DE}{BC}$$

o equivalentemente

$$\frac{P_{56} - e_5}{e_6 - e_5} = \frac{56 \frac{n}{100} - N_5}{N_6 - N_5}$$

$$P_{56} = e_5 + \frac{56}{\frac{100}{N_6 - N_5}} (e_6 - e_5) = 5000 + \frac{56 - 43}{\frac{65 - 43}{N_6 - N_5}} (5200 - 5000) = 5118.1818$$