

**Práctica. Estadística descriptiva**

1) Para los ejemplos 2.1 a 2.5 del material 1 realice lo siguiente:

- Indique un objetivo (si no está en el enunciado invente uno)
- A partir del objetivo, defina la población física. Mencione la unidad elemental y la variable.
- Indique además si se trata de un estudio muestral o poblacional.

2) Clasifique las siguientes variables e indique la unidad elemental en cada caso.

- Se registra la cantidad de tickets diarios generados en un Help Desk
- Se estudian los tiempos en resolver tickets críticos
- Nivel de urgencia asignado a cada ticket
- Se mide la longitud de unas barras de acero
- Se clasifican las barras de acero según su longitud en 5 grupos diferentes
- Se cuentan la cantidad de barras defectuosas por día durante un período de tiempo
- Motivo de detención de un determinado proceso
- Cantidad de días que demora la habilitación para comenzar a construir una casa

3) Dé algunos ejemplos de estudios donde sea imposible estudiar la población completa.

4) Durante el último mes, se revisaron los equipos defectuosos producidos en una empresa y se registró el defecto principal que poseían. Para cada defecto se registró su ubicación. Los valores obtenidos fueron: PQ – RR – SO – MQ – MO – CM (PQ: Procesador Quemado, RR : Ram Rota, SO: Sistema Operativo, MQ: Mother Quemada, MN: Monitor, PE: Periféricos). De los datos se obtuvo la siguiente distribución de frecuencias:

Clase (k)	$n_k$
PQ	15
RR	28
SO	37
MQ	39
MN	21
PE	10

- Defina y clasifique la variable.

- b. Muestre la información en un gráfico e indique sobre qué defecto/s deberán trabajar con más atención para reducir el número de equipos con defectos.

5) Los siguientes datos corresponden a los tiempos de arranque (en segundos) de 100 ordenadores de una Institución de Urgencias.

Tiempo	Frecuencia
$13,5 \leq t < 14,5$	8
$14,5 \leq t < 15,5$	9
$15,5 \leq t < 16,5$	18
$16,5 \leq t < 17,5$	20
$17,5 \leq t < 18,5$	35
$18,5 \leq t < 19,5$	10

- a. Realice un gráfico para visualizar la distribución de frecuencias.  
b. Marque en el gráfico los tiempos de arranque que superen los 15 segundos.

6) Una Soft Factory desea estudiar la relación entre el número de errores y el momento de la jornada de trabajo en que se producen las mismas. Se fijaron tres períodos de la jornada: 6 – 7 horas ; 10 – 11 horas y 13 – 14 horas y en cada uno de ellos se observó durante treinta días el número de defectuosos. Los resultados obtenidos fueron:

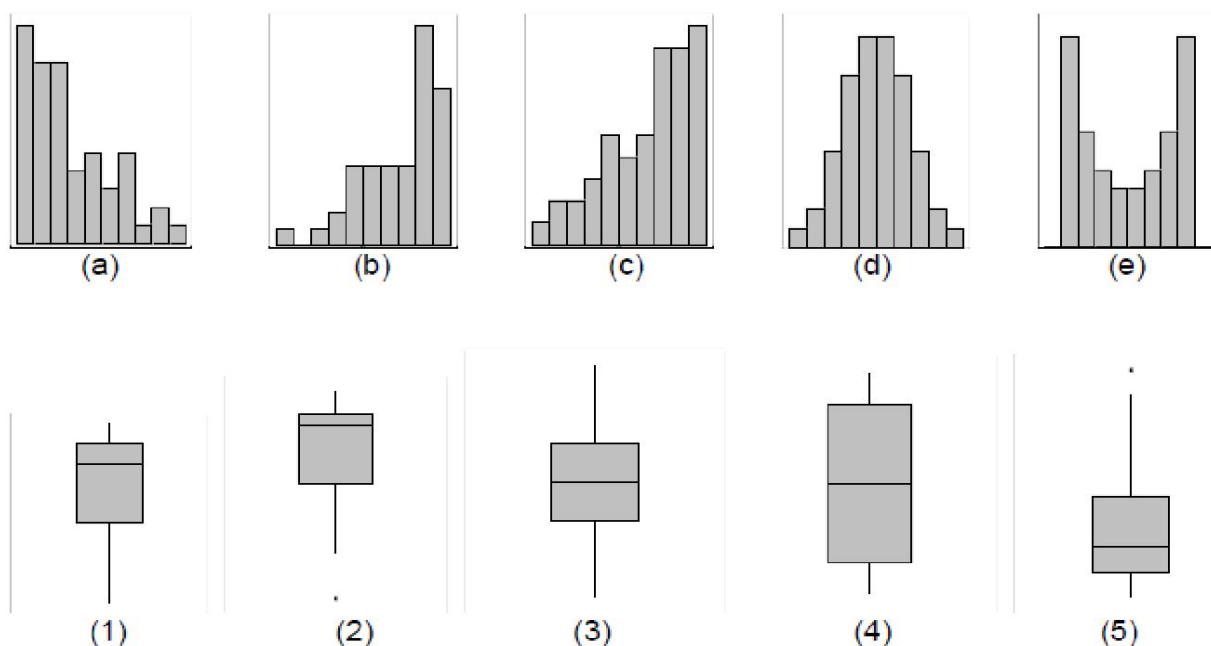
1º período 4 8 3 10 4 6 3 4 5 6 10 5 4 7 8 6 10 6 3 9 7 7 6 5 6 6 5 7 9 5

2º período 6 3 2 6 4 6 5 5 7 9 8 8 5 6 6 7 4 6 6 4 5 7 7 6 4 6 6 5 5 7

3º período 4 11 5 12 6 9 9 14 8 12 10 9 12 11 11 13 12 12 10 11 9 13 9 14 10 12 8 10 7 13

- a. Defina variable, unidad elemental y población.  
b. Realice un gráfico para poder observar cómo evoluciona durante los 30 días el número de errores en los 3 períodos. ¿En algún caso observa alguna tendencia?  
c. Para el período en que no se observe ninguna tendencia, patrón o ciclo, obtenga y grafique la distribución de frecuencias.

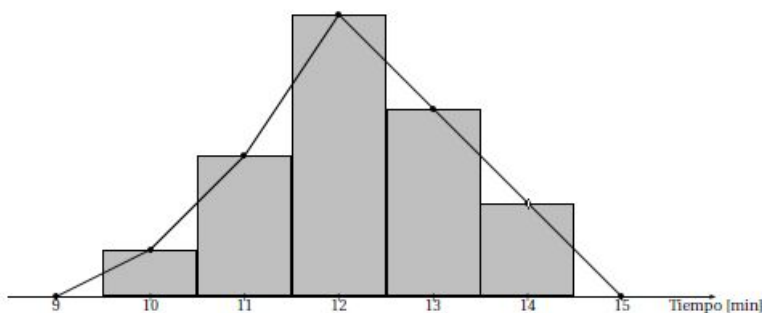
- 7) Realice un diagrama de caja con los datos del ejemplo 2.3 (puede obtener la información directamente de la tabla de frecuencias que ya se encuentra hecha)
- 8) Obtenga la moda en los ejemplos 2.1 a 2.5. (En los casos de variable cuantitativas continuas obtenga el intervalo moda y dé como valor de la moda su punto medio)
- 9) Con la información del problema 5 de esta práctica, obtenga la media y la mediana y explique por qué difieren sus valores.
- 10) En los ejemplos 2.4 y 2.5 obtenga:
- media, mediana y desvío estándar a partir de la distribución de frecuencias que se encuentra en el material teórico. Utilice para el cálculo de la media y el desvío estándar una planilla de datos (excel), o una aplicación o la calculadora en modo estadístico.
  - Interprete media y mediana.
  - Realice un diagrama de caja (obtenga las medidas directamente de la distribución de frecuencias).
- 11) Para cinco conjuntos de datos se construyeron los histogramas y el diagrama de caja correspondiente. Se presentan a continuación los gráficos obtenidos:



- Relacione cada histograma con el diagrama de caja que representa el mismo conjunto de datos.
- Considere los histogramas (d) y (e) e indique en cuál es mayor el desvío estándar y en cual es mayor el rango (no realice cálculos, considere solamente que todas las barras de los dos gráficos tienen el mismo ancho).
- Indique cuál conjunto de datos pareciera tener un valor anómalo.
- En los histogramas (a), (b) y (c) indique en cada caso cuál de las dos medidas de tendencia central, (media y mediana) es mayor (no compare estas medidas entre distintos grupos).
- ¿En cuál caso se verifica la regla empírica?

12) Un estudio sobre las indemnizaciones dictadas por un tribunal civil (daños personales, responsabilidad civil, etc.) determinó que la indemnización mediana es de \$8000 y que la indemnización media es de \$69000. Encuentre alguna explicación a esa diferencia tan grande entre esos dos valores característicos de la tendencia central.

13) El siguiente histograma y polígono de frecuencias relativas corresponden a 100 observaciones de los tiempos, en minutos, que una persona tarda para viajar con su auto desde su casa al trabajo.



Analice cuáles de las siguientes afirmaciones son verdaderas y cuáles son falsas. En cada caso justifique.

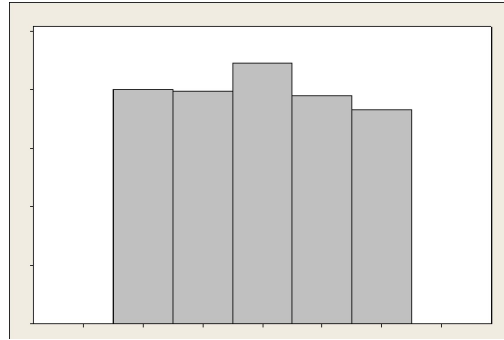
- El valor de la mediana es superior a 12.
- El valor del primer cuartil se encuentra en el primer intervalo de clase.
- El valor del tercer cuartil se encuentra en el cuarto intervalo de clase.
- La proporción de tiempos observados entre 11.5 y 12.5 es menor que la proporción de tiempos inferiores a 11.5.

- 14) Analice cuáles de las siguientes afirmaciones son verdaderas y cuáles son falsas.
- Si la media aritmética de artículos defectuosos que se producen por día es 2,3 entonces son más los días en que se producen 2 artículos defectuosos, que los días en que se producen 3 artículos defectuosos.
  - Si la desviación estándar representa el 1% de la media entonces el coeficiente de variación es 0:01.
  - Una industria tiene 3 fábricas: A, B y C. Los 40 obreros de la fábrica A ganan, en promedio \$280 por día; los 26 obreros de B, \$350 por día y los 34 obreros de C, \$300 por día. Los datos disponibles no son suficientes para calcular el ingreso medio por día en esa industria.
- 15) En la siguiente tabla se muestra la cantidad de accidentes de tránsito (por día), ocurridos en una ciudad durante el mes de septiembre.

Cantidad de accidentes	0	1	2	3	4
Cantidad de días	12	8	5	4	1

Analice si las siguientes afirmaciones referidas a la variable: número de accidentes por día, son verdaderas o falsas. Justifique.

- El valor de la mediana es igual a 2.
  - El valor de la moda es igual a 12.
  - La desviación estándar es menor que la media aritmética.
  - $Q2 - Q1 > Q3 - Q1$  (responda de ser posible sin realizar cálculos).
- 16) El histograma corresponde a 80 observaciones de la variable aleatoria T: (tiempo que una persona demora en trasladarse desde su casa al trabajo), medido en minutos. Un estudiante interpreta que la variable presenta muy poca variabilidad por cuanto las alturas de los distintos rectángulos presentan (poca variación). Analice si la afirmación del estudiante es o no correcta. Fundamente.



17) En una fábrica de láminas de vidrio una característica importante de la calidad es el grosor de las láminas. Para cierto tipo de lámina el grosor óptimo es de 5 mm con una tolerancia de 0.50 mm. Si la lámina tiene un grosor menor que 4.50 mm se considera demasiado delgada y no reunirá las condiciones de resistencia exigidas por el cliente. Si la lámina tiene un grosor mayor que 5.50 mm entonces se gastará demasiado material y elevarán los costos del fabricante. Por lo tanto, es de suma importancia fabricar láminas con el grosor óptimo, y en el peor de los casos dentro de las tolerancias especificadas. Se han seleccionado al azar 80 láminas y medido su grosor obteniéndose los siguientes valores:

4.82 4.87 5.10 4.95 5.52 4.52 5.34 4.48 4.73 5.18 5.03 5.04 4.92 4.90 4.83 4.75 5.10 4.73  
 5.04 5.06 5.33 5.18 4.61 5.03 4.66 4.82 4.77 5.11 5.36 5.10 4.87 5.00 5.02 4.98 5.24 5.33  
 4.93 5.02 4.91 4.72 5.23 5.06 5.48 5.02 4.90 4.62 5.05 4.95 4.88 4.92 4.50 4.68 5.30 5.30  
 5.62 5.00 5.04 4.56 5.30 4.82 4.74 5.40 5.05 5.34 5.28 4.87 5.24 5.20 5.19 5.15 5.14 5.22  
 5.10 4.72 4.58 5.50 4.95 5.42 5.46 4.94

- Defina la población física. Mencione la unidad elemental y la variable. Indique además si se trata de un estudio muestral o poblacional.
- Calcule la media y el desvío y con estas dos medidas obtenga un intervalo que contenga aproximadamente el 100% de los datos. En base a este intervalo y las especificaciones, indique si los grosores de las láminas son adecuados (en cuanto a centrado y variabilidad)