

Trimestre Enero-Marzo 2008
Departamento de Cómputo Científico y Estadística
Estadística para Ingenieros — CO3321
Guía de ejercicios Estadística Descriptiva

Práctica N° 1

1. En cada uno de los siguientes casos describa la población correspondiente, el objetivo inferencial y qué es lo que haría para obtener una muestra.
 - a) Un ingeniero pretende estimar el consumo semanal promedio de agua en casas unifamiliares, en una ciudad determinada.
 - b) Un político desea saber si la mayoría de los votantes de un estado están dispuestos a votar por él.
 - c) Un ingeniero electrónico desea determinar si la duración promedio de cierto tipo de transistores supera las 500 horas
2. Un censor de textos contabiliza el número de erratas que encuentra en cada página. Después de pasar este censor por un texto de 50 páginas, se obtiene el siguiente número de erratas por página:

2	3	5	0	1	4	0	6	2	1
1	0	2	4	5	3	1	2	3	2
2	5	4	1	3	2	6	8	2	0
1	0	2	3	1	5	10	2	1	3
3	1	2	4	4	6	2	0	1	3

- a) A partir del enunciado del problema, ¿qué tipo de variable estadística se está trabajando?.
 - b) Calcule las frecuencias absolutas, relativas y relativas acumuladas.
 - c) ¿Cómo sería el histograma de frecuencias absolutas? ¿Qué se puede decir mirando este gráfico?
 - d) Calcule los siguientes estadísticos descriptivos: media, mediana, moda, varianza y cuartiles. Interprete cada uno de acuerdo al contexto del problema.
 - e) Si se compara la media y la mediana, ¿qué se puede decir de la simetría de los datos? ¿Entre cuáles valores se encuentran el 50% (central) de los datos?
 - f) ¿Qué porcentaje de páginas respecto del total tienen menos de 6 erratas? ¿Y 6 erratas o más?
 - g) ¿Cuántos errores tienen el 10% de las páginas que menos errores presentan?
 - h) Construya el diagrama de caja respectivo y diga que conclusiones se pueden deducir al respecto.
3. Durante una epidemia de gripe, los tiempos de espera en cierto centro de salud fueron mayores a lo habitual. La siguiente tabla resume la distribución de los tiempos de espera para una muestra de 20 pacientes que visitaron el centro de salud durante este período,

Tiempo de espera (horas)	[0-1)	[1-2)	[2-3)	[3-4)
N° pacientes	6	9	4	1

- Calcule el tiempo medio de espera de los pacientes y la variabilidad del mismo.
 - Un nuevo paciente que llega al centro de salud plantea que él esperará para ser atendido si al menos la mitad de los pacientes es atendido en 1 hora y media. ¿Esperará el paciente para ser atendido?. ¿Por qué?
 - Calcule un valor aproximado del tiempo de mayor frecuencia para que un paciente sea atendido.
 - El centro de salud ha decidido que al 20% de los pacientes que más tardan en ser atendidos, los remitirá a una consulta adicional que ha habilitado. ¿Cuánto tiempo tiene que esperar como mínimo un paciente para ser enviado a esta consulta especial?
4. Las siguientes medidas corresponden a las estaturas (metros) de cincuenta niños y niñas.

1.56 1.59 1.63 1.62 1.65 1.61 1.59 1.51 1.62 1.62
 1.53 1.49 1.57 1.54 1.53 1.59 1.58 1.57 1.47 1.64
 1.55 1.59 1.53 1.56 1.53 1.47 1.57 1.60 1.54 1.56
 1.50 1.62 1.59 1.62 1.54 1.68 1.52 1.62 1.59 1.49
 1.65 1.53 1.59 1.56 1.54 1.58 1.52 1.63 1.56 1.62

- A partir del enunciado del problema, ¿qué tipo de variable estadística se está trabajando?.
 - Calcule las frecuencias absolutas, relativas y relativas acumuladas.
 - ¿Cómo sería el histograma de frecuencias relativas? ¿Qué se puede decir mirando este gráfico?
 - Calcule los siguientes estadísticos descriptivos: media, mediana, moda, varianza y cuartiles. Interprete cada uno de acuerdo al contexto del problema.
 - Si se compara la media y la mediana, ¿qué se puede decir de la simetría de los datos? ¿Entre cuáles valores se encuentran el 50% (central) de los datos?
 - ¿Qué número de niños tienen estatura menor de 1.53 metros? ¿Y 1.6 metros o más?
 - ¿Cuánto es la estatura menor del 5% de los niños más altos?
 - Construya el diagrama de caja respectivo y diga que conclusiones se pueden deducir al respecto.
5. En 1979 Michelson obtuvo los siguientes valores para la velocidad de la luz en el aire (se dan resultados restando 299.000 a los datos originales, en Km/seg):

850 740 900 1070 930
 850, 950 980 980 880
 1000 980 930 650 760

En 1882 Newcomb, utilizando otro procedimiento, obtuvo los siguientes resultados:

883	816	778	796	682
711,	611	599	1051	781
578	796	774	820	772

- a) Construya un histograma para cada conjunto de datos.
 b) Calcule la media, la mediana y la desviación típica para cada una de las muestras. Interprete lo observado
 c) ¿Qué conclusiones pueden extraerse?
 d) Construya un boxplot de cada experimentador en un mismo gráfico. ¿Se pueden extraer otras conclusiones a las encontradas en el inciso c)?, ¿Cuáles?.
6. Los siguientes datos corresponden al número de turistas extranjeros que han visitan una zona de Venezuela a lo largo de 12 años, en 2 diferentes temporadas:

Julio-Agosto	108	112	94	144	162	162	76	102	11	79	129	95
Diciembre	30	41	41	40	43	25	32	22	27	64	33	41

Según una agencia de viajes la variabilidad de los datos debe ser mayor en la temporada de Julio-Agosto. ¿Se corresponde esta hipótesis con lo observado en los datos?. Realice los cálculos que Usted considere necesarios para dar respuesta a esta interrogante, y justifique su respuesta.

7. Supongamos que 2 empresas desean repartir beneficios entre sus 4 principales accionistas, y que el reparto se realiza de la siguiente forma (en miles de Bolívares)

Jerarquía	Empresa A	Empresa B
1er. accionista	100	1200
2do. accionista	500	1300
3er. accionista	300	1400
4to. accionista	100	1100

¿Cuál de los 2 repartos es más equitativo?. Justifique su respuesta en base a un análisis estadístico del reparto.

8. Demuestre que la suma de las desviaciones de un conjunto de mediciones con respecto a su media, es igual a cero, es decir,

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

9. Utilizando las propiedades de las sumatorias estudiadas en clase, demuestre que:

$$s^2 = \frac{1}{n} \left[\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right]$$

Esta ecuación le será de utilidad para el cálculo de la varianza muestral s^2 .

10. Escriba cuál sería el efecto de una observación grande (outlier) sobre el intervalo $\bar{x} \pm s$