

## CAJA DE TUKEY

Lic. Jorge Matute, estadístico, INCAP.

La Caja de Tukey es un procedimiento gráfico que está siendo utilizado cada vez más en las ciencias biológicas y médicas, para comparar gráficamente diferentes grupos de población, así como los resultados obtenidos de ensayos científicos.

Como toda gráfica, la Caja de Tukey presenta un resumen de los datos obtenidos, con la ventaja, sobre otro tipo de gráficas, que las distribuciones de grupos graficados de esta manera pueden ser comparados. Comparación que se puede llevar a cabo en forma rápida, y previa o durante la realización del análisis estadístico.

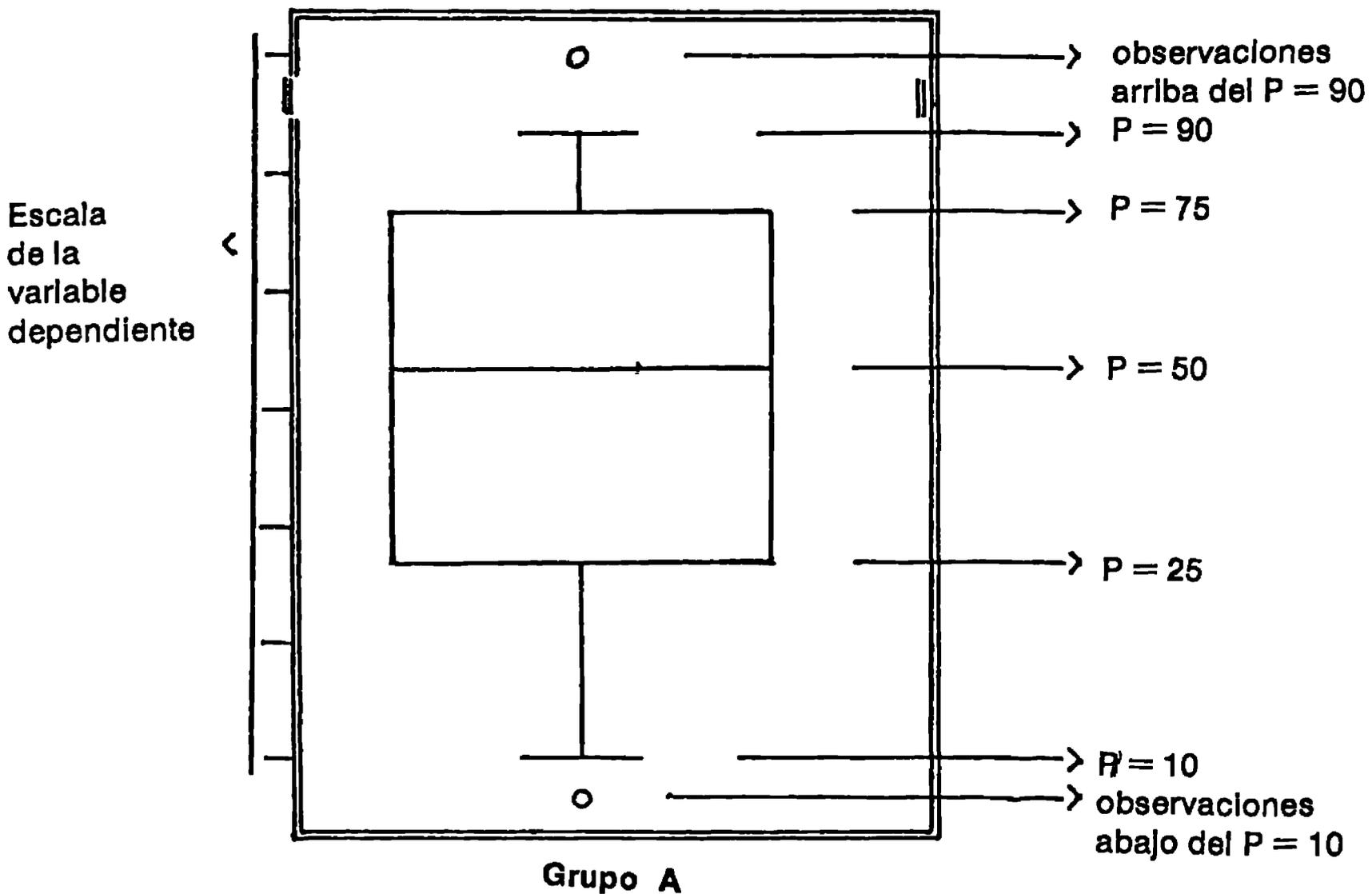
Esta comparación es posible debido a que la Caja de Tukey NO brinda únicamente una medida de tendencia central (en este caso la mediana) para cada uno de los grupos, sino que permite, además, conocer la variabilidad que existe en cada uno de ellos. Por lo que se puede visualizar fácilmente si las variabilidades presentadas por los grupos son semejantes.

La comparación de las distribuciones es posible debido a que lo que se compara, son los correspondientes percentiles de las mismas.

El gráfico de la Caja de Tukey de la distribución de un grupo de datos, es como su nombre lo indica, una caja en la cual se pueden identificar los siguientes percentiles:

- percentil 10
- percentil 25
- percentil 50 (mediana)
- percentil 75
- percentil 90

tal y como se puede apreciar en el siguiente esquema:



Si usted posee un computador tipo IBM o compatible, SAS ofrece este tipo de gráficas a través de PROC UNIVARIATE, si su computador es McIntosh de la APPLE, el paquete STATVIEW 512 se las hace. Si no se posee un computador pa

ra calcular los percentiles y elaborar las gráficas, entonces se puede hacer uso de las siguientes fórmulas:

Definamos:  $n$  = número de observaciones en el grupo  
 $P$  = percentil deseado

$$A = nP/100 + 0.5 = \underbrace{\quad}_{\text{Parte entera del No. = k}} \cdot \underbrace{\quad}_{\text{fracción = f}}$$

entonces el valor que contiene al percentil deseado está dado por :

$$V = X_k(1-f) + X_{k+1}(f)$$

donde  $X_k$  es la observación "k", después de que todas las observaciones han sido arregladas en forma ascendente.

### EJEMPLO:

Se tienen las siguientes observaciones ( $n=15$ ):  
 3, 4, 7, 5, 8, 9, 7, 2, 8, 10, 12, 1, 15, 25, 6.

1. se ordenan las observaciones en forma ascendente:

<u># de orden = k</u>	<u>Observación = x</u>
1	1
2	2
3	3
4	4
5	5

<u># de orden = k</u>	<u>Observación = x</u>
6	6
7	7
8	7
9	8
10	8
11	9
12	10
13	12
14	15
15	25

2. se desea el percentil 10,  $P=10$ :

$$A = 15(10)/100 + 0.5 = 2.00$$

de donde:

$$k=2, \quad f=0.00$$

de donde:

$$\begin{aligned} V &= X_2(1 - 0) + X_{2+1}(0) \\ &= 2(1) + 3(0) = 2 + 0 = 2 \end{aligned}$$

se desea el Percentil 25,  $P=25$ :

$$A = 15(25)/100 + 0.5 = 4.25$$

de donde:

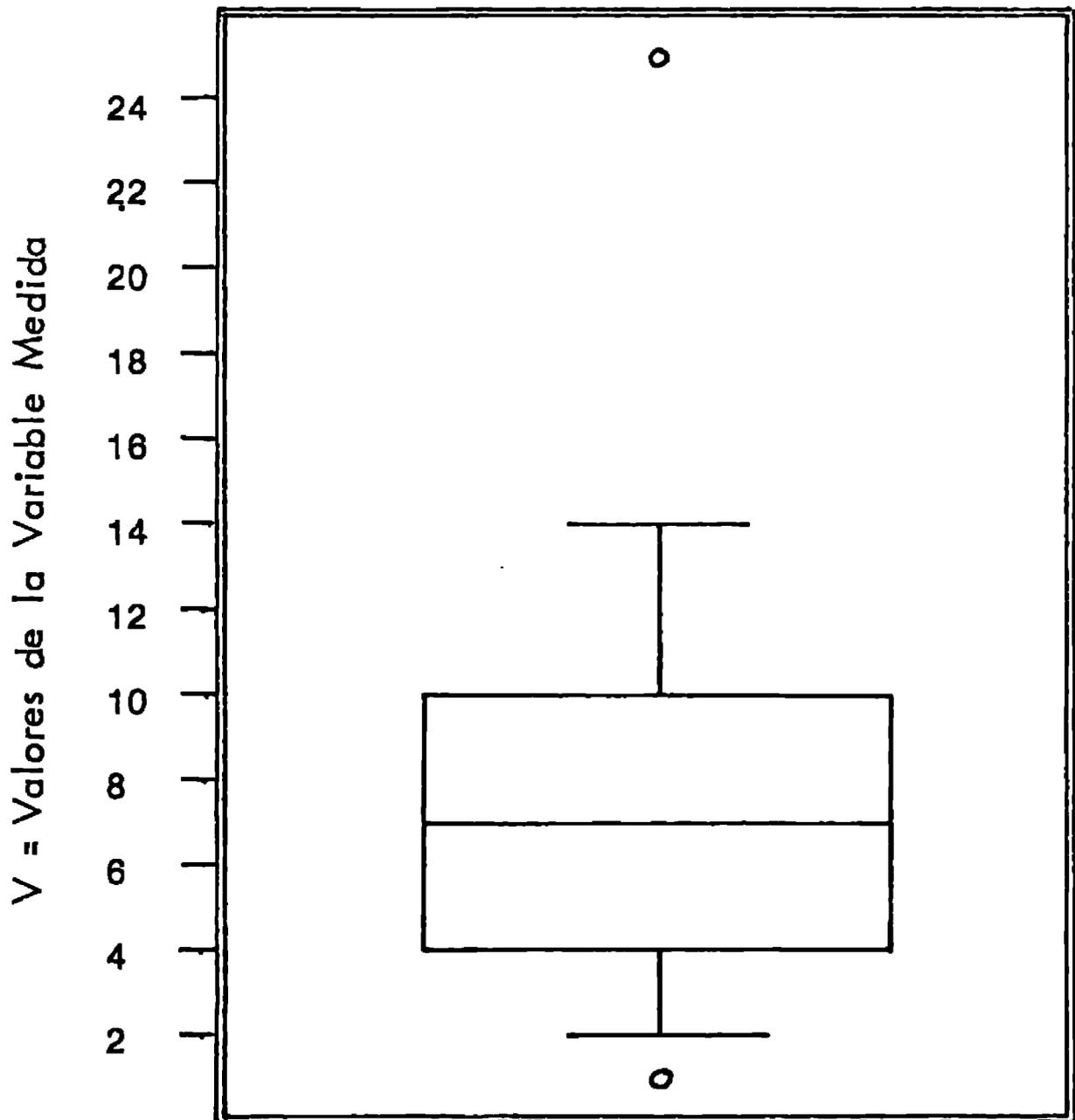
$$k=4, \quad f=0.25$$

de donde:

$$\begin{aligned} V &= X_4(1 - 0.25) + X_{4-1}(0.25) \\ &= 4(0.75) + 5(0.25) = 4.25 \end{aligned}$$

El Percentil 50 está ubicado en  $V=7$   
El Percentil 75 está ubicado en  $V=9.75$   
El Percentil 90 está ubicado en  $V=15$

3. La gráfica correspondiente a este grupo es entonces .



A continuación se presenta un ejemplo con algunos datos obtenidos en el estudio "Efecto Repelente y/o Insecticida de Algunos Extractos Vegetales sobre la Población de Mosca Blano

ca (*Bemisia tabaci* (Genn.), en Plantas de Frijol Común (*Phaseolus vulgaris* L.)", a ser publicado por Virginia Freire como tesis de grado (2).

La gráfica 1 representa la mortalidad de mosca blanca obtenida con ajo al 0.95% durante uno, tres, seis, nueve, y doce días después de hecha la aplicación en plantas de frijol. Los percentiles se muestran en el cuadro 1.

### CUADRO 1

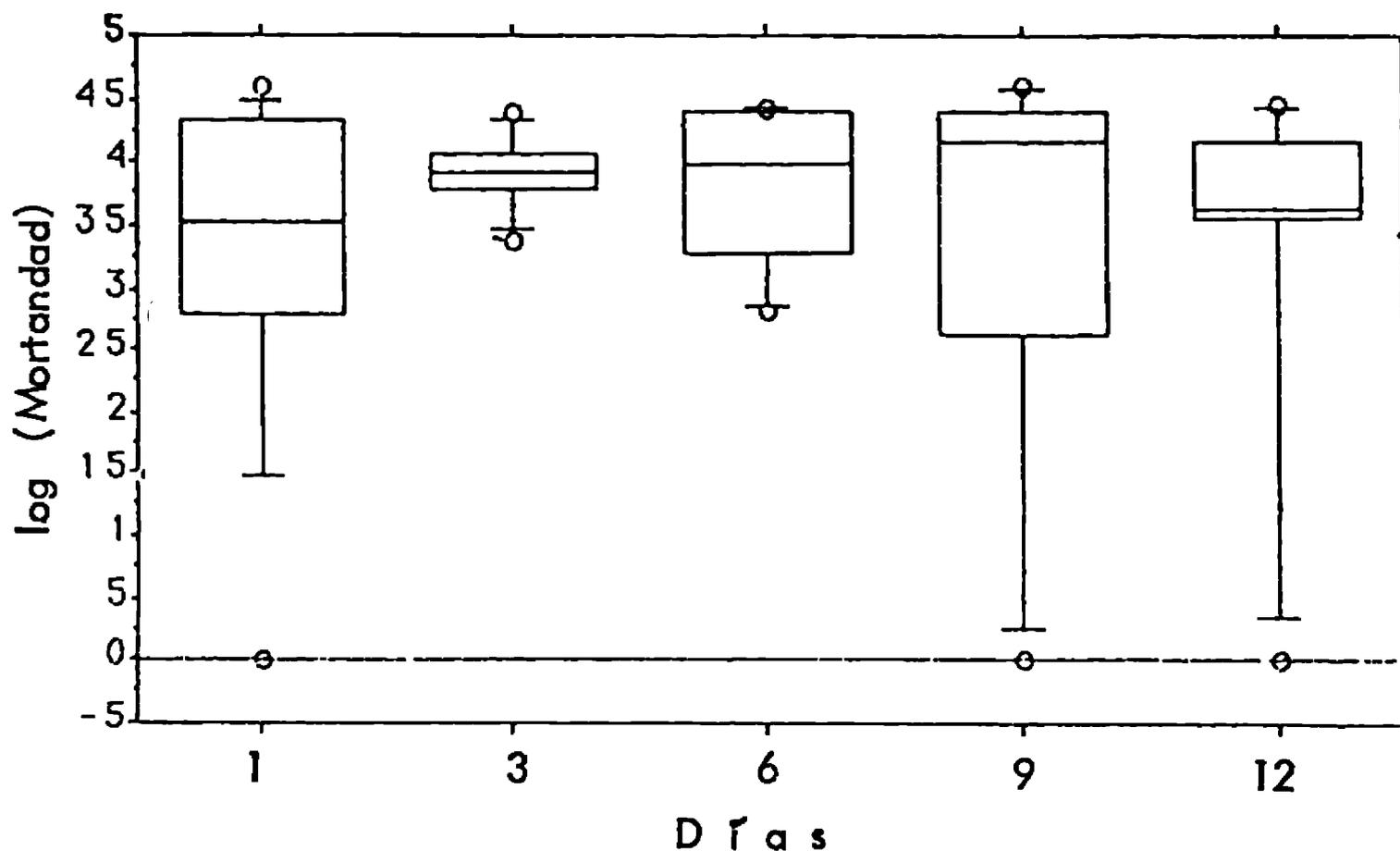
VALORES CORRESPONDIENTES A LOS PERCENTILES 10, 25, 50, 75 Y 90 DE LA MORTALIDAD BRINDADA POR EL AJO DE ACUERDO AL DIA

PERCENTIL	1	3	6	9	12
10	1.50	3.45	2.84	0.26	0.35
25	2.76	3.77	3.26	2.60	3.53
50	3.54	3.93	3.96	4.15	3.63
75	4.33	4.06	4.39	4.39	4.15
90	4.48	4.33	4.43	4.59	4.43

Los percentiles están dados en logaritmos naturales

## GRAFICA 1

MORTALIDAD BRINDADA POR AJO AL 0.95%, POR DIA



Se puede observar que el día tres presentó la menor variabilidad y que los días 3, 6 y 9 presentan las medianas más altas, siendo éstas muy parecidas, pero sus distribuciones no son del todo semejantes. Haciendo uso del siguiente razonamiento: "la gráfica es a un set de datos, lo que la fotografía es al ser humano", es posible darse cuenta que la gráfica indica algunas generalidades obtenidas durante el ensayo que orientarán en mejor forma el análisis estadístico.

## BIBLIOGRAFIA

1. CLEVELAND, William. The elements of graphing data. U.S.A., Wadsworth. 1985. pp. 123-144.
2. FREIRE, Virginia. Efecto repelente y/o insecticida de algunos extractos vegetales sobre la población de mosca blanca (*Bemisia tabaci*, genn.) en plantas de frijol común (*Phaseolus vulgaris*). Tesis de grado a ser publicada. Guatemala, USAC.