

# Caracterización estadística de la duración de la lluvia en Santa Clara, Cuba

PEDRO A. CÁRDENAS PÉREZ y LINO R. NARANJO DÍAZ

## RESUMEN

Usando la información pluviográfica disponible en la estación meteorológica No. 339, en Santa Clara (Cuba) se hace un estudio de duración de las precipitaciones, en el que se trata de obtener una distribución estadística teórica a la que se ajuste la muestra. Se determina que la distribución exponencial negativa es la que presenta mejor ajuste. Posteriormente, se estudia la relación entre los totales de lluvia y la duración de la misma, utilizando el índice de correlación, y se concluye que estas dos variables se relacionan a través de ecuaciones de regresión no lineales, en la mayoría de los meses del año.

## 1. INTRODUCCIÓN

El objetivo fundamental de este trabajo es describir algunos aspectos climatológicos de la precipitación, usando métodos estadísticos y centrando la atención en el aspecto pluviográfico, el que resulta de interés no solo por su importancia de índole científica, sino también por el hecho de que, siendo parámetros imprescindibles para diferentes intereses, los mismos son a la vez poco estudiados, al menos en nuestras condiciones.

En este sentido se busca una distribución teórica que se ajuste al tiempo de duración de la lluvia en los diferentes períodos. Además, se trata de encontrar una relación entre el total de precipitación y la duración de la misma.

## 2. MATERIALES Y MÉTODOS

Se procesaron, en total, 10 años de pluviogramas diarios (3 650 cartas) del período 1969-1978, de la estación meteorológica 339 (Santa Clara). Esta información se extrajo de las mediciones de un pluviógrafo soviético, modelo P-2, instalado sobre plazoleta con césped a altura estándar.

---

Manuscrito aprobado en julio de 1986.

P. A. Cárdenas Pérez y L. R. Naranjo Díaz pertenecen al Instituto de Meteorología, de la Academia de Ciencias de Cuba.

El procesamiento de las cartas se hizo siguiendo la metodología usual, tomando dentro de una lluvia tantas intensidades como trazos existieran dentro de dos puntos de cambio de la pendiente. La Hora de comienzo de la lluvia se tomó como la hora de la primera lluvia del día, por lo que el tamaño de esta muestra es 1 247 casos (días con lluvia).

En cuanto a la duración de la lluvia, algunos autores plantean que puede modelarse de acuerdo con una función de distribución exponencial truncada, del tipo:

$$f(T) = A(T_0) e^{-A(T_0)T} T \leq T_0 \text{ ó } T > T_0 \quad (1)$$

El parámetro  $A(T_0)$  es una constante determinada por el valor  $T_0$ , y la variable  $T$  representa el tiempo total de duración de la precipitación dentro de un plazo de 24 horas (SUSUKI, 1967), de manera que:

$$T_i = \sum_{j=1}^n \Delta T_{ij} \quad (2)$$

donde  $\Delta T_{ij}$  es la duración de  $j$ -ésimo intervalo con precipitación ocurrida en el  $i$ -ésimo día.

Buscar una relación entre el total de precipitación y el tiempo de duración de la misma tiene un gran interés desde muchos puntos de vista; algunos autores (SUSUKI, 1967) han encontrado relaciones matemáticas con buenas aproximaciones, como la relación:

$$R = aT^n + \epsilon \quad (3)$$

donde  $R$  representa el total de precipitación,  $T$  el tiempo de duración, y  $a$  y  $n$  son constantes a estimar a partir de la muestra. El parámetro  $\epsilon$  es un término de corrección que considera las variaciones no explicables.

### 3. RESULTADOS Y DISCUSIÓN

En el análisis de los resultados se incluyen como parámetro adicional las horas de comienzo de la lluvia para diferentes períodos, así como lo planteado de duración y relación entre totales y duración.

#### 3.1 Horas de comienzo de la lluvia

En la Tabla 1 se presentan los resultados de estos cálculos, en los que se puede ver con facilidad cómo en los meses del período seco no existen grandes diferencias entre las horas de comienzo durante todo el día. Esto se debe a que en estos meses las precipitaciones son de origen frontal en su gran mayoría y no tienen una hora preferente de afectar a una región determinada. Por el contrario, para los meses del período lluvioso, las horas de comienzo se agrupan en horarios posteriores al mediodía, en que ocurre el máximo de calentamiento superficial. Este resultado, esperado, pone de manifiesto de una manera indirecta la génesis de la precipitación para estos meses.

TABLA 1. Frecuencias relativas de horas de comienzo de la lluvia en intervalos trihorarios.

Horas	Periodos												Año		
	XI	XII	I	II	III	IV	P.s.	V	VI	VII	VIII	IX		X	P.II.
07-10	11,0	14,5	12,1	7,0	10,5	8,0	10,6	4,4	2,2	1,2	3,9	1,5	8,7	3,7	5,9
10-13	12,3	10,7	15,0	9,6	9,3	10,2	11,5	8,8	9,2	5,9	7,5	10,8	12,5	9,3	10,0
13-16	13,1	14,5	11,3	17,8	10,5	19,0	14,1	34,5	32,5	29,1	30,2	39,8	20,3	31,0	25,8
16-19	14,9	15,3	15,0	27,2	23,3	21,9	18,6	30,1	29,7	40,2	34,5	23,6	20,9	29,1	25,9
19-22	13,6	15,3	12,8	11,4	17,4	19,0	14,7	10,1	16,5	16,9	12,2	15,6	12,2	13,8	14,2
22-01	14,5	9,2	10,6	9,6	9,3	8,0	10,7	4,7	4,8	4,3	5,7	5,1	9,3	5,7	7,3
01-04	11,4	9,2	12,1	10,5	9,3	8,8	10,4	3,4	3,4	2,4	2,5	1,5	9,0	3,8	5,8
04-07	9,2	11,4	11,3	8,8	10,5	5,1	9,3	4,0	1,7	0,0	3,2	2,1	7,0	3,1	5,0

### 3.2 Duración de la lluvia

Las distribuciones teóricas obtenidas que presentan mejor ajuste a esta variable son las del tipo:

$$f(T) = \lambda e^{-\lambda T}$$

donde

$$\lambda = 1/T \quad (4)$$

Este resultado difiere esencialmente del obtenido por SUSUKI (1967) en el hecho de que la distribución encontrada no tiene necesidad de condiciones restrictivas en cuanto al valor del parámetro  $A(T_0)$ . Esto se debe a la identidad con el parámetro  $\lambda$  definido para las distribuciones exponenciales negativas, a las que solo se le puede aplicar la restricción:

$$0 < T_i \leq 24 \quad (5)$$

que surge de las características propias de la muestra.

Para la bondad del ajuste se utilizó la prueba no paramétrica de Kolmogorof-Smirnov, con un nivel de significación  $\alpha = 0,05$  (DIXON y MASSEY, 1965), con el que contrastó la hipótesis:

$H_0$ : La distribución muestral se ajusta a una distribución exponencial negativa.

$H_1$ : La distribución muestral no se ajusta a una distribución exponencial negativa.

La razón por la que se utilizó esta prueba y no la comúnmente utilizada  $\chi^2$  es que en las distribuciones muestrales aparecen intervalos con frecuencias relativas inferiores a 0,01.

En la Tabla 2 se brindan los resultados de los ajustes al año, al período lluvioso, y al período seco, respectivamente. Es de destacar que los ajustes son muy buenos en los tres casos y no existen diferencias marcadas entre los tipos de distribución, como tampoco entre los valores medios de duración de cada lluvia entre los tres períodos. La diferencia que existe no resulta ser la duración de cada lluvia sino el tiempo total que permanece lloviendo en cada mes y período. En la Tabla 3 se presenta este resultado, donde se pone de manifiesto la diferencia señalada. En el período seco permanece lloviendo 29,0% del tiempo total que llueve en el año, mientras que en el período lluvioso éste es de 70,9% del total.

TABLA 2. Distribución de frecuencias de la duración de la lluvia en los períodos anual, lluvioso, y seco, en Santa Clara. En la tabla aparecen los totales de casos, las frecuencias observadas por intervalos (fo), las frecuencias esperadas acumuladas (Fo), y las frecuencias esperadas acumuladas (Fe). Además, aparecen las discrepancias (D) entre Fe y Fo.

Duración (horas)	Total de casos	fo	Fo	Fe	D
Período anual					
0-1	785	0,629	0,629	0,640	0,011
1-2	248	0,199	0,828	0,838	0,010
2-3	107	0,086	0,914	0,928	0,014
3-4	51	0,040	0,954	0,967	0,013
4-5	21	0,017	0,971	0,985	0,014
5-6	14	0,011	0,982	0,993	0,011
6-7	6	0,006	0,987	0,997	0,010
7-8	7	0,005	0,993	0,998	0,005
8-9	3	0,002	0,995	0,999	0,004
9-10	1	0,000	0,996	0,999	0,003
10-11	0	0,000	0,996	0,999	0,003
11-12	3	0,002	0,998	0,999	0,002
12-13	1	0,001	0,999	0,999	0,002
$\bar{X} = 1,25$	$\lambda = 0,80$		$D_{0,05} = 0,038$		$D_{\max} = 0,014$
Período lluvioso					
0-1	556	0,6164	0,6164	0,6410	0,025
1-2	189	0,2095	0,8259	0,8380	0,012
2-3	77	0,0853	0,9112	0,9272	0,016
3-4	42	0,0466	0,9578	0,9672	0,009
4-5	17	0,0188	0,9766	0,9852	0,008
5-6	9	0,0099	0,9866	0,9934	0,007
6-7	3	0,0033	0,9899	0,9970	0,007
7-8	3	0,0033	0,9932	0,9986	0,005
8-9	3	0,0033	0,9965	0,9994	0,003
9-10	0	0,0000	0,9965	0,9995	0,003
10-11	0	0,0000	0,9965	0,9999	0,003
11-12	2	0,0022	0,9987	0,9999	0,001
12-13	1	0,0011	0,9998	0,9999	0,000
$\bar{X} = 1,25$	$\lambda = 0,80$		$D_{0,05} = 0,45$		$D_{\max} = 0,025$

(Continúa)

TABLA 2 (Continuación)

Duración (horas)	Total de casos	fo	Fo	Fe	D
Período seco					
0-1	229	0,6640	0,6640	0,6396	0,024
1-2	59	0,1710	0,8350	0,8397	0,005
2-3	30	0,0870	0,9222	0,9287	0,007
3-4	9	0,0260	0,9482	0,9683	0,020
4-5	4	0,0120	0,9602	0,9859	0,026
5-6	5	0,0140	0,9742	0,9937	0,020
6-7	3	0,0090	0,9832	0,9972	0,010
7-8	4	0,0120	0,9952	0,9988	0,004
8-9	0	0,0000	0,9952	0,9994	0,004
9-10	1	0,0033	0,9985	0,9997	0,002
10-11	0	0,0000	0,9985	0,9999	0,002
11-12	1	0,0033	0,9998	0,9999	0,000
12-13	0	0,0000	0,9998	0,9999	0,000
$\bar{X} = 1,23$	$\lambda = 0,81$		$D_{0,05} = 0,073$		$D_{max} = 0,026$

### 3.3 Relación entre totales y duración de la precipitación

El primer paso en este sentido resultó el probar con la relación (3); para esto se construyeron los diagramas de dispersión de  $R$  contra  $T$  y de  $\log R$  contra  $\log T$ , con el objeto de calcular con métodos gráficos las constantes  $a$  y  $n$ .

Aunque ya de los diagramas de dispersión se podía observar la no existencia de una relación del tipo planteado, se comprobó por el método de los mínimos cuadrados el mal ajuste de los mismos y los resultados no satisfactorios.

El anterior resultado nos llevó a cuestionar la existencia o no de una relación curvilínea de otro tipo.

La relación entre dos variables aleatorias simples se puede poner de manifiesto mediante el cálculo del índice de correlación  $N_{xy}$ , que está definido por:

$$N_{xy} = \sqrt{\frac{\sigma^2(Y_x)}{\sigma^2(Y)}} \quad (6)$$

TABLA 3. Duración total media de la lluvia para cada mes y período, y porcentaje total del tiempo del mes que permanece lloviendo.

Mes	Duración total (horas)	Tiempo del período (%)
Noviembre	15,4	2,1
Diciembre	8,5	1,1
Enero	10,6	1,4
Febrero	8,0	1,2
Marzo	5,2	0,7
Abril	9,1	1,3
Período Seco	56,7	1,3
Mayo	20,7	2,8
Junio	33,6	4,7
Agosto	17,0	2,3
Septiembre	19,2	2,6
Octubre	25,5	3,4
Período Lluvioso	138,7	3,1
Año	195,5	2,2

donde  $\sigma^2(Y_x)$  es la varianza intragrupo o varianza de las medias de los intervalos, y  $\sigma^2(Y)$  es la varianza para la muestra (ULANOVA, 1964). Es evidente que si la relación es lineal, este índice de correlación se convierte en el coeficiente de correlación lineal  $R_{xy}$ .

Los índices de correlación calculados entre totales de precipitación y duración total de cada lluvia fueron los siguientes: enero,

TABLA 4. Coeficientes de la ecuación de regresión parabólica de los diferentes meses, así como los errores de estimación ( $s_{yx}$ ). Los coeficientes corresponden con la ecuación  $Y = Ax^2 + Bx + C$ .

Meses	Coeficientes de la ecuación de regresión			
		B	C	$S_{yx}$
Enero	-0,000 053	0,051	3,59	3,8
Junio	-0,000 066	0,080	3,71	8,6
Julio	-0,000 273	0,135	3,97	5,0
Octubre	-0,000 044	0,152	1,64	13,3

0,712; febrero, 0,434; marzo, 0,703; abril, 0,681; mayo, 0,370; junio, 0,694; julio, 0,820; agosto, 0,370; septiembre, 0,627; octubre, 0,724; noviembre, 0,499; diciembre, 0,950. Se demuestra que, excepto para los meses de mayo, agosto, y febrero, existe una cierta correlación que se hace máxima en el mes de diciembre con los restantes valores en una gama aceptable. Para los períodos seco, lluvioso, y anual, los valores calculados de este índice caen por debajo de 0,4.

Del análisis de los diagramas de dispersión y de las tablas de cálculo del parámetro  $N_{xy}$  se obtuvo la posibilidad de ajustar los valores muestrales a ecuaciones parabólicas del tipo:

$$R = A + BT + CT^2 \quad (7)$$

En la Tabla 4 se muestran los resultados obtenidos para algunos meses tipos de la muestra.

#### 4. CONCLUSIONES

- A) Los momentos de comienzo de la lluvia tienen horas preferentes solo en los meses del período lluvioso.
- B) Las muestras de duración de la lluvia tienen para esta localidad un buen ajuste a distribuciones exponenciales negativas.
- C) La relación hallada por SUSUKI (1967) para Tokio no se ajusta a las muestras estudiadas.
- D) Entre las variables  $R$  y  $T$  existe una relación parabólica de la forma  $R = A + BT + CT^2$ , con excepción de los meses señalados.
- E) No fue posible obtener relaciones entre  $R$  y  $T$  para los períodos seco, lluvioso y anual.

#### REFERENCIAS

- DIXON, W. J., y MASSEY, F. J., Jr. (1965): *Introducción al análisis estadístico*. Ediciones Revolucionarias, La Habana, 1971, 489 pp.
- SUSUKI, E. (1967): A statistical and climatological study on the rainfall in Japan. *Pap. Meteorol. Geophys.*, 18(3):103-181.
- ULANOVA, E. C. (1964): *Priminenie matematicheskoi statistiki v agrometeorologii dlia najodenia uravnenia sviasei* [en ruso]. Moscú, 112 pp.

#### STATISTICAL CHARACTERIZATION OF RAINFALL IN SANTA CLARA, CUBA.

#### ABSTRACT

The available information on rain in station number 339 in Santa Clara, Cuba, was used to study the duration of precipitation, trying to obtain a theoretical statistical distribution to fit the sample. A negative exponential distribution was found to provide the best fitness. Moreover, the relationship between duration and total precipitation is studied using the correlation index. It was found that the relationship was a non-linear regression equation in the majority of the cases.