LA VARIABILIDAD DE LAS LLUVIAS EN ESPAÑA

La terrible sequía padecida en el otoño del año 1953 y el recuerdo de las registradas en los años agrícolas 1944-45 y 1948-49 nos impulsan a estudiar, siquiera ligeramente, la variabilidad de las lluvias—o, más generalmente, las precipitaciones, pues se incluyen en éstas las nevadas y granizadas—registradas en España durante un período lo más largo posible.

En el Calendario Meteoro-Fenológico 1950 se publicaron unas series de datos de lluvias anuales que correspondían a los lugares en que comenzaron antes las observaciones. Algunos empezaron desde hace bastantes años. Algunas, a partir de mediados del siglo pasado o algo antes. Sin embargo, como los años iniciales de esas series de observaciones no eran concordantes y algunas tenían cortes notables, no consideramos aquí sino **once**. Con estas once tratamos de estudiar la **variabilidad** de la lluvia anual sobre España.

Esas once estaciones son: **La Coruña** y **San Sebastián**, en el litoral cantábrico; **Valladolid** y **Soria**, en la cuenca del Duero; **Huesca**, en la del Ebro; **Madrid**, en la del Tajo; **Badajoz**, en la del Guadiana; **Murcia**, en la del Segura; **Barcelona** y **Alicante**, en el litoral mediterráneo, y **San Fernando (Cádiz)**, en el del Golfo de Cádiz. Mayor número de estaciones hubiéramos deseado incluir en esta lista, ya que existen otras, como las de Tortosa (Tarragona), Sevilla, Granada, Zaragoza, etc., que tienen series largas, pero las dificultades antes indicadas lo han impedido.

Así, pues, manejamos sólo esas once y limitamos su período al de 1881-1948, que es el que, en los citados cuadros del "Calendario Meteoro-Fenológico de 1950", aparece ser común a todas.

* * *

La **variabilidad de una serie** se estudia en Estadística del modo siguiente, universalmente empleado:

Ante todo, se ordenan los valores de menor a mayor. El intervalo—o **recorrido**—entre los valores extremos ya nos indica si la serie es muy **dispersa** o no, pues si es muy grande ese recorrido, entre las mayores y las menores lluvias registradas hay gran diferencia, y, por consiguiente, pueden presentarse muy variados valores de ellas. Aunque este primer índice estadístico, el recorrido o intervalo, ya nos da una primera y grosera idea de la dispersión de la serie, no lo consideramos suficiente. Hay que buscar otros índices más expresivos.

Para ello, tenemos que calcular un valor central de la serie. Y, de todos los valores centrales que se han propuesto, el más conocido y usado es el de la **media aritmética**, o sea, el cociente de dividir la suma de todos los términos por el número de ellos, que en este caso es 68.

Este índice estadístico, la media aritmética, se suele representar por **M**, y, aparte de ser el aplicado en todo el mundo desde hace muchos años, tiene una propiedad fundamentalísima que le da un carácter de "centralidad" de la serie que no tienen otros. Esa propiedad es la siguiente:

"Si se resta de cada término la media aritmética de la serie y se suman las **diferencias positivas** obtenidas, el resultado es igual a la suma de las **diferencias negativas**." En la práctica no se cumple exactamente esto porque la media aritmética se halla con varias cifras decimales que, muchas veces, se desprecian; pero esto no tiene importancia alguna en las aplicaciones reales.

Por esa propiedad de centralidad de la media aritmética, se le concede en Estadística Matemática no una representación de la serie total—que no la ostenta—, sino la categoría de firme **punto de apoyo** de todo estudio de esa misma serie.

Las diferencias—que acabamos de citar—entre la media aritmética y cada uno de los términos de la serie resultan, naturalmente, positivas o negativas y se llaman **desviaciones.**

La lista de ellas, prescindiendo de que sean positivas o negativas, puede presentar uno de estos dos caracteres: o son muy grandes, o son muy pequeñas. Si son muy grandes, quiere ello decir que los términos de la serie están muy alejados de la media aritmética, y, por lo tanto, la serie es muy **dispersa.** Al contrario, si las desviaciones son pequeñas, significa esto que los términos están muy agrupados alrededor de la media aritmética, o sea, que la serie es **poco dispersa.**

Por consiguiente, si sumamos todas las desviaciones, prescindiendo ya de su signo, esa suma será: grande si la serie es muy dispersa y pequeña si no lo es.

Se suele dividir esa suma por el número de desviaciones—que es el mismo que el de términos de la serie—y al cociente resultante se le llama **desviación media**, siendo él un nuevo índice de "dispersión", más expresivo que lo era el "recorrido", antes citado. La desviación media suele representarse por la letra griega μ (mu minúscula).

Pero aun podemos afinar más, si bien tendremos que razonar en lo que sigue con argumentos más bien intuitivos que rigurosamente científicos, pues, para hacerlo de esta última manera, sería necesario adentrarse de lleno en las teorías de la Estadística matemática.

Diremos, pues, de un modo elemental que esas desviaciones podemos elevarlas al cuadrado, sumar los resultados obtenidos y dividir esa suma por el número de desviaciones. El nuevo índice que obtenemos—el cual también es un índice de la mayor o menor dispersión de la serie—se llama la **varianza**.

Tomamos los cuadrados de las desviaciones por dos razones: una es la de que así se hacen todas positivas; y otra, la de que las hacemos—pudiéramos decir—más visibles, "de mayor relieve".

Sin embargo, eso tiene inconvenientes y, por ello, una vez obtenido este índice estadístico, la varianza, extraemos la raíz cuadrada de él, llegando así a otro nuevo índice de dispersión, el llamado **desviación típica**—en inglés "standard"—, que se representa en todo el mundo por la letra griega σ (sigma minúscula).

Este índice es el más importante en Estadística Matemática. Puede decirse que es el punto central de todas las teorías de esta moderna ciencia.

Concretemos, pues, que **"desviación típica"** (σ) es la raíz cuadrada de la suma de los cuadrados de las desviaciones de los términos de una serie dada, tomadas con respecto a su media aritmética y dividida esa suma por el número de términos. Si representamos las desviaciones por la letra griega Δ (delta mayúscula) y la colección de ellas por Δ_1, Δ_2, Δ_3, ..., Δ_h, se tendrá:

$$\sigma = \pm \sqrt{\frac{\Delta_1^2 + \Delta_2^2 + \Delta_3^2 + \ldots + \Delta_n^2}{n}}$$

Expliquemos ahora que tomando la media aritmética, que, según dijimos al principio, está representada por M como punto central podemos formar unos límites que sean: el superior $M + \sigma$ y el inferior $M - \sigma$. Entre ellos —y según ciertas teorías estadísticas—deben estar comprendidos el 68 por 100 de los términos de la serie que estamos estudiando, si la distribución de ellos es normal.

Formando ahora unos nuevos límites, el superior $M + 2\sigma$ y el inferior $M - 2\sigma$, entre ellos deben estar comprendidos el 95 por 100 de los términos de la serie.

Finalmente, entre los límites $M + 3\sigma$ y $M - 3\sigma$ deben estar comprendidos el 99 por 100 de los términos, es decir, casi todos ellos.

Si la distribución de los términos de la serie no obedece a la que acabamos de dar, sino que la mayor parte de los términos, es decir, más de un 68 por 100 de ellos está comprendida entre los límites más estrechos, quiere decir ello que la serie es **muy poco dispersa**, o sea, que todos sus términos difieren muy poco de la media aritmética. En este caso, adquiere ésta mucha significación como índice estadístico, porque casi todos los términos están apiñados a su alrededor. Además, se llega a la conclusión de que **es pequeña la variabilidad** de la serie.

Todo lo contrario hay que decir si en el grupo central, entre $M + \sigma$ y $M - \sigma$, hay mucho menos del 68 por 100 del número de términos. Entonces, obtendríamos la consecuencia de que la serie era muy dispersa, **muy variable**, y la media aritmética no puede ostentar la representación de la serie.

Aún se suele multiplicar σ por el factor fijo 0,6745 (muy fácil de recordar, porque tiene las cifras 4, 5, 6 y 7), y así se obtiene $F = 0{,}6745 \times \sigma$, nuevo índice estadístico que se llama **desviación probable.** Fundándose en éste, la distribución de los términos de la serie debe ser la siguiente, si esa distribución es normal:

Entre $M + F$ y $M - F$ 2 debe haber el 50 % de términos.
Entre $M + 2F$ y $M - 2F$ debe haber el 82 % de términos.
Entre $M + 3F$ y $M - 3F$ debe haber el 96 % de términos.

En esta nueva distribución es más fácil que en la anterior comprobar si se cumple o no la ley de normalidad, pues entre los primeros límites debe haber el 50 por 100, o sea, la mitad del número de términos; y entre los segundos, el 82 por 100, ó, aproximadamente, el 80 por 100, o sea, las cuatro quintas partes del número de términos.

Nosotros, sin embargo, adoptaremos en la aplicación al estudio de las series de lluvias el primer método, o sea, el del empleo de $M + \sigma$ y $M - \sigma$, etc.

* * *

Aplicando todo lo que acabamos de decir a las once estaciones españolas citadas al principio, obtenemos el siguiente cuadro I.

CUADRO I

Lluvias en periodo de sesenta y ocho años (1881-1948).

ESTACION	Lluvia media anual (en mm.)	Valor de σ (en mm.)
La Coruña	830	222
San Sebastián	1.386	333
Valladolid	389	101
Soria	551	120
Huesca	535	163
Barcelona	554	130
Badajoz	507	157
Madrid	431	91
Murcia	316	122
Alicante	324	127
San Fernando (Cádiz)...	584	163

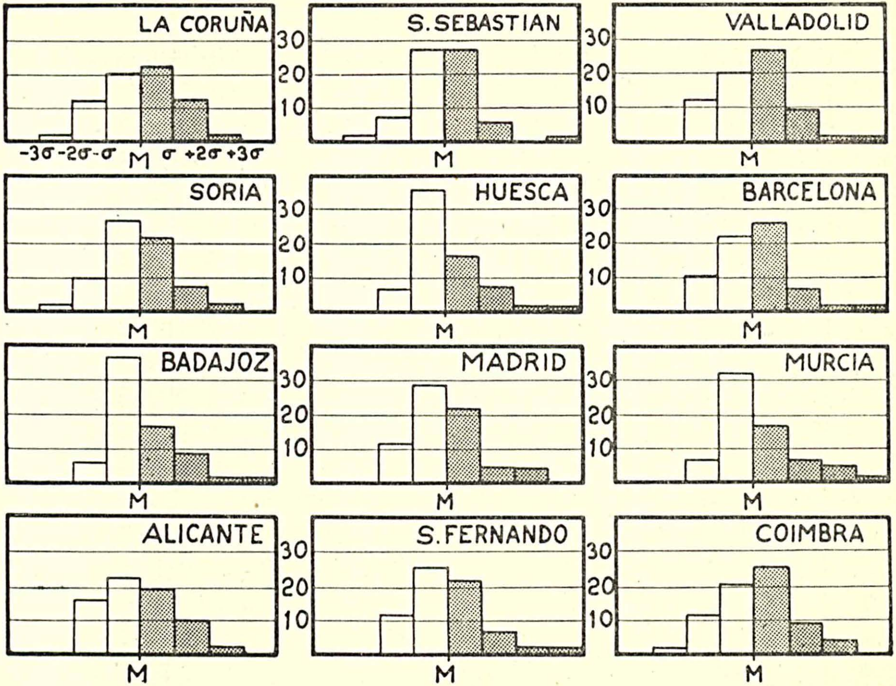
CUADRO II.—**Variabilidad de lluvias en España.**

ESTACION		ENTRE M — σ y M + σ	ENTRE M — 2σ y M + 3σ	ENTRE M — 3σ y M + 3σ
La Coruña.	Límites de cantidad lluvia, mm.	608 a 1.052	386 a 1.074	164 a 1.496
	Número de casos	41	66	68
	Idem íd. (en %)	61	98	100
	Debía ser (en %)	68	95	99
San Sebastián ...	Límites de cantidad lluvia, mm.	1.053 a 1.719	720 a 2.052	387 a 2.385
	Número de casos	54	66	67
	Idem íd. (en %)	80	98	99
	Debía ser (en %)	68	95	99
Valladolid..	Límites de cantidad lluvia, mm.	288 a 490	187 a 591	86 a 692
	Número de casos	45	66	67
	Idem íd. (en %)	67	98	99
	Debía ser (en %)	68	95	99
Soria	Límites de cantidad lluvia, mm.	432 a 671	311 a 791	191 a 911
	Número de casos	48	65	67
	Idem íd. (en %)	71	96	99
	Debía ser (en %)	68	95	99
Huesca	Límites de cantidad lluvia, mm.	372 a 698	209 a 861	46 a 1.024
	Número de casos	51	66	67
	Idem íd. (en %)	76	98	99
	Debía ser (en %)	68	95	99

ESTACION		ENTRE $M - \sigma$ y $M + \sigma$	ENTRE $M - 2\sigma$ y $M + 3\sigma$	ENTRE $M - 3\sigma$ y $M + 3\sigma$
Barcelona.	Límites de cantidad lluvia, mm.	424 a 684	294 a 814	164 a 944
	Número de casos	49	65	66
	Idem íd. (en %)	73	96	98
	Debía ser (en %)	68	95	99
Badajoz	Límites de cantidad lluvia, mm.	350 a 664	193 a 821	36 a 978
	Número de casos	53	66	67
	Idem íd. (en %)	79	98	99
	Debía ser (en %)	68	95	99
Madrid	Límites de cantidad lluvia, mm.	340 a 522	249 a 613	158 a 704
	Número de casos	49	64	68
	Idem íd. (en %)	73	95	100
	Debía ser (en %)	68	95	99
Murcia	Límites de cantidad lluvia, mm.	194 a 438	72 a 560	0 a 682
	Número de casos	50	63	67
	Idem íd. (en %)	74	93	99
	Debía ser (en %)	68	95	99
Alicante	Límites de cantidad lluvia, mm.	197 a 451	70 a 578	0 a 705
	Número de casos	41	65	68
	Idem íd. (en %)	61	96	100
	Debía ser (en %)	68	95	99
San Fernando (Cádiz)..	Límites de cantidad lluvia, mm.	421 a 747	258 a 940	95 a 1.073
	Número de casos	47	66	67
	Idem íd. (en %)	70	98	99
	Debía ser (en %)	68	95	99

Basándonos en estos valores hemos formado el Cuadro II de distribución de lluvia alrededor del valor medio de cada estación.

En La Coruña, por ejemplo, la lluvia media anual es de 830 mm. en el período de sesenta y ocho años (1881-1948) estudiado. El valor de σ resulta 222 mm. Entonces entre $M - \sigma$ y $M + \sigma$ hay cuarenta y un años que tienen lluvias dentro de ese intervalo, es decir, el 61 por 100 de los sesenta y ocho años. Según las leyes de la estadística matemática debía haber el 68 por 100 de casos dentro de ese intervalo; luego las lluvias anuales de La Coruña **no** están muy agrupadas, apiñadas, alrededor de su valor medio. Ahora, consideremos las lluvias anuales que hay entre $M - 2\sigma$ ($830 - 2 \times 222 = 386$ mm.) y $M + 2\sigma$ ($830 + 2 \times 222 = 1.074$ mm.), y resulta que hay sesenta y seis años que están dentro de ese intervalo, o sea el 98 por 100 de los sesenta y ocho años. Como no debía haber sino el 95 por 100, resulta que las lluvias están más extendidas, en realidad, que lo normal. Finalmente, entre $M - 3\sigma$ ($830 - 3 \times 222 = 164$ mm.) y $M + 3\sigma$ ($830 + 3 \times 222 = 1.496$ mm.) está la totalidad de casos, es decir, el 100 por 100; en realidad, debía haber el 99 por 100; luego las lluvias anuales se extienden hacia los límites extremos más de lo debido. Conclusión de todo ello es la siguiente: **En el período de sesenta y ocho años que va de 1881 a 1948, las lluvias anuales de La Coruña no han estado concentradas, agrupadas, alrededor del valor medio de las mismas, sino algo dispersas hacia los límites.** Esto mismo lo indica el **gráfico adjunto**, que da las frecuencias de la distribución de las lluvias a derecha e izquierda de su media aritmética. Pero nótese que el gráfico da las frecuencias no como las hemos calculado antes entre $M - \sigma$ y $M + \sigma$, etc., sino entre M y $M - \sigma$, entre M y $M + \sigma$; luego entre $M - \sigma$ y $M - 2\sigma$; entre $M + \sigma$ y $M + 2\sigma$, etcétera.



Distribución de las lluvias alrededor de la media aritmética (M) del período 1881 a 1948.

De análoga manera deduciríamos del cuadro que San Sebastián tiene sus lluvias **poco** dispersas (pues el grupo $M - \sigma$ a $M + \sigma$ da el 80 por 100, mucho mayor que lo normal, que es 68 por 100); que Valladolid tiene una distribución **casi normal**; Soria **poco** dispersa; Barcelona **poco** dispersa; Badajoz **poco** dispersa; lo mismo Madrid, Murcia y San Fernando (Cádiz); en cambio, resultan algo dispersas las de Alicante.

De esperar es que el método expuesto vaya siendo aplicado por muchos observadores a las series de datos de sus respectivas estaciones.

J. M. L.