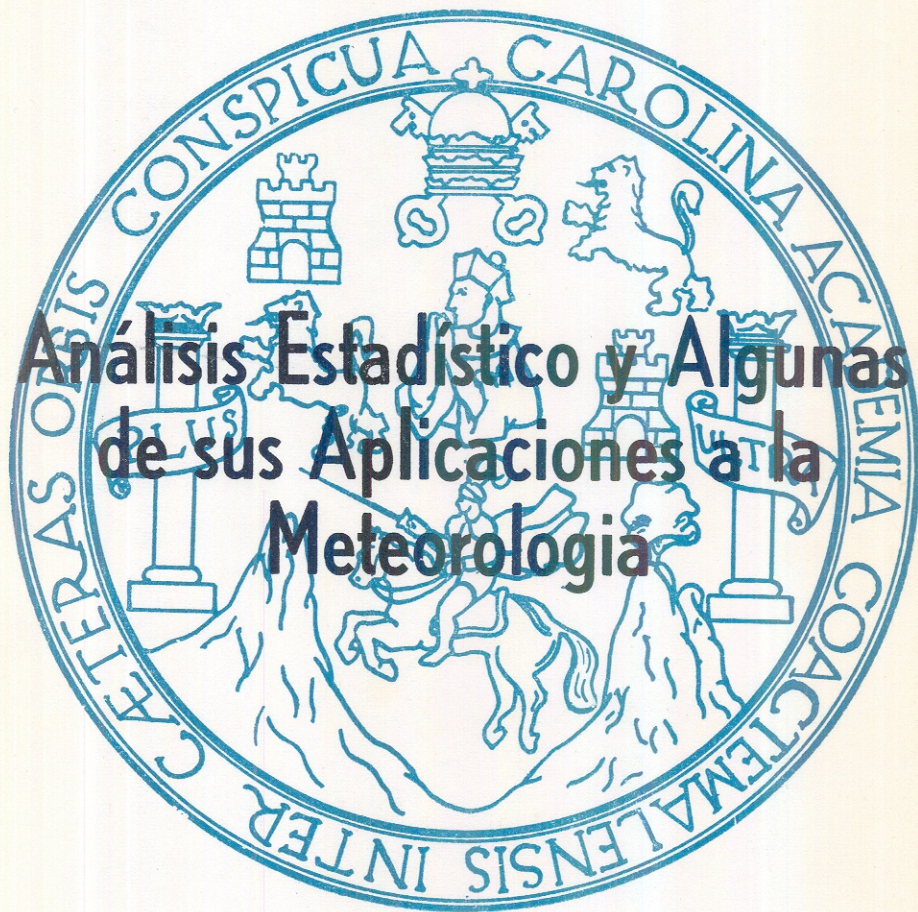


ARTURO ZEA CORDON



**Análisis Estadístico y Algunas
de sus Aplicaciones a la
Meteorología**

BIBLIOTECA CENTRAL-USAC
DEPOSITO LEGAL
PROHIBIDO EL PRESTAMO EXTERNO

GUATEMALA, FEBRERO DE 1968

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

08
T(31)
C 2.

FACULTAD DE INGENIERIA

ANALISIS ESTADISTICO Y ALGUNAS DE SUS
APLICACIONES A LA METEOROLOGIA

T E S I S

PRESENTADA A LA JUNTA DIRECTIVA

DE LA

FACULTAD DE INGENIERIA

DE LA

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

P O R

ARTURO ZEA CORDON

AL CONFERIRSELE EL TITULO DE

INGENIERO CIVIL

CIUDAD DE GUATEMALA

Febrero de 1968

JUNTA DIRECTIVA

DE LA

FACULTAD DE INGENIERIA

DE LA

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

DECANO:	Ing. Amando Vides T.
VOCAL 1º.	Ing. Otto E. Becker M.
VOCAL 2º.	Ing. Francisco UBIETO B.
VOCAL 3º.	Ing. Leonel Pinot L.
VOCAL 4º.	Br. Rolando Rousselin S.
VOCAL 5º.	Br. Javier Godoy A.
SECRETARIO:	Ing. José A. Massanet P.

TRIBUNAL QUE PRACTICO EL EXAMEN
GENERAL PRIVADO

DECANO:	Ing. Amando Vides T.
VOCAL 3º	Ing. Leonel Pinot L.
EXAMINADOR:	Ing. Joaquin Lottmann E.
EXAMINADOR:	Ing. Hugo Quan Ma.
SECRETARIO:	Ing. José A. Massanet P.

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con lo establecido por la ley de la Universidad de San Carlos de Guatemala, tengo el honor de someter a vuestra consideración mi trabajo de tesis, titulado

ANALISIS ESTADISTICO Y ALGUNAS
DE SUS APLICACIONES A LA METEOROLOGIA

tema que me fué asignado por la Honorable Junta Directiva de la Facultad de Ingeniería.

.....DEDICO ESTE ACTO:

A mi madre Rosario Cordón v. de Zea

A mi hermana María del Rosario Zea Cordón

como un reconocimiento a sus inquebrantables empeños por ayudarme.

A la Facultad de Ingeniería de la Universidad de San Carlos.

Agradezco la ayuda prestada en la realización
ción de este trabajo a:

Sr. José Vassaux P.

Sr. Claudio Urrutia E.

Sr. Ernesto López G.

Sr. Raul Villalobos L.

Personal de la Biblioteca del ICAITI.

CONTENIDO

	Pág.
PARTE I <u>INTRODUCCION</u>	1
PARTE II <u>ANALISIS ESTADISTICO</u>	2
II.1 <u>Conceptos Matemáticos</u>	2
II.1.1 Espacio Muestral	2
II.1.2 Probabilidad	3
II.1.3 Probabilidad Condicional	4
II.1.4 Independencia	5
II.1.5 Analisis Combinatorio	9
II.1.6 Variable Aleatoria	12
II.1.7 La distribución binomial	13
II.1.8 Funciones Continuas	14
II.1.9 La distribución Normal	18
II.1.10 Concepto de Expectativa	19
II.1.11 El concepto de Variancia	24
II.1.12 Propiedades de la distribu- ción normal estandarizada	28
II.1.13 Probabilidad de la intersec- ción de dos variables alea- torias	31
II.1.14 Covariancia y coeficiente - de correlación	33
II.1.15 Variancia y expectativa de una media aritmética	34
II.2 <u>Conceptos Estadísticos</u>	36
II.2.1 Teoría del muestreo	36
II.2.2 Estimación puntual y estima- ción de intervalos	36
II.2.3 El método de máxima verosi- militud	39
II.2.4 Hipótesis Estadísticas	45
II.2.5 Un criterio para escoger en tre pruebas alternas	49
	xi

	Pág.	
II.2.6	Prueba de una observación proveniente de una pobla- ción normal con variancia conocida	52
II.2.7	Hipótesis compuestas	59
II.2.8	Otra prueba de un grupo de observaciones provenientes de una población normal con variancia conocida	62
II.2.9	Prueba de un grupo de ob- servaciones provenientes de una población normal con variancia no conocida	68
II.2.10	Comparación de dos medias aritméticas	72
II.2.11	El valor P	73
II.2.12	Límites de Confianza (ó confidencia)	74
PARTE III	APLICACION A LA METEOROLOGIA	79
III.1	<u>Introducción</u>	79
III.2	<u>Información en Meteorología</u>	79
III.3	<u>Colección, concentración y es- crutinio de la información Me- teorológica</u>	80
III.4	<u>Uso de parámetros como cantida- des representativas</u>	88
III.4.1	Regresión y correlación	90
III.4.2	Asociación y contingencia	92
III.5	<u>La distribución de Poisson</u>	93
III.6	<u>La distribución gamma y chi- cuadrado</u>	94

		Pág.
III.7	<u>La distribución "t" ó de Student</u>	100
III.8	<u>La distribución de Fisher ó distribución F</u>	103
III.9	<u>Series de tiempo</u>	103
III.10	<u>Persistencia</u>	105
III.11	<u>Interpretación de datos estadísticos en Meteorología</u>	107
III.11.1	Pruebas respecto a la "bondad" del ajuste	107
III.11.2	Análisis estadístico de temperaturas de la ciudad de Guatemala para un mes dado	108
III.11.3	Análisis de Variancia	113
III.12	<u>Los pronósticos de tiempo</u>	121
III.12.1	Predictores y predictandos	121
III.12.2	Métodos de regresión gráfica	122
III.12.3	Método Residual	130
III.12.4	Método de discriminación gráfica	132
III.12.5	Métodos utilizados en Guatemala	134
III.13	<u>Resumen</u>	138
III.13.1	La temperatura del aire	139
III.13.2	La precipitación	140
III.13.3	El viento	142
III.13.4	La evaporación	144
III.14	<u>Recomendaciones</u>	149

PARTE IV APENDICE Y BIBLIOGRAFIA

Tabla I: Ordenadas de la normal	iv-1
Tabla II: Areas de la normal	iv-2
Tabla III: Distribución "t"	iv-3
Tabla IV: Chi-Cuadrado	iv-4
Tabla V: Distribución de Fisher	iv-10
Tabla VI: Valores de $e^{-\lambda}$	iv-13
Tabla VII: Función Gamma	iv-14
Bibliografía	iv-16

Los fenómenos Meteorológicos y su predicción dentro de períodos de tiempo finitos han atraído la atención de los estudiosos en todas las épocas de la historia de la humanidad. Antiguamente se hacía uso de experiencias pasadas para determinar hechos futuros y esto, en cierto modo, constituyó un método estadístico para predicción del tiempo. Actualmente con las técnicas modernas de la Estadística, se trata de desarrollar, en base a información pasada, métodos que permitan encontrar las relaciones entre variables y que, asimismo, den lugar a sacar conclusiones a partir de ellas.

La Estadística moderna puede ser usada exitosamente en la elaboración de pronósticos de tiempo dado el carácter aleatorio de los fenómenos Meteorológicos; por tanto, una buena base estadística asegura la comprensión de sus aplicaciones en el campo de la Meteorología.

El presente trabajo no tiene la pretensión de ser un manual en donde la persona encuentre todo lo relacionado con el tópico; tampoco se ha tratado de hacer de él un anuario o recopilación de datos Meteorológicos existentes en Guatemala. El verdadero carácter que se ha intentado imprimirle ha sido el de una modesta guía que permita a los interesados en la materia introducirse sin mayor dificultad en el nuevo campo de la Meteorología Estadística.

II.1 CONCEPTOS MATEMATICOS

II.1.1 Espacio Muestral

Se denomina espacio muestral al conjunto de todos los posibles resultados de un experimento. Por experimento se entiende cualquier procedimiento cuyos resultados están sujetos a una serie de factores. Cada uno de los posibles resultados se denomina punto muestral. El espacio muestral puede ser:

- a) Discreto: si el número total de puntos muestrales es finito.
- b) Continuo: si el número de puntos muestrales es infinito.

El conjunto de puntos muestrales con propiedades específicas se denomina evento representado simbólicamente con la letra E ; el conjunto de puntos muestrales que no pertenecen a E pero si al espacio muestral se denomina complemento de E y se representa por " \bar{E} ".

Si dos eventos E_1 y E_2 dentro de un espacio muestral no tienen puntos muestrales comunes se denominan mutuamente exclusivos y el evento que resulta de la intersección de los dos eventos (representado $E_1 \cap E_2$) se llama evento nulo ó vacío ya que no contiene puntos muestrales.

Cada punto muestral A_i lleva asociado un número llamado la probabilidad de A_i que se expresa $P(A_i)$. Es un número positivo tal que debe cumplir la condición de

$$P(A_1) + P(A_2) + \dots = 1 \quad (2.1)$$

hecha la suma para todo el espacio muestral. Aún mas se puede suponer que la probabilidad de un evento E es la suma de las probabilidades de los puntos muestrales que constituyen el evento.

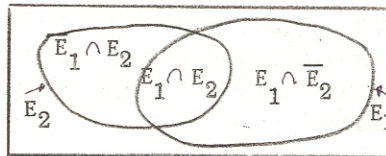
Si dos eventos E_1 y E_2 son mutuamente exclusivos, entonces la probabilidad del evento que constituye la suma ó unión de E_1 y E_2 (expresado $E_1 \cup E_2$) es

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) \quad \begin{array}{l} \text{Eventos} \\ \text{mutuamente} \\ \text{exclusivos} \end{array} \quad (2.2)$$

Para dos eventos que no sean mutuamente exclusivos, por ejemplo los eventos E_1 y E_2 de la fig. 1a vemos que la región E_2 puede descomponerse en dos eventos mutuamente exclusivos $\bar{E}_1 \cap E_2$ y $E_1 \cap E_2$ y la región E_1 puede también descomponerse en dos eventos mutuamente exclusivos $E_1 \cap \bar{E}_2$ y $E_1 \cap E_2$.

FIGURA 1a

Espacio Muestral



$$P(E_1 \cup E_2) = P(E_2 \cap \bar{E}_1) + P(E_1 \cap \bar{E}_2) + P(E_1 \cap E_2) \quad (2.3)$$

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \quad (2.4)$$

II.1.3 Probabilidad Condicional

Para que un evento se realice es suficiente obtener un punto muestral del evento.

Si se tienen dos eventos E_1 y E_2 como los de la fig. 1 y se realiza el evento E_1 , que puede decirse respecto a la probabilidad de E_2 ?. En otras palabras si hemos obtenido un punto muestral de E_1 , cual es la probabilidad de que ese punto muestral corresponda a E_2 ? (esto naturalmente para dos eventos dentro de un mismo espacio muestral). Decimos cual es la probabilidad de E_2 una vez realizado E_1 y lo expresamos así:

$$P(E_2/E_1)$$

La probabilidad de E_1 es, de acuerdo al artículo anterior:

$$P(E_1) = \frac{\text{suma de las probab. de los puntos muestrales de } E_1}{\text{suma de las probab. de todos los ptos. en el esp. muestral.}}$$

Sin embargo, si sabemos que E_1 se ha realizado, sabemos que el resultado del experimento estará únicamente dentro de los puntos muestrales de E_1 , luego es razonable pensar que la probabilidad de E_2 una vez que E_1 se ha realizado es

$$P(E_2/E_1) = \frac{\text{suma de las probabilidades de los } 5 \text{ puntos muestrales de } E_1 \cap E_2}{\text{suma de las probabilidades de los puntos muestrales de } E_1}$$

ó poniéndolo en la forma

$$P(E_2/E_1) = \frac{\text{suma de las probabilidades de los puntos muestrales de } E_1 \cap E_2}{\text{suma de las probabili. de todos los puntos en el esp. muestral}} = \frac{\text{suma de las probabili. de los puntos muestrales de } E_1 \cap E_2}{\text{suma de las probabili. de todos los puntos en el esp. muestral}}$$

es decir

$$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} \quad (2.5)$$

ó

$$P(E_1 \cap E_2) = P(E_1) P(E_2/E_1) \quad (2.6)$$

e igualmente

$$P(E_1/E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)} \quad (2.7)$$

ó

$$P(E_1 \cap E_2) = P(E_1/E_2) P(E_2) \quad (2.8)$$

II.1.4 Independencia

Si la probabilidad de un evento E_1 es la misma, se realice ó no otro evento E_2 , se dice

que los eventos E_1 y E_2 son independientes. Esta condición puede expresarse así

$$P(E_1/E_2) = P(E_1/\bar{E}_2) \quad (2.9)$$

igualdad evidente para dos eventos realizados en diferentes espacios muestrales. Para muchos experimentos, sin embargo, es importante conocer la probabilidad de que ambos eventos se realicen

$$\text{de (2.7): } P(E_1/E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)} \quad (2.10)$$

sustituyendo E_2 por \bar{E}_2

$$P(E_1/\bar{E}_2) = \frac{P(E_1 \cap \bar{E}_2)}{P(\bar{E}_2)} \quad (2.11)$$

de (2.9)

$$\frac{P(E_1 \cap E_2)}{P(E_2)} = \frac{P(E_1 \cap \bar{E}_2)}{P(\bar{E}_2)}$$

ó

$$P(E_1 \cap E_2) P(\bar{E}_2) = P(E_1 \cap \bar{E}_2) P(E_2)$$

pero

$$P(\bar{E}_2) = 1 - P(E_2)$$

luego

$$P(E_1 \cap E_2) = P(E_1 \cap E_2) P(E_2) \neq P(E_1 \cap \bar{E}_2) P(E_2)$$

$$P(E_1 \cap E_2) = P(E_2) \left[\frac{P(E_1 \cap E_2)}{P(E_1 \cap \bar{E}_2)} \right]$$

pero

$$P(E_1 \cap E_2) \neq P(E_1 \cap \bar{E}_2) = P(E_1)$$

y finalmente

$$P(E_1 \cap E_2) = P(E_2) \cdot P(E_1) \quad \text{para eventos independientes} \quad (2.12)$$

Conviene hacer notar lo siguiente: el signo de intersección (\cap) representa la condición de que ambos eventos se lleven a cabo y para dos eventos que pueden ó no realizarse, el número de modos en que se pueden ó no efectuar es en forma tabulada

CUADRO I

Modos	Probabilidad	Significación
$E_1 \cap E_2$	$P(E_1) \cdot P(E_2)$	E_1 y E_2 se realicen
$E_1 \cap \bar{E}_2$	$P(E_1) \cdot P(\bar{E}_2)$	E_1 se realice y E_2 no
$\bar{E}_1 \cap E_2$	$P(\bar{E}_1) \cdot P(E_2)$	E_1 no se realice y E_2 sí
$\bar{E}_1 \cap \bar{E}_2$	$P(\bar{E}_1) \cdot P(\bar{E}_2)$	E_1 no se realice y E_2 tampoco
	$\sum \text{Prob} = 1$	co

Por otro lado el signo de unión (\cup) significa que puede realizarse solo un evento, y el otro no ó pueden realizarse ambos eventos.

Las formas en que pueden unirse los eventos ó sus intersecciones nuevamente es en forma tabulada

C U A D R O II

Eventos	Modos en que se cumple	Probabilidad
$E_1 \cup E_2$	$E_1 \cup E_2$	$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
$E_1 \cap E_2$	$E_1 \cap E_2$	$P(E_1 \cap E_2)$
$\overline{E_1} \cup \overline{E_2}$	$\overline{E_1} \cup \overline{E_2}$	$P(\overline{E_1} \cup \overline{E_2}) = 1 - P(E_1 \cap E_2)$
$\overline{E_1} \cap \overline{E_2}$	$\overline{E_1} \cap \overline{E_2}$	$P(\overline{E_1} \cap \overline{E_2}) = 1 - P(E_1 \cup E_2)$
$E_1 \cup \overline{E_2}$	$E_1 \cup \overline{E_2}$	$P(E_1 \cup \overline{E_2}) = P(E_1) + P(\overline{E_2}) - P(E_1 \cap \overline{E_2})$
$\overline{E_1} \cup E_2$	$\overline{E_1} \cup E_2$	$P(\overline{E_1} \cup E_2) = P(\overline{E_1}) + P(E_2) - P(\overline{E_1} \cap E_2)$
$E_1 \cup E_2 \cup \overline{E_1} \cup \overline{E_2}$	siempre se realice	$P(E_1 \cup E_2 \cup \overline{E_1} \cup \overline{E_2}) = 1$
$E_1 \cup \overline{E_2} \cup \overline{E_1} \cup E_2$	siempre se realice	$P(E_1 \cup \overline{E_2} \cup \overline{E_1} \cup E_2) = 1$
$E_1 \cup \overline{E_2} \cup \overline{E_1} \cup \overline{E_2}$	E_1 y $\overline{E_2}$ se realicen ó no se realicen	$P(E_1 \cup \overline{E_2} \cup \overline{E_1} \cup \overline{E_2}) = P(E_1 \cup \overline{E_2})$
$\overline{E_1} \cup E_2 \cup \overline{E_1} \cup E_2$	un evento se realice y otro no	$P(\overline{E_1} \cup E_2 \cup \overline{E_1} \cup E_2) = P(E_2)$
$\overline{E_1} \cup \overline{E_2} \cup E_1 \cup E_2$	E_2 nunca se realice	$P(\overline{E_1} \cup \overline{E_2} \cup E_1 \cup E_2) = P(\overline{E_1} \cup E_1) = P(E_1)$
$\overline{E_1} \cup \overline{E_2} \cup \overline{E_1} \cup \overline{E_2}$	E_1 nunca se realice	$P(\overline{E_1} \cup \overline{E_2} \cup \overline{E_1} \cup \overline{E_2}) = P(\overline{E_2})$

II.1.5 Análisis Combinatorio

El análisis combinatorio tiene por objeto la ordenación de las cosas ó elementos. Inicialmente una agrupación de cosas se puede distinguir de otras agrupaciones:

- 1) por el orden en que están colocados los elementos.
- 2) por la naturaleza de alguno de ellos.

Si atendemos a los dos caracteres anteriores juntamente ó sólo el primero de ellos obtendremos las llamadas permutaciones. Cuando se atiende al segundo carácter solamente se tendrán las combinaciones. Supongamos que tenemos "m" objetos diferentes; estos objetos los podemos arreglar en grupos de 1, 2, ... m objetos y para generalizar diremos que vamos a arreglar los "m" objetos en grupos de n objetos colocados en línea recta, es decir uno atrás del otro. En la primera posición podemos colocar cualquiera de los m objetos, en la segunda posición podemos colocar cualquiera de (m-1) objetos puesto que hay uno colocado en la primera posición, en la tercera posición podemos colocar cualquiera de (m-2) objetos y en la posición n (la última) podemos colocar cualquiera de $\lfloor m-(n-1) \rfloor$ objetos. Si atendemos a la generalización del principio fundamental que dice:

"Si un evento puede ocurrir de h modos diferentes, y si una vez que ha ocurrido en uno de esos modos, otro evento puede ocurrir de k modos diferentes, entonces ambos eventos pueden ocurrir, en el orden especificado, de $h \cdot k$ diferentes modos".

Entonces el número total de modos en que podemos agrupar los m objetos en grupos de n objetos, llamado el número de permutaciones de m objetos en grupos de n objetos es

$$P_n^m = m(m-1)\dots \lfloor m-(n-1) \rfloor \quad (2.13)$$

Si se define el factorial de m como $m! = m(m-1)\dots 3 \cdot 2 \cdot 1$ tendremos

$$\begin{aligned} P_n^m &= m(m-1)\dots (m-n+1) \frac{(m-n) \cdot (m-n-1) \cdot \dots \cdot 2 \cdot 1}{(m-n) \cdot (m-n-1) \cdot \dots \cdot 2 \cdot 1} \\ &= \frac{m!}{(m-n)!} \end{aligned} \quad (2.14)$$

en el caso especial de que $n = m$

$$P_m^m = m(m-1)\dots 2 \cdot 1 = m! = \frac{m!}{0!} \quad (2.15)$$

y para que se cumpla la igualdad $0! = 1$

Si los m objetos consisten en m_1 objetos iguales, de tipo T_1 , m_2 objetos iguales de tipo T_2 , ..., m_k objetos iguales de tipo T_k tales que $\sum_{i=1}^k m_i = m$, el número de permutaciones de m objetos en grupos de m objetos es

$$\frac{m!}{m_1! \cdot m_2! \cdot \dots \cdot m_k!} \quad (2.16)$$

Si se le asignan a p elementos de una permutación (P_m^m) lugares determinados, el número de permutaciones que pueden formarse es P_{m-p}^{m-p} . Si las " p " cosas especificadas pueden a su vez permutarse entre sí, entonces el número es P_{m-p}^{m-p} multiplicado por $p!$

Por otro lado, el número de permutaciones de " m " cosas en grupos de m , tomados alrededor de un círculo es $(m-1)!$ con tal que se cuenten en un solo sentido. Esto porque como solo hay que tener en cuenta las posiciones relativas resulta, por el modo de formar las permutaciones de m cosas, que al ser colocadas en el círculo se distribuyen en grupos de " m " idénticas.

Se trata ahora de encontrar el número de combinaciones de m objetos en grupos de n objetos, representado C_n^m . Cada combinación de n objetos puede permutarse entre sí de $n!$ modos; luego el producto $C_n^m \cdot n!$ será igual al número de permutaciones de m objetos en grupos de n objetos; es decir

$$C_n^m \cdot n! = P_n^m$$

ó

$$C_n^m = \frac{P_n^m}{n!} \quad (2.17)$$

Utilizando ecuación 2.14 obtendremos

$$C_n^m = \frac{m!}{n! (m-n)!} \quad (2.17a)$$

Si se exige que p elementos se encuentren en todas las combinaciones de m elementos tomados en grupos de n elementos, el número de ellos será C_{n-p}^{m-p} . Esto porque teniendo p elementos fijos quedan $n-p$ en cada combinación que pueden ser ocupados por las $m-p$ restantes.

El número de combinaciones de " m " objetos tomados en grupos de " n " objetos que no encierren ninguno de los " p " elementos es C_{n-p}^{m-p} porque separando los p objetos quedan $(m-p)$; pero como las agrupaciones son en grupos de n objetos se tendrán que combinar los $(m-p)$ en grupos de n elementos.

II.1.6 Variable Aleatoria

Una variable aleatoria (llamada también - estocástica) es una función que, para cada punto en el espacio muestral, adquiere:

- a) un valor definido
- b) un valor de probabilidad también definido

Es usual para una variable aleatoria, graficar los valores obtenidos contra las respectivas probabilidades. También es común graficar una distribución de probabilidades acumuladas cuya ordenada es la probabilidad de que la variable sea menor ó igual al valor obtenido.

II.1.7 La distribución binomial

Se trata de calcular la probabilidad de que un evento A se realice exactamente x veces en "n" pruebas independientes siendo $P(A) = \theta$ una constante.

Supongamos que se realizan 2 pruebas únicamente con lo que A puede efectuarse 0, 1, ó 2 veces; si representamos por $P_2(x)$ como la probabilidad de que un evento se realice exactamente x veces en n pruebas tendremos

$$P_2(0) = P(\bar{A} \cap \bar{A}) = P(\bar{A}) \cdot P(\bar{A}) = (1-\theta)^2$$

$$P_2(1) = P(A \cap \bar{A} \cup \bar{A} \cap A) = P(A) \cdot P(\bar{A}) + P(\bar{A}) \cdot P(A) \\ = 2\theta(1-\theta)$$

$$P_2(2) = P(A \cap A) = P(A) \cdot P(A) = \theta^2$$

En cualquiera de los casos puede notarse que la probabilidad se encuentra multiplicando $P(A)$ "x" veces (θ^x), por $P(\bar{A})$ (n-x) veces ($(1-\theta)^{n-x}$) y este producto multiplicado por el número de combinaciones de n objetos en grupos de x objetos (ver pag 11), luego generalizando tendremos:

$$P_n(x) = C_x^n \cdot \theta^x (1-\theta)^{n-x} \quad (2.18)$$

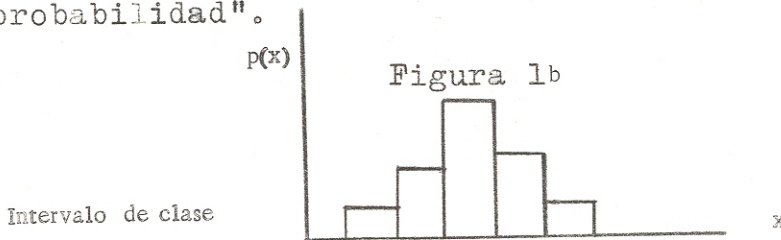
$$P_n(x) = \frac{n!}{x! (n-x)!} \theta^x (1-\theta)^{n-x} \quad (2.19)$$

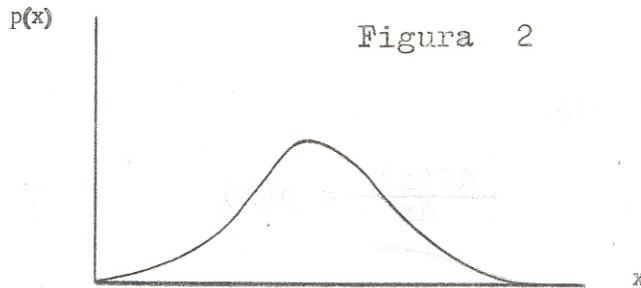
II.1.8 Funciones Continuas

Una variable continua es aquella que puede tener cualquier valor intermedio entre dos valores dados; una variable discreta es aquella que, generalmente, solo puede tener un valor intermedio entre dos valores dados. En general el proceso de "contar" da lugar a variables discretas y el proceso de "medir" produce teóricamente, variables continuas.

La diferencia entre los dos valores extremos que puede tener una variable se llama intervalo de clase y a medida que el intervalo de clase se hace más pequeño la variable tiende a ser continua.

Una forma más conveniente de representar una variable repetitiva es graficar en las abscisas los valores que adquiere la variable y en las ordenadas graficar, ya no la probabilidad para cada valor sino la "densidad" de probabilidad que no es más que la probabilidad dividida entre el intervalo de clase. Para variables discretas el gráfico obtenido se denomina histograma y tiene la forma de la figura 1 y para las continuas tiene una forma parecida a la curva 2 llamada "función de densidad de probabilidad".





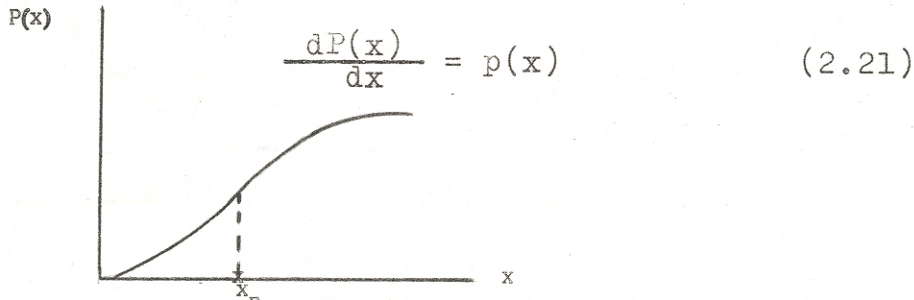
Ambas figuras tienen la característica de que el área entre dos valores de x es la probabilidad de que x esté entre esos límites. Lógicamente, el área total de la figura es igual a uno.

En el caso de variables continuas el valor de

$$P(x \leq x_n) = \int_{-\infty}^{x_n} p(x) dx \quad (2.20)$$

Para diferentes valores de x_n da lugar a una distribución acumulada de probabilidad como la de la figura 2ª con las características siguientes:

- a) la pendiente nunca puede ser negativa puesto que la densidad de probabilidad de la que está derivada nunca es negativa.
- b) la pendiente de la curva en x_n es igual a la densidad de probabilidad para x_n ; es decir



Cualquier función $y = f(x)$ tal que "y" nunca sea negativa y que $\int_{-\infty}^{\infty} y dx = 1$ puede considerarse como una función de densidad de probabilidad, siendo "y" la densidad de probabilidad y "x" los valores que adquiere la variable. Luego se estudiarán algunas funciones especiales.

Algunas veces se tiene una función $y = f(x)$, y se desea conocer la distribución de densidad de probabilidad de una de las variables conociendo la de la otra. Sea $p(x)$ la densidad de probabilidad de "x" y $q(y)$ la densidad de probabilidad de "y". Para variables discretas cuando "x" toma un valor particular x_i , "y" toma simultáneamente un valor definido y_i . Siendo "x" una variable aleatoria, "y" es también una variable aleatoria con densidad de probabilidad

$$q[f(x)] = q(y) = p(x) \quad (2.22)$$

Para variables continuas consideremos las

figuras 3, 4 y 5:

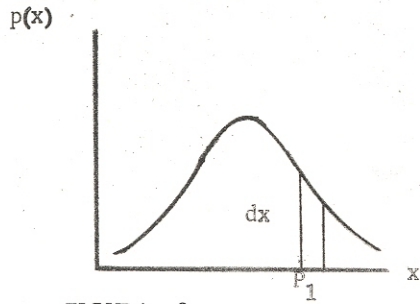


FIGURA 3

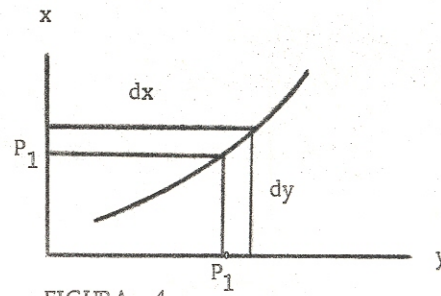


FIGURA 4

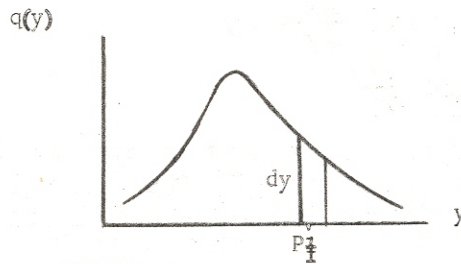


FIGURA 5

Para un valor de $x = P_1$ y un intervalo de clase dx , la probabilidad es

$$dP(x) = p(x).dx \quad (2.23)$$

Simultáneamente para $x = P_1$, "y" adquiere un valor de P_1 con probabilidad

$$dP(y) = q(y).dy \quad (2.24)$$

y aproximadamente

$$p(x).dx = q(y).dy \quad (2.25)$$

pero

$$dx = (x \int dx) - x$$

$$x \int dx = x \int dx/dy \cdot dy$$

ó

$$dx = \frac{dx}{dy} \cdot dy \quad (2.26)$$

sustituyendo 2.26 en ecuación 2.25

$$q(y).dy = p(x) \frac{dx}{dy} \cdot dy$$

$$q(y) = p(x) \frac{dx}{dy} \quad (2.27)$$

Hay que hacer notar que la ecuación 2.27 es aproximadamente cierta para funciones de x crecientes para valores crecientes de x .

II.1.9 La distribución Normal

La función de densidad de probabilidad para la distribución normal es

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp \left[-(x-\mu)^2 / 2\sigma^2 \right] \quad (2.28)$$

Donde \exp . significa exponente del número natural e , y " y " es una función diferente para cada par de valores de σ y μ . Todas estas funciones pueden, no obstante, estandarizarse mediante la transformación

$$y = f(x) = (x-\mu)/\sigma \quad (2.29)$$

la función inversa

$$x = g(y) = \sigma y + \mu$$

cuya derivada es

$$\frac{dx}{dy} = \sigma$$

sustituyendo valores en ecuación 2.28

$$q(y) = \frac{1}{\sqrt{2\pi}} \cdot \exp -y^2/2 \quad (2.30)$$

ó

$$\phi(u) = \frac{1}{\sqrt{2\pi}} \cdot \exp -u^2/2 \quad (2.31)$$

y la forma acumulada

$$\Phi(u) = \int_{-\infty}^u \phi(u) \cdot du \quad (2.32)$$

II.1.10 Concepto de Expectativa

La expectativa de una variable aleatoria se define como

$$E(x) = \sum_i x_i \cdot P(x_i) \text{ para variables discretas} \quad (2.33)$$

$$E(x) = \int_{-\infty}^{\infty} x \cdot p(x) \cdot dx \quad \text{para variables continuas} \quad (2.34)$$

y corresponde con la idea intuitiva del valor medio. La expectativa de la función lineal ($a + bx$), siendo x una variable aleatoria continua y "a" y "b" constantes es

$$\begin{aligned} E(a + bx) &= \int_{-\infty}^{\infty} (a + bx) \cdot p(x) \cdot dx \\ &= a \int_{-\infty}^{\infty} p(x) \cdot dx + b \int_{-\infty}^{\infty} x \cdot p(x) \cdot dx \\ &= a + b \cdot E(x) \end{aligned} \quad (2.35)$$

Si $b = 0$ la expectativa de una constante es la constante misma; si $a = 0$ entonces $E(bx) = bE(x)$. También puesto que el proceso de hallar $E(x)$ dá lugar a una constante, la expectativa de una expectativa es la expectativa misma. Es decir

$$E[E(x)] = E(x) \quad (2.36)$$

Aún mas, la expectativa de la diferencia de una variable aleatoria respecto a su expectativa es

$$\begin{aligned} E[x - E(x)] &= E(x) - E[E(x)] \\ &= E(x) - E(x) = 0 \end{aligned} \quad (2.37)$$

es decir la expectativa de las desviaciones respecto a ella es cero.

Necesitamos conocer también la expectativa de una función de una variable aleatoria. Sea $y = f(x)$, $p(x)$ la densidad de probabilidad de "x" y $q(y)$ la densidad de probabilidad de "y". De ecuación 2.34

$$E [y] = \int_{-\infty}^{\infty} y \cdot q(y) \cdot dy \quad (2.38)$$

sustituyendo 2.27 en ecuación 2.38

$$\begin{aligned} E(y) &= \int_{-\infty}^{\infty} f(x) \cdot p(x) \cdot \frac{dx}{dy} \cdot dy \\ &= \int_{-\infty}^{\infty} f(x) \cdot p(x) \cdot dx \end{aligned} \quad (2.39)$$

y para variables discretas

$$E(y) = \sum_i f(x_i) \cdot P(x_i) \quad (2.40)$$

La expectativa de la distribución binomial (2.19) es

$$E(x) = \sum_{x=0}^n x \cdot C_x^n \cdot \theta^x \cdot (1-\theta)^{n-x} \quad (2.41)$$

pero

$$x \cdot C_x^n = \frac{x \cdot n!}{x! (n-x)!} = \frac{(n-1)!}{(x-1)! (n-x)!} \cdot n = C_{x-1}^{n-1} \cdot n$$

luego podemos escribir 2.41 como

$$E(x) = n\theta \sum_{x=1}^n C_{x-1}^{n-1} \cdot \theta^{x-1} \cdot (1-\theta)^{n-x} \quad (2.42)$$

Se ha cambiado el límite inferior de la sumatoria, de 0 en 2.41 a 1 en 2.42 ya que para $p_{n-1}(x-1) = p_{n-1}(0)$ tendremos $x=1$. La ecuación 2.42 podría ponerse en mejor forma si hacemos $n=m/1$ y $x=y/1$. En 2.42 x varía entre 1 y n , ó entre 1 y $(m/1)$ así que "y" varía entre "0" y "m", luego

$$E(x) = n\theta \cdot \sum_{y=0}^m C_y^m \cdot \theta^y \cdot (1-\theta)^{m-y} \quad (2.43)$$

La sumatoria de todos los términos de una distribución binomial es igual a uno, entonces

$$E(x) = n\theta \quad (2.44)$$

La expectativa de la distribución normal (2.28) es

$$E(x) = \int_{-\infty}^{\infty} x \cdot p(x) \cdot dx = x \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp. \quad -(x-\mu)^2/2\sigma^2 \cdot dx$$

ó

$$E(x) = \int_{-\infty}^{\infty} [\mu + (x-\mu)] \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp. \quad -(x-\mu)^2/2\sigma^2 \cdot dx \quad (2.45)$$

$$E(x) = \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp \left[-(x-\mu)^2 / 2\sigma^2 \right] \cdot dx$$

$$+ \int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp \left[-(x-\mu)^2 / 2\sigma^2 \right] \cdot dx \quad (2.46)$$

El integral del primer sumando de la derecha es 1. Para evaluar el segundo integral hágase $(x-\mu) = y$

$$\frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} y \cdot \exp \left[-y^2 / 2\sigma^2 \right] \cdot dy = - \frac{\sigma}{\sqrt{2\pi}} \cdot \exp \left[-y^2 / 2\sigma^2 \right] \Big|_{-\infty}^{\infty} = 0 \quad (2.47)$$

luego

$$E(x) = \mu \quad (2.48)$$

Para la distribución normal estandarizada obtenida por la sustitución $u=(x-\mu)/\sigma$ tendremos, usando 2.35

$$E(u) = E\left(\frac{x-\mu}{\sigma}\right) = \frac{E(x) - \mu}{\sigma} = \frac{\mu - \mu}{\sigma} = 0$$

Es decir la distribución normal estandarizada tiene expectativa cero.

La expectativa ó valor esperado no es necesariamente el valor mas frecuente (conocido como la moda de la distribución); de hecho en distribuciones discretas el valor medio puede no ocurrir. Un tercer valor de medidas de tendencia central es la mediana que es aquel va-

lor de x que divide el área de la distribución en dos partes iguales.

II.1.11 El Concepto de Variancia

La variancia, como medida de dispersión se define como la expectativa del cuadrado de la desviación de una variable aleatoria x respecto a su expectativa. Es decir

$$V(x) = E(\lfloor x - E(x) \rfloor^2) \quad (2.49)$$

Algunas veces es mas conveniente, para de terminar valores de variancias, utilizar las siguientes identidades:

$$\begin{aligned} V(x) &= E(\lfloor x - E(x) \rfloor^2) = E \lfloor x^2 - 2 \cdot x \cdot E(x) + \\ &\quad (E(x))^2 \rfloor \\ &= E(x^2) - 2E\lfloor x \cdot E(x) \rfloor + E\lfloor (E(x))^2 \rfloor \\ &= E(x^2) - 2\lfloor E(x) \rfloor^2 + \lfloor E(x) \rfloor^2 \\ &= E(x^2) - \lfloor E(x) \rfloor^2 \end{aligned} \quad (2.50)$$

también

$$\begin{aligned} V(x) &= E(x^2) - \lfloor E(x) \rfloor^2 = E(x^2) - E(x) - \\ &\quad \lfloor E(x) \rfloor^2 + E(x) \\ &= E\lfloor x(x-1) \rfloor - E(x)\lfloor E(x)-1 \rfloor \end{aligned} \quad (2.51)$$

La variancia de una función lineal de x por ejemplo (a/bx) es

$$V(a/bx) = E\lfloor (a/bx)^2 \rfloor - \lfloor E(a/bx) \rfloor^2$$

$$\begin{aligned}
&= E[a^2 + 2ab \cdot x + (bx)^2] - [a + b \cdot E(x)]^2 \\
&= E(a^2) + 2ab \cdot E(x) + b^2 \cdot E(x^2) - a^2 - \\
&\quad 2ab \cdot E(x) - b^2 [E(x)]^2 \\
&= b^2 (E(x^2) - [E(x)]^2) = b^2 \cdot V(x) \quad (2.52)
\end{aligned}$$

Es decir que el efecto de añadir una constante "a" a una variable aleatoria deja la variancia igual, pero la variancia de bx es b^2 veces la variancia de x .

Para el cálculo de la variancia de la distribución binomial (2.19) se usará la ecuación 2.51. El valor de $E[x(x-1)]$ se computa considerando $x(x-1)$ como una función de x , luego de 2.40:

$$\begin{aligned}
E[x(x-1)] &= E_n(y) = \sum_i x(x-1) C_x^n \cdot \theta^x \cdot (1-\theta)^{n-x} \\
&= \sum_{x=0} x(x-1) C_x^n \cdot \theta^x \cdot (1-\theta)^{n-x} \quad (2.53)
\end{aligned}$$

pero

$$\begin{aligned}
x(x-1) C_x^n &= x(x-1) \frac{n!}{x! (n-x)!} = n(n-1) \frac{(n-2)!}{(x-2)! (n-x)!} \\
&= n(n-1) C_{x-2}^{n-2}
\end{aligned}$$

luego

$$E(y) = n(n-1) \theta^2 \cdot \sum_{x=2}^n C_{x-2}^{n-2} \cdot \theta^{x-2} \cdot (1-\theta)^{n-x} \quad (2.54)$$

Se ha cambiado el límite inferior de la sumatoria de 2.53 de 0 a 2 en 2.54 ya que cuando

do x es cero ó uno los correspondientes términos de la sumatoria son cero. Al igual que en ecuación 2.43 la ecuación 2.54 puede ponerse en forma mas objetiva si hacemos $x=y/2$ y $n=m/2$, entonces

$$E[y] = n(n-1) \cdot \theta^2 \sum_{y=0}^m C_y^m \cdot \theta^y \cdot (1-\theta)^{m-y} \quad (2.55)$$

La sumatoria de 2.55 es igual a uno, luego $E(y) = n(n-1)\theta^2$. Sustituyendo este valor y el de la igualdad 2.44 en 2.51 obtendremos para la variancia de la distribución binomial

$$V(x) = n(n-1)\theta^2 - n\theta(n\theta-1) = n\theta(1-\theta) \quad (2.56)$$

y si consideramos la variancia de la proporción de sucesos $x/n = h$ tendremos por 2.52

$$V(h) = V\left(\frac{x}{n}\right) = \frac{1}{n^2} \cdot V(x) = \frac{\theta(1-\theta)}{n} \quad (2.57)$$

La variancia de la distribución normal en su forma estandarizada (ecuación 2.31), es, recordando que $E(u) = 0$

$$\begin{aligned} V(u) &= E([u - E(u)]^2) = E(u^2) = \int_{-\infty}^{\infty} u^2 \cdot p(u) \cdot du \\ &= \int_{-\infty}^{\infty} u^2 \frac{1}{\sqrt{2\pi}} \exp -u^2/2 \cdot du \quad (2.58) \\ &\text{haciendo} \end{aligned}$$

$$dy = u \cdot \exp -u^2/2 \cdot du; \quad y = -\exp -u^2/2$$

$$x = u; \quad dx = du$$

e integrando por partes

$$\int x \cdot dy = xy - \int y \cdot dx$$

$$\begin{aligned} \int_{-\infty}^{\infty} u(u \cdot \exp -u^2/2 \cdot du) &= u(-\exp -u^2/2) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \exp -u^2/2 \cdot du \\ &= 0 - \int_{-\infty}^{\infty} \exp -u^2/2 \cdot du \end{aligned}$$

$$\int = \sqrt{2\pi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp -u^2/2 \cdot du = \sqrt{2\pi} \quad (2.59)$$

Puesto que el último integral es precisamente el de la distribución normal estandarizada a todo lo largo de la línea y por tanto igual a uno.

Sustituyendo 2.59 en 2.58, se obtiene la variancia de la distribución normal estandarizada

$$V(u) = 1 \quad (2.60)$$

La variancia para la variable aleatoria "x" en 2.28, despejada de 2.29

$$x = y\sigma + \mu$$

es

$$V(x) = V(y\sigma + \mu) = \sigma^2 \cdot V(y) \quad (2.61)$$

pero "y" (ó "u"), es la variable aleato-

ria para la distribución normal estandarizada, luego

$$V(x) = \sigma^2 \quad (2.62)$$

Así pues, los parámetros μ y σ^2 de la distribución normal (ecuación 2.28) son su expectativa y variancia respectivamente.

La forma estandarizada de cualquier variable aleatoria "x" se define como

$$w = \frac{x - E(x)}{\sqrt{V(x)}} \quad (2.63)$$

luego

$$E(w) = \frac{E(x) - E\left[\frac{x - E(x)}{\sqrt{V(x)}}\right]}{\sqrt{V(x)}} = 0 \quad (2.64)$$

y

$$\begin{aligned} V(w) &= E\left[\left(\frac{x - E(x)}{\sqrt{V(x)}}\right)^2\right] = E(w^2) \\ &= \frac{E\left[\left(x - E(x)\right)^2\right]}{V(x)} = \frac{V(x)}{V(x)} = 1 \end{aligned} \quad (2.65)$$

luego cualquier variable aleatoria puesta en su forma estandarizada tiene expectativa cero y variancia uno.

II.1.12 Propiedades de la distribución normal estandarizada

La distribución normal estandarizada

$$\phi(u) = \frac{1}{\sqrt{2\pi}} \cdot \exp -u^2/2$$

tiene la propiedad de ser simétrica respecto a $u = 0$ ya que $-(-u)^2 = -u^2$ de tal manera que

$$\phi(-u) = \phi(u)$$

además

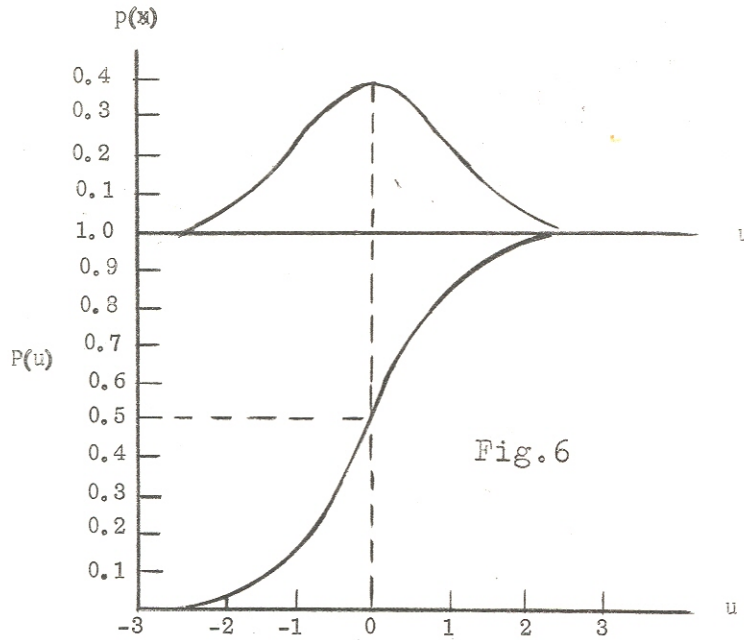
$$\frac{d \phi(u)}{du} = \frac{1}{\sqrt{2\pi}} \cdot \exp -u^2/2 \cdot (-1/2)2u = -u\phi(u) \quad (2.66)$$

y

$$\frac{d^2 \phi(u)}{du^2} = -\phi(u) - (-u)(-u\phi(u)) = (u^2 - 1)\phi(u) \quad (2.67)$$

La fig. 6 es una gráfica de la ecuación 2.31, en la parte superior, y de la ecuación 2.32 en la parte inferior. Valores de las ordenadas de estas gráficas están en tablas I y II del apéndice.

Por las reglas usuales de cálculo obtendremos el valor máximo de $\phi(u)$ para $u=0$ y puntos de inflexión (puntos en que la pendiente deja de aumentar y comienza a disminuir) para $u = 1$ y $u = -1$



Las operaciones de encontrar la probabilidad de que una variable aleatoria sacada de una distribución normal, sea menor que un cierto valor y viceversa son fáciles de efectuar si conocemos la media (μ) y la variancia (σ^2) de la distribución. Simplemente en la distribución acumulada de 2.28

$$P(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp \left(-\frac{(x-\mu)^2}{2\sigma^2} \right) dx \quad (2.68)$$

si sustituimos $u = \frac{x-\mu}{\sigma}$ con lo cual

$$P(x) = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{(x-\mu)/\sigma} \exp \left(-\frac{u^2}{2} \right) \cdot \sigma \cdot du = \Phi \left(\frac{x-\mu}{\sigma} \right) \quad (2.69)$$

Por ejemplo si se dan $\mu = 100$, $\sigma = 10$ y se quiere hallar la probabilidad de que una observación tomada al azar de la distribución sea menor que 120, tendremos

$$P(x < 120) = P(120) = \Phi\left(\frac{120-100}{10}\right) = \Phi(2) = 0.97725$$

Una pregunta inversa sería por ejemplo que valor tiene "x" tal, que una observación al azar sea menor que él, con una probabilidad de 0.95?

entonces

$$\Phi(u) = 0.95$$

$$u_{.95} = 1.645 = \frac{x_{.95} - \mu}{\sigma}$$

y

$$x_{.95} = 1.645(10) + 100 = 116.45$$

II.1.13 Probabilidad de la intersección de dos variables aleatorias

Consideremos el experimento "E" que consiste en sacar tres bombillas, una después de otra. Cada bombilla puede salir: A = Sin defectos, B = defectuosa; cada bombilla defectuosa puede a su vez salir: C = con el cristal roto D = con el filamento roto, F = con el cristal y el filamento roto. Cada bombilla que se saca tiene pues 4 posibles resultados. Si E_1 = sacar la primera bombilla, E_2 = sacar la segunda bombilla y E_3 = sacar la tercera bombilla, entonces $E = E_1 \cap E_2 \cap E_3$. Además, si cada uno de los e

ventos E_1 , E_2 y E_3 puede realizarse de cuatro maneras, el espacio muestral de E consta de $4^3 = 64$ puntos muestrales. Definamos ahora dos variables aleatorias

x = número de bombillas enteras

y = número de bombillas con el cristal y el filamento rotos

y hagamos

$p(A) = 1/2$; $p(B) = 1/2$; $p(C) = 2/9$; $p(D) = 2/9$ y

$p(F) = 1/18$

Si recordamos la distribución binomial (II.1.7), veremos que $p_n(x)$ ó $p_n(y)$ es una distribución de tal naturaleza donde $n=3$ y " x " ó " y " pueden ser 0, 1, 2 ó 3. Sin embargo ahora nos interesa $p(x \cap y)$; es decir la probabilidad de la intersección de las dos variables aleatorias. De ecuación 2.6

$p(x \cap y) = p(x) \cdot p(y/x)$ (2.70)
pero las variables " x " e " y " son independientes, luego

$p(x \cap y) = p(x) \cdot p(y)$ (2.71)
 $p(x)$, $p(y)$ y $p(x \cap y)$ pueden en forma general, representarse como en el cuadro 3

CUADRO 3

	x_1	x_2	...	x_m	
y_1	$p(x_1 \cap y_1)$	$p(x_2 \cap y_1)$...	$p(x_m \cap y_1)$	$p(y_1)$
y_2	$p(x_1 \cap y_2)$	$p(x_2 \cap y_2)$...	$p(x_m \cap y_2)$	$p(y_2)$
⋮	⋮	⋮		⋮	⋮
y_n	$p(x_1 \cap y_n)$	$p(x_2 \cap y_n)$...	$p(x_m \cap y_n)$	$p(y_n)$
	$p(x_1)$	$p(x_2)$...	$p(x_m)$	

II.1.14 Covariancia y Coeficiente de Correlación

La covariancia de dos variables aleatorias "x" e "y" se definió como la expectativa del producto de las desviaciones de "x" e "y" de sus respectivas expectativas:

$$\text{Cov}(x,y) = E[(x - E[x])(y - E[y])] \quad (2.72)$$

siendo, para propósitos de cómputo mas conveniente

$$\text{Cov}(x,y) = E[xy - yE(x) - xE(y) + E(x)E(y)] \quad (2.73)$$

La covariancia mide ó representa una clase especial de asociación entre las dos variables. Es de hacer notar que la variancia es un tipo especial de covariancia de una variable respecto a si misma.

$$\begin{aligned} \text{Cov}(x,x) &= E[(x - E[x])(x - E[x])] \\ &= E[(x - E[x])^2] = V(x) \end{aligned} \quad (2.74)$$

El coeficiente de correlación es la covariancia de dos variables aleatorias puestas en forma estandarizada (ecuación 2.63)

$$\rho_{xy} = \text{Cov}\left[\frac{x - E(x)}{\sqrt{V(x)}}, \frac{y - E(y)}{\sqrt{V(y)}}\right] \quad (2.75)$$

Utilizando la definición de covariancia (ecuación 2.72) se convierte en

$$\rho_{xy} = \frac{\text{Cov}(x,y)}{\sqrt{V(x)V(y)}} \quad (2.76)$$

si $\rho_{xy} = 0$ decimos que x e y no están correlacionadas, esto implica que $\text{Cov}(x, y) = 0$.

II.1.15 Variación y expectativa de una media aritmética

Supongamos x_1, x_2, \dots, x_n variables aleatorias y consideremos la combinación lineal

$$z = a_0 + a_1 x_1 + \dots + a_n x_n \quad (2.77)$$

siendo las a_i constantes. Entonces

$$E(z) = a_0 + a_1 E(x_1) + \dots + a_n E(x_n) \quad (2.78)$$

y

$$\begin{aligned} V(z) &= E[(z - E[z])^2] \\ &= E[(a_0 + a_1 x_1 + \dots + a_n x_n - a_0 - a_1 E[x_1] - \dots - a_n E[x_n])^2] \\ &= E[(a_1 x_1 - E[x_1]] + \dots + [a_n x_n - E[x_n]])^2] \\ &= E\left[\sum_{i=1}^n a_i^2 (x_i - E[x_i])^2 + \sum_{\substack{i, j \\ i \neq j}} a_i a_j (x_i - E[x_i])(x_j - E[x_j]) \right] \\ &= \sum_{i=1}^n a_i^2 E[(x_i - E[x_i])^2] + \sum_{\substack{i, j \\ i \neq j}} a_i a_j E[(x_i - E[x_i])(x_j - E[x_j])] \end{aligned}$$

$$V(z) = \sum_i^n a_i^2 \cdot V(x_i) + \sum_i^n \sum_{\substack{j \\ i \neq j}}^n a_i a_j \text{Cov}(x_i, x_j) \quad (2.79)$$

La doble sumatoria del segundo término excluye el caso $i=j$ que está incluido en el primer término.

Supongamos ahora el caso de tener obser-
vaciones independientes de la misma distribu-
ción, teniendo ésta media μ y variancia σ^2 , la
ecuación 2.77 puede escribirse

$$z = \frac{x_1}{n} + \dots + \frac{x_n}{n} = \frac{1}{n} \sum_i^n x_i = \bar{x}$$

es decir que z es la media de los n valo-
res; usando ecuaciones 2.78 y 2.79 tendremos

$$E(\bar{x}) = \frac{1}{n} E(x_1) + \dots + \frac{1}{n} E(x_n) = \frac{\mu}{n} + \dots + \frac{\mu}{n} = \mu \quad (2.80)$$

$$y \quad V(\bar{x}) = \frac{1}{n^2} V(x_1) + \dots + \frac{1}{n^2} V(x_n) = \frac{\sigma^2}{n^2} \cdot n = \frac{\sigma^2}{n} \quad (2.81)$$

Puede probarse que cualquier combinación
lineal de variables aleatorias sacadas de una
población normal está también normalmente dis-
tribuida; en particular esto se aplica a la me-
dia de n variables sacadas de la población nor-
mal $N(\mu, \sigma^2)$, la media tendrá una distribu-
ción normal $N(\mu, \sigma^2/n)$

II.2 CONCEPTOS ESTADÍSTICOS

II.2.1 Teoría del Muestreo

En muchos problemas estadísticos los datos que pueden obtenerse se refieren como una "muestra aleatoria" sacada de un universo, entendiéndose por muestra aleatoria (ó random) al conjunto de datos tales que cualquier dato del universo tiene la misma probabilidad de estar incluido en la muestra. Las cantidades representativas de la muestra se denominan estadísticas a diferencia de las cantidades representativas del universo, a las cuales se les llama parámetros.

Supongamos que, de una muestra aleatoria de 100 observaciones se ha computado una estadística - por ejemplo la media -. Este valor obtenido no es una cantidad absolutamente fija que pueda reproducirse exactamente al tomar otra muestra aleatoria con el mismo número de muestras. De hecho, si tal experimento se repitiera muchas veces obtendríamos valores de medias que formarían una distribución de frecuencia. Es decir que de un grupo de estadísticas similares obtendríamos una distribución de frecuencias llamada distribución muestral.

II.2.2 Estimación puntual y estimación de intervalos

Problemas relativos a estimar los valores reales de los parámetros del universo se dice que pertenecen a la "estimación puntual". Igualmente importante es conocer dentro de que

límites podemos esperar que estará un parámetro θ , con una cierta confianza (llamada grado de confianza); estos problemas se dice que se refieren a la "estimación de intervalos". Pueden resolverse si conocemos la distribución muestral de la estadística usada para estimar el parámetro θ .

Supongamos que se tiene un conjunto de muestras aleatorias de "n" observaciones cada una, cuyo universo ó población tiene una función conocida

$$p(x) = f(\theta_1, \dots, \theta_k) \quad (2.82)$$

dependiente de k parámetros. $p(x)$ podría ser por ejemplo la distribución normal en cuyo caso los parámetros serían σ y μ .

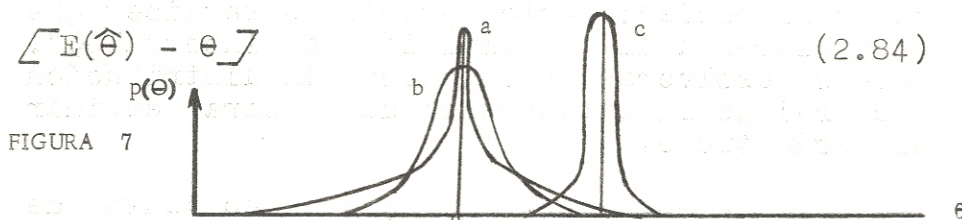
El problema general consiste en estimar u no ó mas de los parámetros. La función escogida a partir de las muestras aleatorias $\hat{\theta}(x_1 \dots x_n)$ se conoce como el "estimador" del parámetro, y cada uno de los valores numéricos de esta función se llama un "estimado".

Intuitivamente se quiere que el estimador $\hat{\theta}$ dé estimados cuya distribución se acerque en algún sentido a " θ " aunque este criterio no es suficiente para resolver ambigüedades. Por ejemplo en la figura 7 se muestran tres estimadores $\hat{\theta}_a$, $\hat{\theta}_b$ y $\hat{\theta}_c$; los dos primeros tienen la ventaja de que sus expectativas son iguales a θ ; es decir

$$E(\hat{\theta}) = \theta \quad (2.83)$$

ambos estimadores se llaman "insesgados". El "sesgo" de un estimador $\hat{\theta}$ se define como

$$E(\hat{\theta}) - \theta \quad (2.84)$$



El estimador $\hat{\theta}_a$ es preferible al $\hat{\theta}_b$ porque desviaciones no muy grandes ocurren menos frecuentemente; por otra parte, desviaciones extremas son más frecuentes con $\hat{\theta}_a$ que con $\hat{\theta}_b$. Si usamos la variancia $V(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2$ como medida de dispersión el estimador $\hat{\theta}_c$ es el mejor de los tres, sin embargo es un estimador "sesgado".

Un criterio útil para comparar estimadores es el "error-cuadrado medio" definido como el valor esperado del cuadrado de la desviación del estimador respecto al parámetro $E[(\hat{\theta} - \theta)^2]$.

$$\begin{aligned} E[(\hat{\theta} - \theta)^2] &= E[(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta)]^2 \\ &= E[(\hat{\theta} - E(\hat{\theta}))^2] + E[(E(\hat{\theta}) - \theta)^2] + 2[E(\hat{\theta}) - \theta] \cdot E[\hat{\theta} - E(\hat{\theta})] \\ &= V(\hat{\theta}) + (\text{sesgo})^2 \end{aligned} \quad (2.85)$$

Otro criterio que implica el comportamiento de un estimador a medida que el tamaño de la muestra aleatoria crece indefinidamente es la consistencia. Si la probabilidad del valor absoluto de la desviación de $\hat{\theta}$ respecto a θ

tiende a 1 a medida que "n" tiende a infinito, es decir

$$p(|\hat{\theta} - \theta|) \rightarrow 1 \text{ cuando } n \rightarrow \infty \quad (2.86)$$

entonces se dice que el estimador " $\hat{\theta}$ " es consistente.

Entre los estimadores consistentes habrá uno (ó varios) con variancia mínima. A este se le llama el estimador eficiente. La eficiencia de cualquier otro estimador se define como la razón de la variancia del estimador eficiente a la variancia del otro estimador.

II.2.3 El método de la máxima verosimilitud

Se trata de escoger aquel valor de θ que maximice la función de verosimilitud definida como

$$L = f(x_1, \theta) f(x_2, \theta) \dots f(x_n, \theta) \quad (2.87)$$

en donde $f(x, \theta)$ es la función de distribución de "x", de forma conocida y conteniendo al parámetro θ .

En la función L los valores x_1, \dots, x_n de la muestra aleatoria se consideran constantes y el parámetro θ se considera variable.

Estimadores escogidos por este método son a menudo sesgados pero tienen la ventaja de ser asintóticamente normales y asintóticamente

eficientes bajo condiciones generales.

Supóngase por ejemplo que se toma una muestra de n objetos de una población cuya distribución se sabe es binomial con parámetro θ . (θ podría ser por ejemplo la proporción de objetos defectuosos). De los " n " objetos de la muestra se encuentran x defectuosos y se quiere estimar θ . La función de verosimilitud es

$$L = f(x; n, \theta) = C_x^n \cdot \theta^x (1-\theta)^{n-x} \quad (2.88)$$

Al maximizar la función es más conveniente maximizar su logaritmo con lo cual se llega a lo mismo; es decir aquellos valores de los parámetros que maximizan $\ln L$ maximizarán a L . El logaritmo de 2.88 es

$$\ln L = \ln C_x^n + x \cdot \ln \theta + (n-x) \cdot \ln (1-\theta) \quad (2.89)$$

derivando respecto al parámetro

$$\frac{d \ln L}{d\theta} = \frac{x}{\theta} + (n-x) \frac{1}{1-\theta} \quad (2.90)$$

e igualando a cero se obtiene $\hat{\theta} = \frac{x}{n}$

Cuando la función de distribución tiene varios parámetros pueden encontrarse los estimadores de máxima verosimilitud ($\hat{\theta}_1, \hat{\theta}_2, \text{ etc}$) de cada uno de los parámetros, derivando respecto a los parámetros e igualando a cero, es decir

$$\frac{\partial \ln L}{\partial \theta_1} = \frac{\partial \ln L}{\partial \theta_2} = \dots = 0 \quad (2.91)$$

Algunas veces el valor de un parámetro (por ejemplo θ_1) que maximiza a $\ln L$ depende de otro parámetro (por ejemplo θ_2) de tal manera que para estimar θ_1 habremos de estimar θ_2 ; sin embargo este no es siempre el caso.

Como ejemplo adicional considérese la estimación del parámetro " μ " de una distribución normal a partir de una muestra de x_1, x_2, \dots, x_n observaciones. La función de verosimilitud se obtiene sustituyendo la función densidad (ecuación 2.28) en 2.87:

$$L = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-(x_1 - \mu)^2 / 2\sigma^2 \right] \dots \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-(x_n - \mu)^2 / 2\sigma^2 \right] \quad (2.92)$$

$$L = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left[-\frac{1}{2\pi\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right] \quad (2.93)$$

su \ln es

$$\ln L = -\frac{n}{2} \ln (2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \quad (2.94)$$

derivando respecto a μ

$$\frac{\partial \ln L}{\partial \mu} = -\frac{1}{2\sigma^2} \sum_{i=1}^n 2(x_i - \mu)(-1) \quad (2.95)$$

igualando a cero la derivada se obtiene

$$\sum_{i=1}^n (x_i - \mu) = \left(\sum_{i=1}^n x_i \right) - n\mu = 0$$

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.96)$$

que es por supuesto la media de la muestra. Hay que notar que se ha obtenido $\hat{\mu}$ sin tener que estimar σ^2 .

Para obtener el estimador de máxima verosimilitud de σ^2 derivamos 2.94 respecto a σ^2

$$\frac{\partial \ln L}{\partial \sigma^2} = - \frac{n}{2} \frac{1}{\sigma^2} - \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \mu)^2 \quad (2.97)$$

Si μ es conocida se iguala 2.97 a cero y se despeja σ^2 :

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (2.98)$$

Si μ no se conoce, se igualan a cero las derivadas parciales 2.95 y 2.97 obteniéndose $\hat{\mu} = \bar{x}$ como antes y

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n} \quad (2.99)$$

Si se hubiera querido obtener el estimador de σ se deriva 2.94 respecto a σ

$$\frac{\partial \ln L}{\partial \sigma} = - \frac{n}{\sigma} - \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 \quad (2.100)$$

igualando a cero y despejando σ :

$$\hat{\sigma} = \sqrt{\sum_1^n (x - \hat{\mu})^2 / n} = \sqrt{\sum_1^n (x - \bar{x})^2 / n} \quad (2.101)$$

puede verse entonces que

$$(\hat{\sigma})^2 = \hat{\sigma}^2 \quad (2.102)$$

propiedad general de los estimadores de máxima similitud, llamada INVARIANCIA.

Veamos ahora si el estimador $\hat{\sigma}^2$ es ó no insesgado; de 2.99

$$\begin{aligned} n\hat{\sigma}^2 &= \sum_1^n (x - \hat{\mu})^2 = \sum_1^n (x - \bar{x})^2 \\ &= \sum_1^n [(x - \mu) - (\bar{x} - \mu)]^2 \\ &= \sum_1^n (x - \mu)^2 - 2(\bar{x} - \mu) \sum_1^n (x - \mu) + n(\bar{x} - \mu)^2 \\ &= \sum_1^n (x - \mu)^2 - n(\bar{x} - \mu)^2 \quad (2.103) \end{aligned}$$

ya que

$$2(\bar{x} - \mu) \sum_1^n (x - \mu) = 2(\bar{x} - \mu) \left(\sum_1^n x - n\mu \right)$$

$$= 2(\bar{x}-\mu)(n\bar{x}-n\mu) = 2n(\bar{x}-\mu)^2$$

entonces

$$E(n\hat{\sigma}^2) = \sum_1^n E[(x-\mu)^2] - nE[(\bar{x}-\mu)^2] \quad (2.104)$$

pero

$$E[(x-\mu)^2] = V(x) = \sigma^2 \text{ y } E[(\bar{x}-\mu)^2] = V(\bar{x}) = \sigma^2/n \text{ puesto que } E(\bar{x}) = \mu. \text{ (2.80 y 2.81). Luego:}$$

$$nE(\hat{\sigma}^2) = n\sigma^2 - n \cdot \frac{\sigma^2}{n} = \sigma^2(n-1)$$

$$E(\hat{\sigma}^2) = \sigma^2 \left(\frac{n-1}{n} \right) \quad (2.105)$$

Luego el estimador de máxima similitud (ó verosimilitud) de la variancia de una distribución normal es sesgado. Si definimos s^2 como

$$s^2 = \sum_1^n \frac{(x-\bar{x})^2}{n-1} \quad (2.106)$$

su expectativa es

$$E(s^2) = E\left[\sum_1^n \frac{(x-\bar{x})^2}{n-1} \right] = E\left[\frac{1}{n-1} \frac{\sum_1^n (x-\bar{x})^2}{n} \right]$$

$$E(s^2) = \frac{n}{n-1} E(\hat{\sigma}^2) = \frac{n}{n-1} \sigma^2 \frac{n-1}{n} = \sigma^2 \quad (2.107)$$

luego es insesgado

II.2.4 Hipótesis Estadísticas

Las dos principales ramas de la inferencia estadística se refieren una a la estimación de parámetros y la otra a probar hipótesis estadísticas. La primera parte se esbozó brevemente en los artículos anteriores; algunos aspectos sobre las hipótesis estadísticas y sus pruebas se tratarán en este y los siguientes artículos.

Supóngase por ejemplo que se realiza un experimento cuyo resultado x se sabe que tiene la función de densidad de probabilidad $f(x; \theta)$

Se sabe además por resultados en experimentos anteriores que $\theta = a$. No obstante, debido a cambios en la forma de realizar el experimento se sospecha que θ ya no es igual a "a" sino igual a "b"; siendo a y b valores constantes; no existe sin embargo evidencia experimental de que $\theta = b$. Por tanto la declaración $\theta = b$ es una conjetura ó una HIPOTESIS ESTADISTICA. Al admitir que la hipótesis $\theta = b$ puede ser falsa permitimos la posibilidad de que $\theta = a$. Realmente hay pues, dos hipótesis. Primera que el parámetro $\theta = a$; es decir que no ha habido cambio en θ . Segunda, que el parámetro $\theta = b$. La primera hipótesis estadística es llamada hipótesis nula porque sostiene que no hay diferen-

cia entre el valor de θ en experimentos anteriores y aquellos que se lleven a cabo después. La hipótesis nula se representa por los símbolos $H_0 : \theta_0 = a$. La segunda hipótesis es llamada hipótesis alterna y se representa por $H_1 : \theta_1 = b$. Hipótesis de este tipo se conocen como hipótesis simples. Si otros parámetros no están completamente especificados, la hipótesis se llama compuesta.

Efectuamos una observación "x", y deseamos un criterio para aceptar la hipótesis nula (y rechazar la alterna) ó aceptar la alterna rechazando la nula.

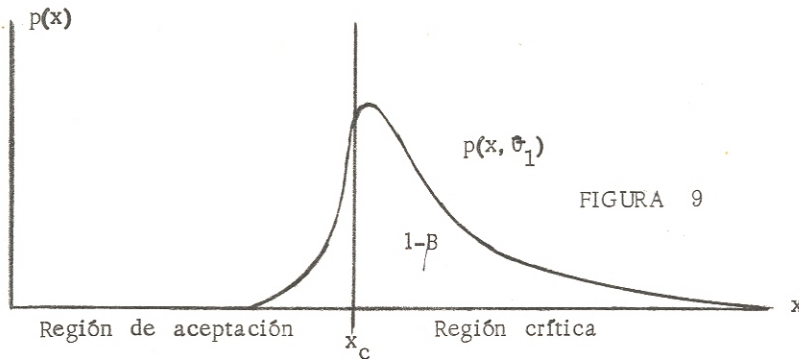
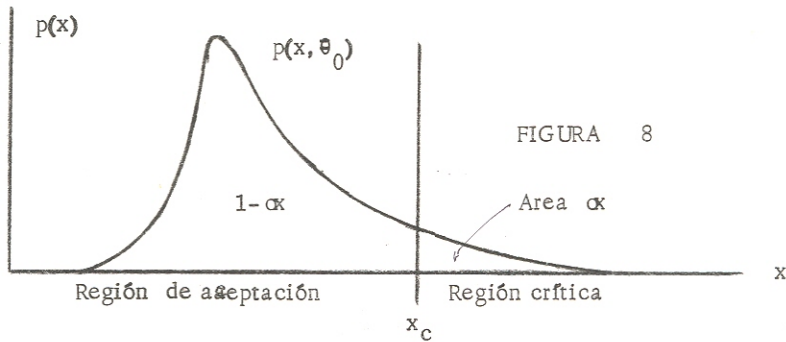
Se comienza por definir una región sobre la curva tal que si la observación cae en esa región entonces se acepta la hipótesis nula; esta es la llamada "región de aceptación". La región complementaria donde rechazamos la hipótesis nula es la región crítica.

En que nos basamos para escoger esas regiones?. El primer criterio es que solo una pequeña fracción α del tiempo rechazaremos la hipótesis nula como falsa cuando de hecho es verdadera. Rechazar la hipótesis nula como falsa siendo verdadera significa cometer un error del tipo I. Obviamente deseamos cometer tal error sólo raramente; valores comunes de α son .05 ó .01. El valor escogido para α se conoce como el nivel de significación de la prueba.

Mas específicamente, supongamos que bajo la hipótesis alterna, la distribución se corre

hacia la derecha sin cambiar de forma de tal manera que la expectativa de x bajo la hipótesis alterna $E_{\theta_1}(x)$ es mayor que $E_{\theta_0}(x)$, la expectativa bajo la hipótesis nula. Entonces es razonable determinar la región crítica como aquella parte sobre la curva a la derecha de algún valor x_c , donde x_c se escoge de tal manera que, si la hipótesis nula es correcta y x tiene la distribución $p(x, \theta_0)$, la probabilidad de que una observación al azar caiga en la región crítica es α ; es decir

$$\alpha = \int_{x_c}^{\infty} p(x; \theta_0) dx \quad (2.108)$$



Consideremos ahora que pasará si H_1 es verdadera y x tiene la distribución $p(x, \theta_1)$. Puesto que habíamos supuesto anteriormente que la diferencia entre $p(x, \theta_0)$ y $p(x, \theta_1)$ consistía en que la distribución $p(x, \theta_1)$ se corría a la derecha, la situación es como en la figura 9.

Aún bajo la condición H_1 nuestra observación " x " puede caer en la región de aceptación $x < x_c$. La probabilidad de que esto suceda es el área sombreada en la figura 9 representada por el símbolo β :

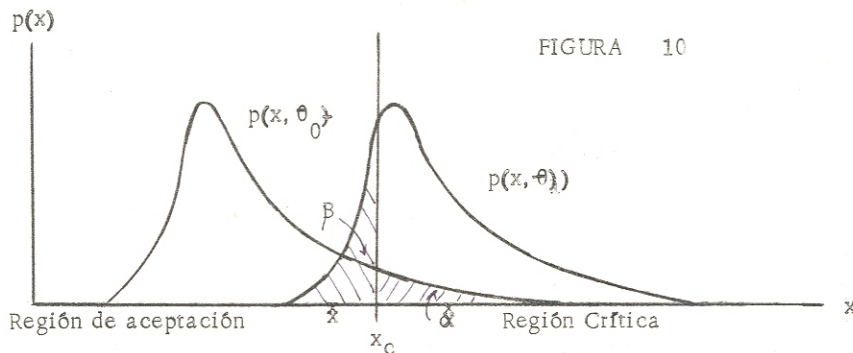
$$\beta = \int_{-\infty}^{x_c} p(x, \theta_1) dx \quad (2.109)$$

Aceptar H_0 como verdadera siendo falsa es un error, llamado error del tipo II. La probabilidad de que x caiga en la región crítica viniendo de la distribución $p(x, \theta_1)$ de tal manera que H_1 es cierta y H_0 falsa, se conoce como la potencia de la prueba, y está representada en la figura 9 por el área $(1-\beta)$. La potencia se representa por la letra griega π (pi); es decir

$$\pi = 1-\beta = \int_{x_c}^{\infty} p(x, \theta_1) dx \quad (2.110)$$

El procedimiento usual es fijar un valor pequeño de α conociendo la distribución $p(x, \theta)$ y los valores de los parámetros θ_0 y θ_1 . El valor de β queda entonces determinado. Debemos sin embargo, ser cuidadosos respecto a rechazar H_1 , ya que, aún aceptando H_0 existe la po-

sibilidad de que la observación pertenezca a la distribución H_1 . Tal situación está ilustrada en la figura 10 en donde, si el valor de la observación cae en el punto indicado está en la región de aceptación y por tanto aceptamos H_0 . Sin embargo tal valor de x podría ser perfectamente parte de la distribución alterna $p(x, \theta_1)$; por tanto en un caso como el presente es arriesgado descartar la posibilidad de que H_1 sea cierta. Sin embargo el procedimiento acostumbrado es aceptar H_0 en primera instancia pudiendo revisar nuestra opinión al obtener mayor información.



Resumiendo entonces: α es la probabilidad de rechazar la hipótesis nula como falsa siendo cierta; es decir α es la probabilidad de cometer un error del tipo I. β es la probabilidad de aceptar la hipótesis nula como cierta siendo falsa; es decir es la probabilidad de cometer un error del tipo II.

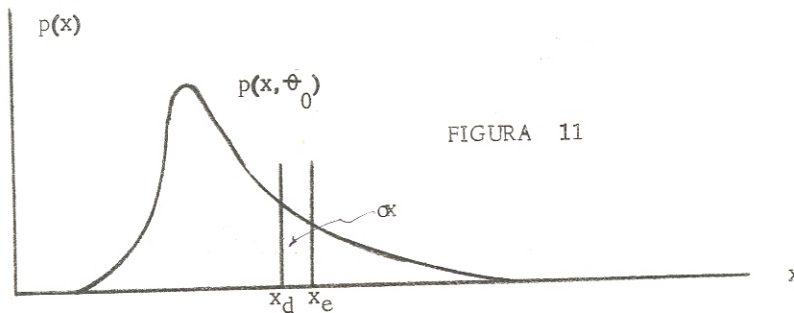
II.2.5 Un criterio para escoger entre pruebas alternas

En el artículo anterior construimos una prueba dividiendo el eje x en dos regiones, la región de aceptación y la región crítica, habiendo escogido para esta última la cola supe-

rior de la distribución $p(x, \theta_0)$ sin discutir si esta era la mejor elección. Consideremos ahora una región alterna crítica, por ejemplo la definida por $x_d \leq x \leq x_e$ en la figura 11 donde los límites x_d y x_e satisfacen la condición de que la probabilidad de cometer un error de tipo I sigue siendo α ; es decir

$$\int_{x_d}^{x_e} p(x, \theta_0) dx = \alpha \quad (2.111)$$

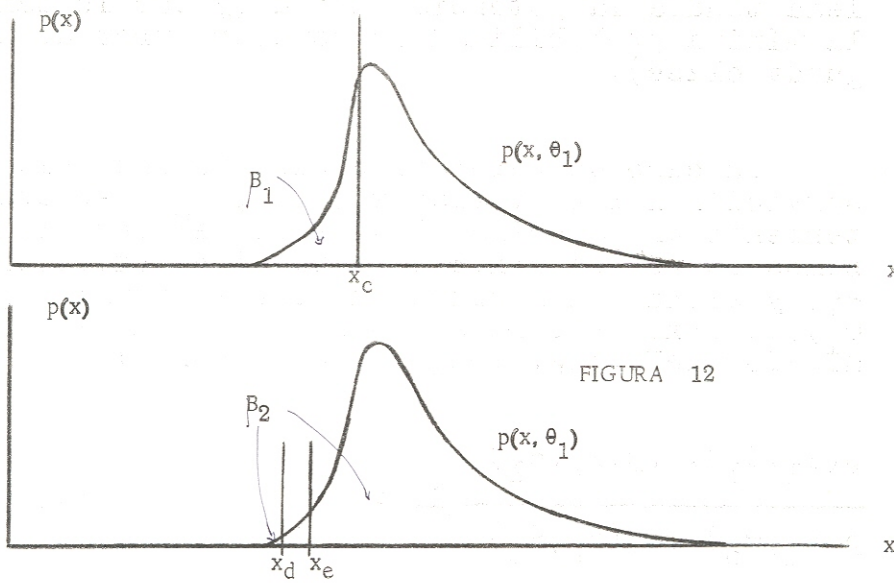
Puesto que ambas pruebas tienen la misma probabilidad de cometer errores del tipo I una forma de compararlas es por las probabilidades de cada una de cometer errores del tipo II



La figura 12 muestra en forma gráfica la probabilidad de cometer errores del tipo II para las dos regiones críticas representada por las áreas β_1 y β_2

$$\beta_1 = \int_{-\infty}^{x_c} p(x, \theta_1) dx \quad (2.112)$$

$$\beta_2 = \int_{-\infty}^{x_d} p(x, \theta_1) dx + \int_{x_e}^{\infty} p(x, \theta_1) dx \quad (2.113)$$



Puede verse que β_2 es mucho mayor que β_1 , luego es preferible la primera prueba.

De un modo más académico se utilizaría el lema de Pearson y Neyman que dice que la región crítica debe incluir aquellos valores de x tales que la relación $p(x, \theta_1)/p(x, \theta_0)$ sea un máximo. En general se pone

$$\frac{p(x, \theta_1)}{p(x, \theta_0)} \geq K \quad (2.114)$$

en donde K se deduce a partir de un valor del nivel de significación (α), previamente seleccionado.

Una prueba construida de acuerdo con este

lema tendrá la potencia máxima (y por lo tanto la mínima probabilidad de cometer error de segunda clase).

En caso de haber efectuado no una sola observación x sino varias x_1, \dots, x_n ; es decir teniendo una muestra al azar de valores utilizamos en vez de $p(x, \theta_1)$ la función $p(x_1, \dots, x_n, \theta_1)$ y siendo la muestra al azar es igual a $p(x_1, \theta_1) \dots p(x_n, \theta_1)$. Lo mismo hacemos con el denominador de 2.114. Luego podemos poner:

$$\frac{p(x_1, \theta_1) \dots p(x_n, \theta_1)}{p(x_1, \theta_0) \dots p(x_n, \theta_0)} \geq K \quad (2.115)$$

II.2.6 Prueba de una observación proveniente de una población normal con variancia conocida

Se trata de probar la hipótesis nula de una población normal con expectativa $\mu = \mu_0$ contra la hipótesis alterna $\mu_1 \geq \mu_0$. La prueba es hacia un lado de la curva.

De acuerdo con el lema de Pearson y Neyman la prueba óptima es aquella para la cual el alcance del valor de x dentro de la región crítica satisface la desigualdad

$$\frac{p(x, \mu_1, \sigma^2)}{p(x, \mu_0, \sigma^2)} \geq K \quad (2.116)$$

pero

$$\begin{aligned} \frac{p(x_1, \mu_1, \sigma^2)}{p(x, \mu_0, \sigma^2)} &= \frac{\frac{1}{\sqrt{2\pi} \sigma} \exp -(x-\mu_1)^2/2\sigma^2}{\frac{1}{\sqrt{2\pi} \sigma} \exp -(x-\mu_0)^2/2\sigma^2} \\ &= \exp -\left[(x-\mu_1)^2 - (x-\mu_0)^2 \right] / 2\sigma^2 \\ &= \exp x(\mu_1 - \mu_0) / \sigma^2 \cdot \exp -(\mu_1^2 - \mu_0^2) / 2\sigma^2 \end{aligned} \quad (2.117)$$

El segundo exponencial de 2.117 es una cantidad constante y positiva ya que μ_1 y μ_0 son dados y σ^2 es conocida; llamemosle a esta cantidad "c". El primer exponencial tiene la forma e^{ax} siendo $a = (\mu_1 - \mu_0) / \sigma^2$ una constante positiva. Luego la función $p(x, \mu_1, \sigma^2) / p(x, \mu_0, \sigma^2)$ tiene la forma ce^{ax} y entonces aumenta al aumentar x . Así, para escoger todos los valores de x para los cuales esta función es mayor que K se halla el mínimo valor de x dentro de la región crítica por

$$\frac{p(x_c, \mu_1, \sigma^2)}{p(x_c, \mu_0, \sigma^2)} = \exp x(\mu_1 - \mu_0) / \sigma^2 \cdot \exp -(\mu_1^2 - \mu_0^2) / 2\sigma^2 = K \quad (2.118)$$

y por logaritmos

$$\ln K = x_c \frac{\mu_1 - \mu_0}{\sigma^2} - \frac{\mu_1^2 - \mu_0^2}{2\sigma^2} \quad (2.119)$$

luego

$$x_c = \frac{2\sigma^2 \ln K / (\mu_1^2 - \mu_0^2)}{2(\mu_1 - \mu_0)} \quad (2.120)$$

El valor de K debe ajustarse de tal manera que se cumpla

$$\int_{x_c}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} \exp -(x-\mu_0)^2/2\sigma^2 \cdot dx = \alpha \quad (2.121)$$

la ecuación anterior implica que

$$1 - \alpha = \int_{-\infty}^{x_c} \frac{1}{\sqrt{2\pi} \sigma} \exp -(x-\mu_0)^2/2\sigma^2 \cdot dx \quad (2.122)$$

pero de acuerdo con el artículo II.1.9

$$1 - \alpha = \Phi\left(\frac{x_c - \mu_0}{\sigma}\right) = \Phi(u)$$

$$x_c - \mu_0 = u_{1-\alpha}$$

$$x_c = \mu_0 + \sigma u_{1-\alpha} \quad (2.123)$$

es decir que podemos evaluar x_c sin necesidad de calcular K previamente. Por ejemplo si se escoge un nivel de significación $\alpha=0.05$ se encuentra en tabla II del apéndice $u_{.95} = 1.645$ y $x_c = \mu_0 + 1.645\sigma$; es decir que el límite inferior de la región crítica está corrido 1.645 veces la desviación estándar respecto a la expectativa μ_0 bajo la hipótesis nula.

Hay que hacer notar que de ecuación 2.123 el valor de x_c es independiente de μ_1 ; luego la región crítica será la misma cualquiera que sea el valor de μ_1 siempre que sea mayor de μ_0 .

La potencia de la prueba es la probabilidad de rechazar la hipótesis nula $\mu = \mu_0$ cuando μ es realmente igual a μ_1 . En otras palabras es la probabilidad de que la observación x caiga en la región crítica bajo la hipótesis alterna. Entonces la potencia es

$$\pi(\mu_1) = \int_{x_c}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp -(x-\mu_1)/2\sigma^2. dx = 1 - \beta \quad (2.124)$$

$$= 1 - \Phi\left(\frac{x_c - \mu_1}{\sigma}\right) \quad (2.125)$$

sustituyendo x_c de 2.123 en 2.125

$$\pi(\mu_1) = 1 - \Phi\left(\frac{\mu_0 + \sigma u_{1-\alpha} - \mu_1}{\sigma}\right) = 1 - \Phi\left(u_{1-\alpha} - \frac{\mu_0 - \mu_1}{\sigma}\right)$$

$$\pi(\mu_1) = \Phi\left(-u_{1-\alpha} - \frac{\mu_0 - \mu_1}{\sigma}\right) = \Phi\left(u_{\alpha} + \frac{\mu_1 - \mu_0}{\sigma}\right) \quad (2.126)$$

Lo anterior debido al hecho de que $1 - \Phi(u) = \Phi(-u)$ y $-u_{1-\alpha} = u_{\alpha}$ por la simetría de la distribución normal.

Hasta este punto se han considerado pruebas basadas en una sola observación. Si se efectúan n observaciones utilizamos la ecuación 2.115. Para el caso del presente artículo tenemos una muestra de n observaciones independientes sacadas de una distribución normal con variancia conocida y media no conocida. Deseamos probar la hipótesis nula $H_0: \mu = \mu_0$ contra la alterna $\mu > \mu_0$ a un nivel α . Sustituyendo en ec. 2.115

$$\begin{aligned} \frac{p(x_1, \mu_1, \sigma^2) \dots p(x_n, \mu_1, \sigma^2)}{p(x_1, \mu_0, \sigma^2) \dots p(x_n, \mu_0, \sigma^2)} &= \frac{\frac{1}{\sqrt{2\pi} \sigma} \cdot \exp}{\frac{1}{\sqrt{2\pi} \sigma} \cdot \exp} \\ &= \frac{\exp\left[-(x_1 - \mu_1)^2 / 2\sigma^2 \dots - (x_n - \mu_1)^2 / 2\sigma^2\right]}{\exp\left[-(x_1 - \mu_0)^2 / 2\sigma^2 \dots - (x_n - \mu_0)^2 / 2\sigma^2\right]} \\ &= \frac{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_1)^2\right]}{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2\right]} = \exp\left[\bar{x} \frac{(\mu_1 - \mu_0)}{\sigma^2/n} - \frac{\mu_1^2 - \mu_0^2}{2\sigma^2/n}\right] \quad (2.127) \end{aligned}$$

Esta ecuación es similar a 2.117 con la diferencia de que es función de \bar{x} en vez de x y de σ^2/n en vez de σ^2 luego puede tratarse - similarmente y el límite inferior de la región crítica es, al igual que la 2.123:

$$\bar{x} = \mu_0 + \frac{u_{1-\alpha}\sigma}{\sqrt{n}} \quad (2.128)$$

asimismo la potencia es

$$\pi(\mu_1) = \Phi\left(u_{\alpha} + \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\right) = \Phi(u_{\alpha} + \delta) \quad (2.129)$$

definiendo

$$\delta = \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}$$

Por ejemplo para una prueba a un nivel de $\alpha = .05$, $u_{\alpha} = -1.645$ y para $\delta = 1$

$$\pi(\mu_1) = \Phi(-1.645 + 1) = \Phi(-.645) = 0.259$$

Graficando los valores de potencias contra valores de " δ " a un nivel de significación dado se obtienen curvas como la representada en la figura 13 graficada para un nivel de .05

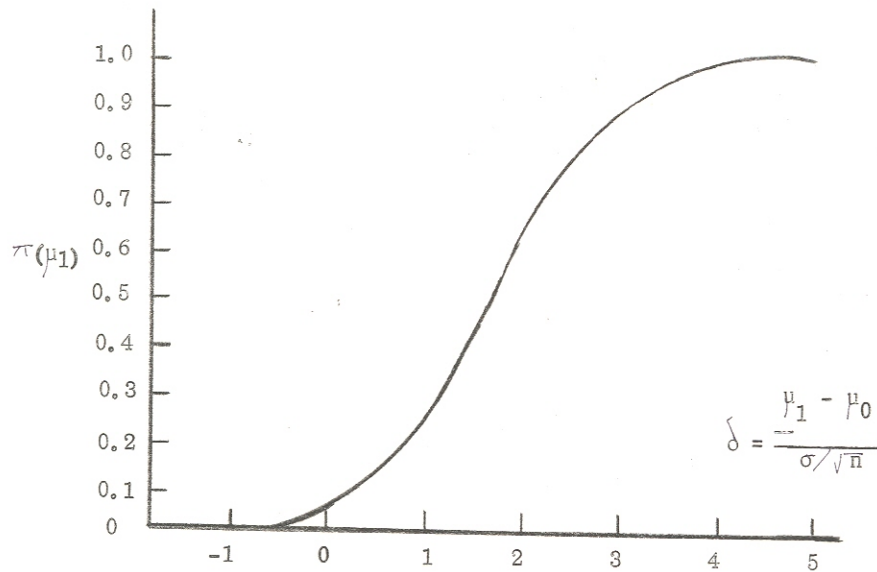


Figura 13

La figura anterior muestra que cuando $\mu_1 = \mu_0$ la potencia coincide con el nivel de significación. Además cuando $\mu_1 < \mu_0$ hay pocas probabilidades de rechazar $H_0: \mu = \mu_0$ en favor de $\mu = \mu_1 < \mu_0$ ó dicho en otras palabras hay pocas probabilidades de NO cometer un error de tipo II.

Puede utilizarse la ecuación 2.129 para calcular el número de observaciones necesarias para dar una potencia específica, asumiendo que σ es conocida:

$$\Phi\left(u_{\alpha} + \frac{\mu_1 - \mu_0}{\sigma / \sqrt{n}}\right) = 1 - \beta$$

pero en general

$$\Phi(u_p) = P; \text{ y por tanto}$$

$$\Phi(u_{1-\beta}) = 1 - \beta$$

luego

$$u_{\alpha} \neq \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} = u_{1-\beta}$$

despejando n

$$n = (u_{1-\beta} - u_{\alpha})^2 \left(\frac{\sigma}{\mu_1 - \mu_0}\right)^2 \quad (2.130)$$

pero $u_{\alpha} = -u_{1-\alpha}$ y entonces

$$n = (u_{1-\beta} \neq u_{1-\alpha})^2 \left(\frac{\sigma}{\mu_1 - \mu_0}\right)^2 \quad (2.131)$$

II.2.7 Hipótesis Compuestas

En el artículo II.2.5 se estudió un criterio para escoger entre pruebas alternas de hipótesis simples. Para funciones con varios parámetros, la ecuación 2.115 puede ponerse

$$\frac{p(x_1, \dots, x_n, \theta_{1,a}, \theta_2, \dots, \theta_k)}{p(x_1, \dots, x_n, \theta_{1,0}, \theta_2, \dots, \theta_k)} \geq K \quad (2.132)$$

en donde los valores de θ son parámetros de los cuales $\theta_2, \dots, \theta_k$ son especificados y θ_1 toma el valor $\theta_{1,0}$ bajo la hipótesis nula y $\theta_{1,a}$ bajo la hipótesis alterna.

Los parámetros $\theta_1, \dots, \theta_k$ pueden considerarse situados en un sistema de coordenadas. Un punto es suficiente para definir una hipótete

sis simple mientras que una región es necesaria para especificar una hipótesis compuesta. Por ejemplo para una distribución normal hay dos parámetros μ y σ^2 representados por las abscisas y ordenadas de la figura 14 respectivamente

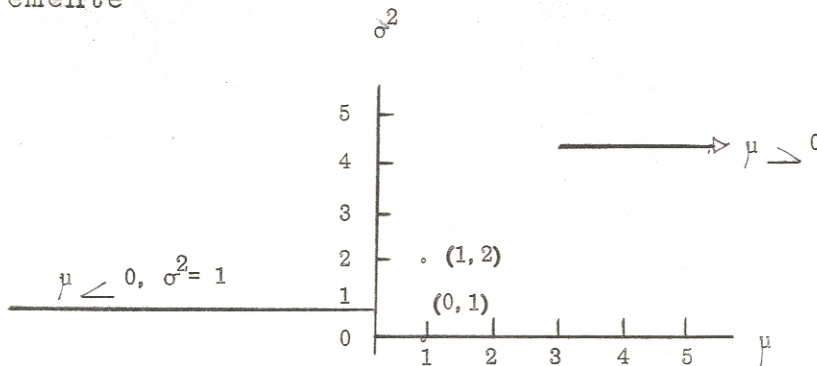


FIGURA 14

Las hipótesis simples $H_0: \mu = 0, \sigma^2 = 1$ y $H_1: \mu = 1, \sigma^2 = 2$ están representadas por los puntos $(0, 1)$ y $(1, 2)$ en el sistema. La hipótesis $H_3: \mu \leq 0, \sigma^2 = 1$ es compuesta (μ no está completamente definido) y está representada por la horizontal $\sigma^2 = 1$ a la izquierda del eje σ^2 . La hipótesis $H_4: \mu \geq 0$ está representada por todo el primer cuadrante y es naturalmente una hipótesis compuesta. La región admisible de las hipótesis nula y alterna se representa por el símbolo Ω . En el ejemplo la región abajo del eje μ es inadmisibles puesto que σ^2 no puede jamás ser negativo. La región que comprende una hipótesis compuesta que sea a la vez nula se representa por w .

Supongamos una hipótesis nula y compuesta; la función

$$\begin{aligned}
 & p(x_1, \theta_1, \dots, \theta_k) p(x_2, \theta_1, \dots, \theta_k) p(x_n, \theta_1, \dots, \theta_k) \\
 & = p(x_1, \dots, x_n, \theta_1, \dots, \theta_k) \quad (2.133)
 \end{aligned}$$

es la función de verosimilitud definida por ecuación 2.87. Por el método de máxima verosimilitud pueden encontrarse los valores de los parámetros que maximizen a 2.133 que será representada por $L(w)$. Supongamos ahora que los parámetros varían en la región Ω es decir en la región de la hipótesis nula y la alterna. Bajo esta condición, podemos encontrar los estimadores de máxima verosimilitud que al ser insertados en 2.133 dan lugar a la función de verosimilitud bajo ambas hipótesis. Llámese a esta función $L(\Omega)$. La ecuación 2.115 siendo para hipótesis simples es, para hipótesis compuestas la razón $L(\Omega)/L(w)$. Sin embargo se usa mas frecuentemente el inverso de esta razón representada por

$$\lambda = \frac{L(w)}{L(\Omega)} \quad (2.134)$$

y llamada módulo de verosimilitud. λ es una cantidad positiva ya que es la razón de productos de funciones de probabilidad que siempre son positivas. Además λ no puede ser mayor de 1 puesto que el valor máximo de la función de verosimilitud al variar los parámetros en una región w no puede ser mayor que la función al variar los parámetros en la región Ω , siendo que $\Omega \supset w$.

Un valor pequeño de λ indica que la verosimilitud computada usando w (región para la hipótesis nula) es improbable así que probablemente lo mejor sería rechazar la hipótesis nu-

la. Por otro lado un valor de λ próximo a uno indica una buena escogencia de la hipótesis nu la y por tanto debiera ser aceptada.

II.2.8 Otra prueba de un grupo de observaciones provenientes de una población normal con variancia conocida

En la sección II.2.6 se estudió una prueba hacia un solo lado de la distribución. En la presente sección se esbozará una prueba pero hacia ambos lados de la curva.

Para aplicar el procedimiento explicado en la sección anterior en probar si un grupo de "n" observaciones proviene de una población normal

a) con media μ_0 y variancia conocida σ^2

b) con media $\mu_1 \neq \mu_0$ y variancia conocida σ^2

repetiremos la función de verosimilitud expresada en ecuación 2.93

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot \exp - \frac{1}{2\sigma^2} \sum_1^n (x-\mu)^2$$

En el artículo II.2.3 se maximizó L derivando respecto a μ obteniéndose $\mu = \sum x/n = \bar{x}$ (ecuación 2.96); es decir se permitió a μ variar luego

$$L(\Omega) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot \exp - \frac{1}{2\sigma^2} \sum_1^n (x-\bar{x})^2$$

Bajo la hipótesis nula no hay parámetros que varíen puesto que σ^2 se supone conocida y μ tiene el valor fijo μ_0 , luego

$$L(w) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot \exp - \frac{1}{2\sigma^2} \sum_1^n (x-\mu_0)^2$$

sustituyendo estos valores en ecuación 2.134 tendremos

$$\lambda = \frac{\exp - \frac{1}{2\sigma^2} \sum_1^n (x-\mu_0)^2}{\exp - \frac{1}{2\sigma^2} \sum_1^n (x-\bar{x})^2}$$

pero

$$\begin{aligned} \sum_1^n (x-\mu_0)^2 &= \sum_1^n [(x-\bar{x}) + (\bar{x}-\mu_0)]^2 \\ &= \sum_1^n (x-\bar{x})^2 + n(\bar{x}-\mu_0)^2 \end{aligned}$$

esto porque

$$\begin{aligned} 2 \sum_1^n (x-\bar{x})(\bar{x}-\mu_0) &= 2(\bar{x}-\mu_0) \sum_1^n (x-\bar{x}) \\ &= 2(\bar{x}-\mu_0) \cdot 0 = 0 \end{aligned}$$

luego el módulo de verosimilitud es

$$\lambda = \exp - \frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2 \quad (2.135)$$

Si definimos una región crítica por

$$0 < \lambda < C \quad (2.136)$$

donde C es una constante escogida de tal manera que el nivel de significación de la prueba sea α , entonces el límite de la región crítica \bar{x}_c está determinado por

$$\exp - \frac{n}{2\sigma^2} (\bar{x}_c - \mu_0)^2 = C$$

y despejando \bar{x}_c

$$\bar{x}_c = \mu_0 \pm \sqrt{-\ln C} \sqrt{2} \frac{\sigma}{\sqrt{n}} \quad (2.137)$$

Siendo C menor de "1", $\ln C$ es negativo y \bar{x}_c puede ponerse en la forma

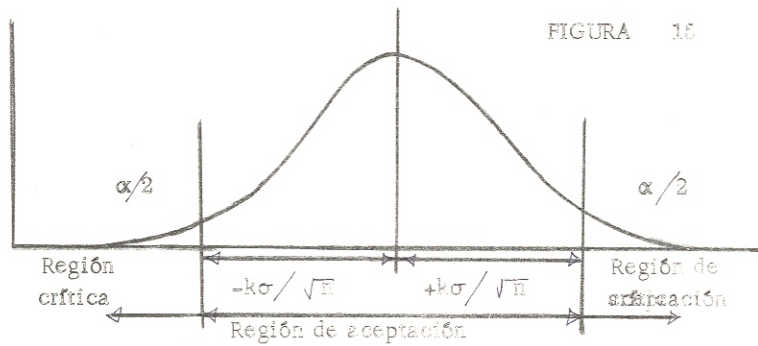
$$\bar{x}_c = \mu_0 \pm k \frac{\sigma}{\sqrt{n}} \quad (2.138)$$

donde

$$k = \sqrt{-2 \ln C}$$

Si \bar{x} se aleja mucho de μ_0 en mas ó en menos, $(\bar{x} - \mu_0)^2$ en ecuación 2.135 será grande y la razón de verosimilitud tendrá un valor pequeño. Así, valores grandes de $\bar{x} - \mu_0$ en valor absoluto presuponen un rechazo de la hipótesis nula.

En la figura 15 están las regiones críticas a ambos lados de la curva, siendo cada una de ellas $\alpha/2$ por simetría de la distribución



La región crítica está dada por

$$\bar{x} < \mu_0 - u_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

y

$$\bar{x} > \mu_0 + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (2.139)$$

Por la simetría de la curva podemos simplificar ecuación 2.139 diciendo: rechazar la hipótesis siempre que

$$\left| \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \right| > u_{1-\alpha/2} \quad (2.140)$$

La potencia de la prueba, es decir la probabilidad de NO cometer un error del tipo II es

$$\pi(\mu_1) = p\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} < u_{\alpha/2}\right) + p\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > u_{1-\alpha/2}\right)$$

ecuación que puede ponerse en la forma

$$\begin{aligned}
 \pi(\mu_1) &= p\left(\frac{\bar{x}-\mu_1}{\sigma/\sqrt{n}} \neq \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} \angle u_{\alpha/2}\right) \\
 &\neq p\left(\frac{\bar{x}-\mu_1}{\sigma/\sqrt{n}} \neq \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} \triangleright u_{1-\alpha/2}\right) \\
 &= p\left(\frac{\bar{x}-\mu_1}{\sigma/\sqrt{n}} \angle u_{\alpha/2} - \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}\right) \\
 &\neq 1 - p\left(\frac{\bar{x}-\mu_1}{\sigma/\sqrt{n}} \angle -u_{\alpha/2} - \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}\right)
 \end{aligned}
 \tag{2.141}$$

Bajo la hipótesis alterna $\frac{\bar{x}-\mu_1}{\sigma/\sqrt{n}}$ es una u nidad estandarizada, luego 2.141 puede ponerse

$$\pi(\mu_1) = p(u \angle k_1) \neq 1 - p(u \angle k_2)$$

pero del artículo II.1.9:

$$p(u \angle k_1) = \Phi(k_1) = \int_{-\infty}^{k_1} \phi(u) \cdot du$$

y

$$p(u \angle k_2) = \Phi(k_2) = \int_{-\infty}^{k_2} \phi(u) \cdot du$$

Definiendo $\frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} = \delta$ al igual que en artículo II.2.6

$$\pi(\mu_1) = \Phi(u_{\alpha/2} - \delta) + 1 - \Phi(-u_{\alpha/2} - \delta)$$

pero

$$1 - \Phi(-u) = \Phi(u)$$

y

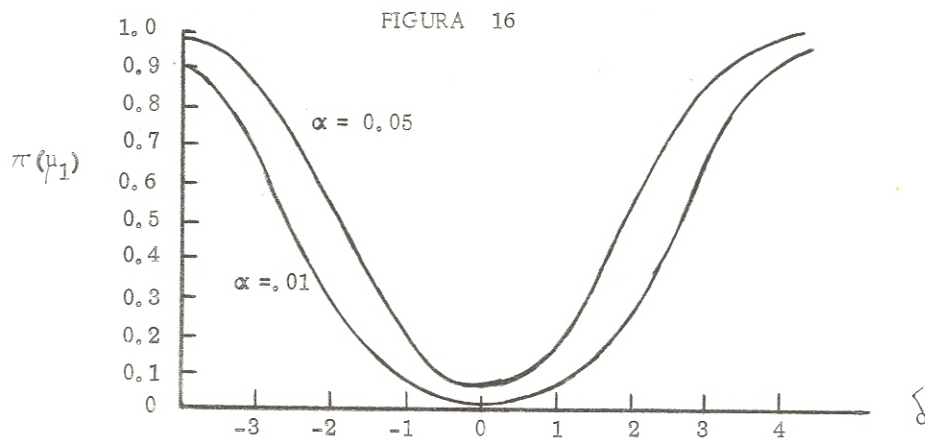
$$\pi(\mu_1) = \Phi(u_{\alpha/2} - \delta) + \Phi(u_{\alpha/2} + \delta) \quad (2.142)$$

Por ejemplo $\delta = 1$ y $\alpha = .05$ de tabla II en apéndice; $u_{\alpha/2} = -1.96$ y la potencia

$$\pi(\mu_1) = \Phi(-1.96 - 1) + \Phi(-1.96 + 1) = \Phi(-2.96) + \Phi(-.96)$$

$$= .0015 + .1685 = 0.17$$

En la figura 16 están graficados los valores de potencia contra δ para $\alpha = .05$ y para $\alpha = .01$



Una comparación de las curvas nos dice lo siguiente:

a) cuando $\mu_1 = \mu_0$, $\delta = 0$ y la potencia coincide con el nivel de significación

b) Para una misma prueba podemos aumentar la potencia (probabilidad de NO cometer error II) solo a costa de aumentar α (probabilidad de cometer error I)

Se vió que el primer sumando del miembro de la derecha de ecuación 2.142 hacía variar en muy poco la potencia $\pi(\mu_1)$; luego podemos hallar aproximadamente y en forma simple el número de observaciones necesarias para dar una potencia específica despreciando el primer sumando en 2.60, es decir

$$\pi(\mu_1) \approx \Phi\left(u_{\alpha/2} + \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\right)$$

ó

$$u_{\alpha/2} + \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} \approx u_{1-\beta}$$

$$n = \left(u_{1-\beta} - u_{\alpha/2}\right)^2 \left(\frac{\sigma}{\mu_1 - \mu_0}\right)^2 \quad (2.143)$$

II.2.9 Prueba de un grupo de observaciones provenientes de una población normal con variancia no conocida

Se trata en esta sección de encontrar una

forma de probar la hipótesis nula $H_0: \mu = \mu_0$ contra la alterna $H_1: \mu \neq \mu_0$ con la variancia σ^2 no especificada. La función de verosimilitudes, de 2.93

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot \exp - \frac{1}{2\sigma^2} \sum_1^n (x-\mu)^2 \quad (2.144)$$

Sustituyendo $\mu = \mu_0$ en la ecuación anterior y derivando $\ln L$ respecto a σ^2 para obtener el estimador de máxima verosimilitud de σ^2 en la región w

$$\frac{d(\ln L)}{d(\sigma^2)} = -\frac{n}{2} \frac{1}{\sigma^2} - \frac{1}{2(\sigma^2)^2} \sum_1^n (x-\mu_0)^2$$

igualando a cero y despejando $\hat{\sigma}^2$ se obtiene

$$\hat{\sigma}^2 = \frac{\sum_1^n (x-\mu_0)^2}{n} \quad (2.145)$$

introduciendo este valor y $\mu = \mu_0$ en 2.144 obtenemos

$$L(w) = \left(\frac{n}{2\pi \sum_1^n (x-\mu_c)^2}\right)^{n/2} \cdot \exp -n/2 \quad (2.146)$$

Si no imponemos restricciones a los parámetros estaremos en la región Ω (ver sección II.2.7); además de ecuaciones 2.96 y 2.99: $\hat{\mu} = \sum_1^n x/n$ y $\hat{\sigma}^2 = \sum_1^n (x-\hat{\mu})^2/n$. Insertando es-

tos valores en 2.144 se obtiene

$$L(\Omega) = \left(\frac{n}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right)^{n/2} \cdot \exp -n/2 \quad (2.147)$$

luego el módulo de verosimilitud es

$$\lambda = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2} \right)^{n/2} \quad (2.148)$$

Puede obtenerse un valor de prueba similar a λ que probablemente sea mas fácil de tabular tomando en cuenta que cualquier función monótona de λ puede utilizarse en vez de λ para definir la región crítica; luego

$$\lambda^{2/n} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}$$

esto porque

$$\begin{aligned} \sum_{i=1}^n (x_i - \mu_0)^2 &= \sum_{i=1}^n (x_i^2 - 2x_i\mu_0 + \mu_0^2) \\ &= \sum_{i=1}^n x_i^2 - 2\mu_0 \sum_{i=1}^n x_i + n\mu_0^2 \end{aligned}$$

sumando y restando $\frac{2}{n} \left(\sum_1^n x \right)^2$

$$\begin{aligned} \sum_1^n (x - \mu_0)^2 &= \sum_1^n x^2 - \frac{2}{n} \left(\sum_1^n x \right)^2 + 2n \left(\frac{\sum_1^n x}{n} \right)^2 - 2\mu_0 \sum_1^n x \\ &\quad + n\mu_0^2 \\ &= \sum x^2 - 2\bar{x} \cdot \sum x + n\bar{x}^2 - 2\mu_0 n\bar{x} + n\mu_0^2 + n\bar{x}^2 \\ &= \sum (x^2 - 2\bar{x}x + \bar{x}^2) + n(\bar{x}^2 - 2\mu_0 \bar{x} + \mu_0^2) \\ &= \sum (x - \bar{x})^2 + n(\bar{x} - \mu_0)^2 \end{aligned}$$

también

$$\chi^2/n = \frac{1}{1/\frac{n(\bar{x} - \mu_0)^2}{\sum_1^n (x - \bar{x})^2}} = \frac{1}{1/\frac{1}{(n-1) \frac{\sum_1^n (x - \bar{x})^2}{(n-1)}}} = \frac{1}{\frac{1}{n}}$$

si

$$t = \frac{\bar{x} - \mu_0}{\sqrt{\frac{\sum_1^n (x - \bar{x})^2}{(n-1)}} / \sqrt{n}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \quad (2.149)$$

$$\chi^2/n = \frac{1}{1/\frac{t^2}{n-1}} \quad (2.150)$$

Con su definición por ecuación 2.106. La hipótesis nula se rechazará para valores pequeños de λ que es lo mismo que decir para valores grandes de $|t|$. La ecuación 2.149 puede entonces ser utilizada para probar la hipótesis nula $H_0: \mu = \mu_0$ contra la alterna $H_1: \mu \neq \mu_0$ siendo σ^2 desconocida.

II.2.10 Comparación de dos medias aritméticas

Supóngase que se tienen dos grupos de muestras con medias \bar{x}_1 y \bar{x}_2 cada uno, el número de muestras es n_1 y n_2 respectivamente y han sido sacadas de dos distribuciones normales $N(x, \mu_1, \sigma_1^2)$ y $N(x, \mu_2, \sigma_2^2)$. Defínase $d = \bar{x}_1 - \bar{x}_2$; de ecuación 2.78 $E(d) = \delta = E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$. Las medias \bar{x}_1 y \bar{x}_2 tendrán variancias σ_1^2/n_1 y σ_2^2/n_2 y covariancia cero, entonces por ecuación 2.79 $V(d) = (\sigma_1^2/n_1 + \sigma_2^2/n_2)$. "d" está normalmente distribuida (ver último párrafo de sección II.1.15); luego

$$\frac{d - E(d)}{\sqrt{V(d)}} = \frac{d - \delta}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \quad (2.151)$$

es una variable normal estandarizada u (recordar ecuación 2.63). Podemos probar la hipótesis nula: d con variancia $(\sigma_1^2/n_1 + \sigma_2^2/n_2)$ -- proviene de una distribución normal con expectativa δ_0 . Para una cola de la curva establecemos el límite de la región crítica (ecuación 2.123) con $\delta = \delta_0$

$$d = \delta_0 + u_{1-\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (2.152)$$

y para la prueba en las dos colas de la curva con $\delta = \delta_1 \neq \delta_0$ estableceremos la región crítica (ecuación 2.139) por

$$d \leq \delta_0 \leq u_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} ; \quad d \geq \delta_0 \geq u_{1-\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (2.153)$$

II.2.11 El valor P

Ya se ha visto en artículos anteriores la forma de probar algunas hipótesis mediante el establecimiento de una región crítica con nivel α . El procedimiento inverso ofrece no obstante algunas ventajas de apreciación. En vez de prefijar α , lo que se hace es, al hacer una observación x , reemplazar el signo de desigualdad por una igualdad y despejar α . Este valor así encontrado se conoce como el valor P.

Supongamos que observamos $x=125$ de una población normal con variancia 100, y deseamos probar la hipótesis $\mu_0 = 100$ a un nivel $\alpha = 0.05$ en las dos colas de la curva, entonces

$$\left| \frac{125-100}{\sqrt{100}} \right| = 2.5 \geq u_{.975} = 1.96$$

y rechazamos la hipótesis nula a un nivel de 0.05. Alternativamente podemos escribir

$$\left| \frac{125-100}{\sqrt{100}} \right| = 2.5 = u_{1-P/2}$$

despejando hallamos $P = 0.012$. A menudo es mas satisfactorio saber que la hipótesis nula puede ser rechazada a un nivel tan pequeño como 0.012 que saber que se rechazará a un nivel de 0.05

II.2.12 Límites de Confianza (ó confianza)

En el artículo II.2.2 se explicaron brevemente los límites de confianza, en el presente se tratará sobre ellos con mas detenimiento.

El método general para determinar los límites de confianza es como sigue. Suponemos que tenemos un universo con una función de densidad de probabilidad conocida $p(x, \theta)$, siendo "x" la variable aleatoria θ el parámetro en cuestión. Supondremos también que tenemos un estimado "T" de θ , que T es una función de "x" y que podemos derivar una función de densidad que involucre a T y a θ . Por ejemplo si $p(x; \theta)$ es $N(\mu, \sigma^2)$ con σ^2 conocida, sabemos (ecuación 2.96) que \bar{x} es un estimado apropiado de μ y la función de densidad de \bar{x} es

$$p(\bar{x}; \mu, \frac{\sigma^2}{n}) = \frac{1}{\sqrt{2\pi} (\sigma/\sqrt{n})} \exp \left[- \frac{(\bar{x}-\mu)^2}{2\sigma^2/n} \right]$$

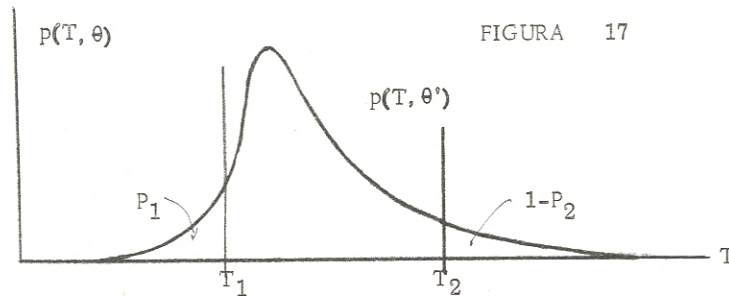
Volviendo al asunto, la función de densidad $p(T, \theta)$ variará así como varíe θ y para un valor particular de $\theta = \theta'$ existirán dos puntos

T_1 y T_2 que limitan dos regiones definidas por

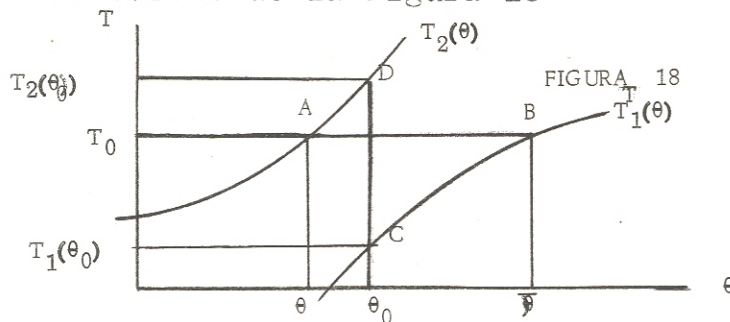
$$p(T \leq T_1; \theta = \theta') = \int_{-\infty}^{T_1} p(T; \theta') dT = P_1 \quad (2.154)$$

$$p(T \geq T_2; \theta = \theta') = \int_{T_2}^{\infty} p(T; \theta') dT = 1 - P_2 \quad (2.155)$$

e indicadas en la figura 17



Haciendo constantes P_1 y P_2 , los valores de T_1 y T_2 variarán al variar θ . Son por tanto funciones de θ y se expresarán $T_1(\theta)$ y $T_2(\theta)$. Podríamos graficar estas funciones contra θ y para simplificar supondremos que toman las formas de las curvas de la figura 18



Asumamos ahora que el verdadero valor del parámetro es θ_0 , entonces $T_1(\theta)$ y $T_2(\theta)$ toman

los valores $T_1(\theta_0)$ y $T_2(\theta_0)$ respectivamente y

$$p\left[T \leq T_1(\theta_0)\right] = P_1$$

$$p\left[T \geq T_2(\theta_0)\right] = 1 - P_2$$

Supongamos que de una observación muestral se ha obtenido un valor numérico de "T", por ejemplo T_0 . Con un grado de confianza dado obtendríamos dos valores límites del parámetro θ y $\bar{\theta}$. Hay que notar que θ_0 puede estar en cualquier punto pero si AB intersecta a CD, entonces T_0 debe estar entre $T_1(\theta_0)$ y $T_2(\theta_0)$ y simultáneamente " θ " debe estar incluido en el intervalo $(\underline{\theta}, \bar{\theta})$. Luego si la probabilidad de que T_0 esté entre $T_1(\theta_0)$ y $T_2(\theta_0)$ es $(P_2 - P_1)$ entonces la probabilidad de que θ_0 esté entre $\underline{\theta}$ y $\bar{\theta}$ es también $(P_2 - P_1)$ y podemos escribir

$$p(\underline{\theta} < \theta < \bar{\theta}) = P_2 - P_1 \quad (2.156)$$

El método se ilustrará encontrando los límites de confianza para la media de una distribución normal a partir de un valor computado \bar{x}_0 . Asumimos que la variancia de la distribución σ^2 es conocida. De ecuaciones 2.168 y 2.69 tenemos

$$\begin{aligned} p(x < x_a) &= \int \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{(x-\mu)^2}{2\sigma^2} \right] dx \\ &= \left(\frac{x_a - \mu}{\sigma} \right) \end{aligned} \quad (2.157)$$

En forma similar puesto que \bar{x} tiene variancia σ^2/n

$$\begin{aligned}
 p(\bar{x} < x_a) &= \int_{-\infty}^{x_a} \frac{1}{\sqrt{2\pi} \sigma/\sqrt{n}} \exp \left[-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n} \right] d\bar{x} \\
 &= \Phi \left(\frac{x_a - \mu}{\sigma/\sqrt{n}} \right) \quad (2.158)
 \end{aligned}$$

El estimado de μ es \bar{x}_0 y podemos escribir

$$p(\bar{x} < \bar{x}_0 : \mu = \bar{\mu}) = \Phi \left(\frac{\bar{x}_0 - \bar{\mu}}{\sigma/\sqrt{n}} \right) = P_1$$

$$u_{P_1} = \frac{\bar{x}_0 - \bar{\mu}}{\sigma/\sqrt{n}}$$

de tal manera

$$\bar{\mu} = \bar{x}_0 - u_{P_1} \frac{\sigma}{\sqrt{n}} \quad (2.159)$$

para obtener μ tendremos

$$p(\bar{x} > \bar{x}_0 : \mu = \bar{\mu}) = 1 - P_2$$

y

$$p(\bar{x} < \bar{x}_0 : \mu = \bar{\mu}) = P_2$$

entonces

$$p(\bar{x} < \bar{x}_0 : \mu = \bar{\mu}) = \Phi \left(\frac{\bar{x}_0 - \bar{\mu}}{\sigma/\sqrt{n}} \right)$$

$$u_{P_2} = \frac{\bar{x}_0 - \mu}{\sigma/\sqrt{n}}$$

es decir

$$\mu = \bar{x}_0 - u_{P_2} \frac{\sigma}{\sqrt{n}} \quad (2.160)$$

Los límites expresados en ecuaciones 2.159 y 2.160 se escriben más comúnmente así:

$$\begin{aligned} p(\bar{x} - u_{P_2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - u_{P_1} \frac{\sigma}{\sqrt{n}}) \\ = P_2 - P_1 \quad (2.161) \end{aligned}$$

PARTE III : APLICACION A LA METEOROLOGIA

III.1 Introducción

La definición clásica de Meteorología dice "la parte de la física que estudia la atmósfera y los fenómenos que en ella ocurren". Si los fenómenos ocurren en períodos cortos se estará hablando del "tiempo". Si por el contrario se habla del estado físico de la atmósfera dentro de períodos largos nos estaremos refiriendo al "clima".

El análisis estadístico es aplicable a la Meteorología si los datos son arreglados en forma tal que formen series climatológicas porque en este caso las secuencias de ellos se comportan como variables aleatorias. Se define una serie climatológica como una serie de muestras de un fenómeno meteorológico, cada muestra tomada por año.

III.2 Información en Meteorología

La escala en tiempo y espacio de los fenómenos meteorológicos es prácticamente continua; en tiempo va desde unos cuantos segundos (oscilaciones turbulentas muy rápidas) hasta miles de años con los cambios seculares del clima. - En espacio varía desde los movimientos casi moleculares que son del orden de milímetros hasta los movimientos de los sistemas planetarios que son del orden de miles de kilómetros. Hay

sin embargo efectos que por cuestiones de empirismo en las medidas son despreciados lo cual dá lugar a que la información meteorológica tenga un cierto caracter estadístico.

La información meteorológica que llega a un servicio puede dividirse en dos clases principalmente. Aquella que se obtiene varias veces al día por teletipos ú otros medios llamada información "presente" y la información coleccionada por los climatólogos en forma de cuadros, anuarios, gráficos, etc. que podría llamarse "pasada". Toda esta información es traducida por las diferentes ramas del servicio en datos útiles para cada propósito. Así por ejemplo la rama de pronósticos elaborará predicciones de elementos meteorológicos, de campos de presión, etc. Mientras que la rama de Climatología se encargará de producir pronósticos de clima, etc. La estadística es pues aplicada primordialmente en el campo climatológico siendo especialmente importante en nuestro medio el estudio de las técnicas mas avanzadas de la Estadística por cuanto vivimos en una región que cuenta con pocas estaciones que tengan registros antiguos. La estadística es entonces la mejor herramienta de que podemos disponer para sacar el provecho óptimo a los datos que actualmente pueden obtenerse.

III.3 Colección, Concentración y Escrutinio de la información Meteorológica

La información que se recibe en un servicio Meteorológico necesita ser procesada y luego concentrada. Esto implica eliminar la información que no sea necesaria para los propósi

tos que se persiguen y también simplificar los datos útiles. Es importante recordar que en Climatología el tiempo es a menudo la variable que mas interesa; por ello es necesario saber en todo momento la relación de tiempo—espacio contra el ó los fenómenos estudiados.

Los datos obtenidos pueden ser aplicados en dos formas. En primer lugar deseamos obtener una relación descriptiva de los fenómenos lo cual puede lograrse en forma concentrada por medio de tablas ó gráficos; por otra parte necesitamos establecer predicciones sobre el comportamiento de los fenómenos. Esto se logra por computaciones que dan lugar a parámetros representativos ó a modelos matemáticos.

Considérese en primer término la representación de un solo fenómeno a través del tiempo tal como está en la figura 19 que muestra las temperaturas mínimas en Enero para los años 1950-54 en la ciudad de Guatemala

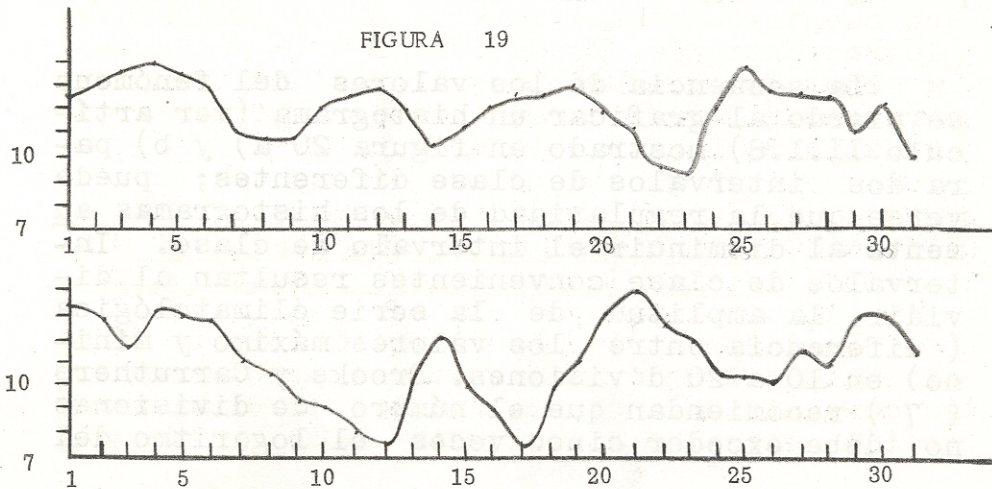
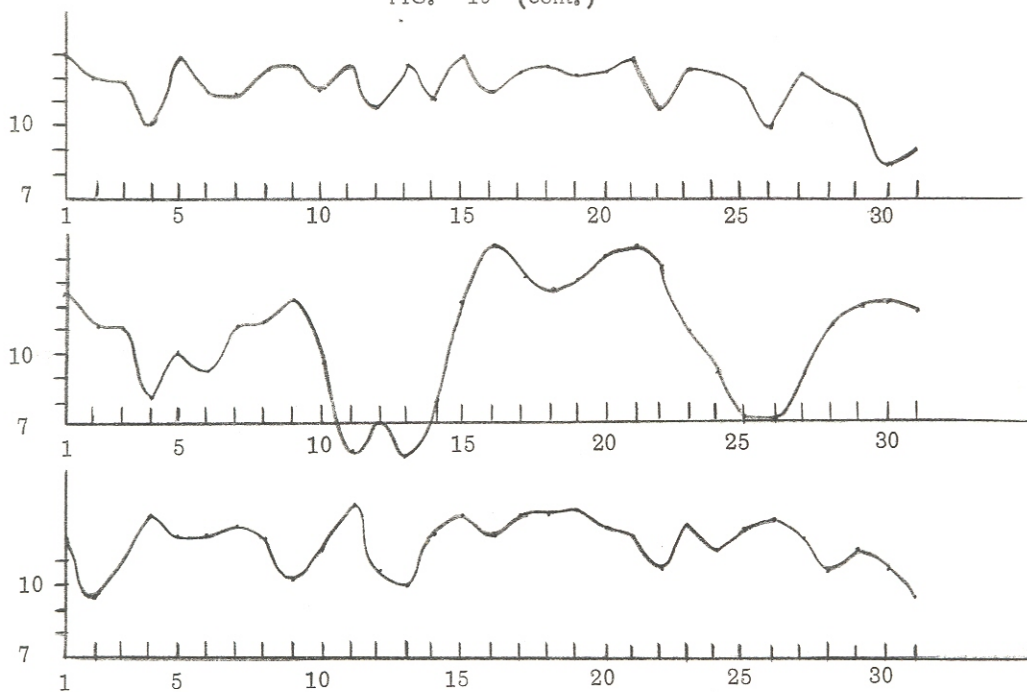


FIG. 19 (cont.)



La secuencia de los valores del fenómeno se pierde al graficar un histograma (ver artículo II.1.8) mostrado en figura 20 a) y b) para dos intervalos de clase diferentes; puede verse que la regularidad de los histogramas aumenta al disminuir el intervalo de clase. Intervalos de clase convenientes resultan al dividir la amplitud de la serie climatológica (diferencia entre los valores máximo y mínimo) en 10 a 20 divisiones. Brooks y Carruthers (7) recomiendan que el número de divisiones no debe exceder cinco veces el logaritmo del

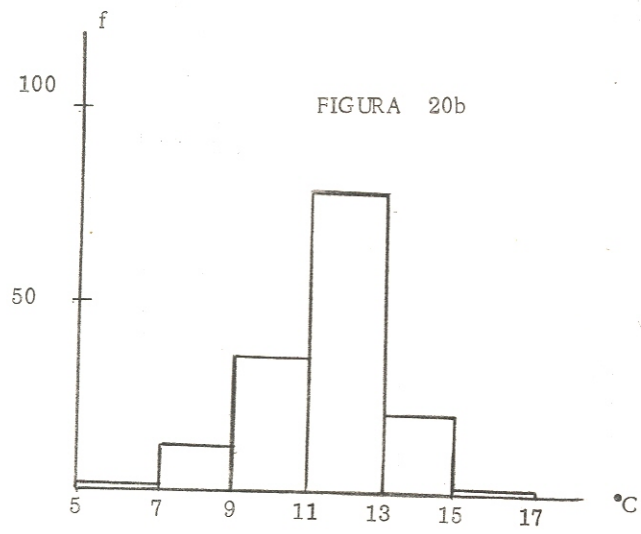
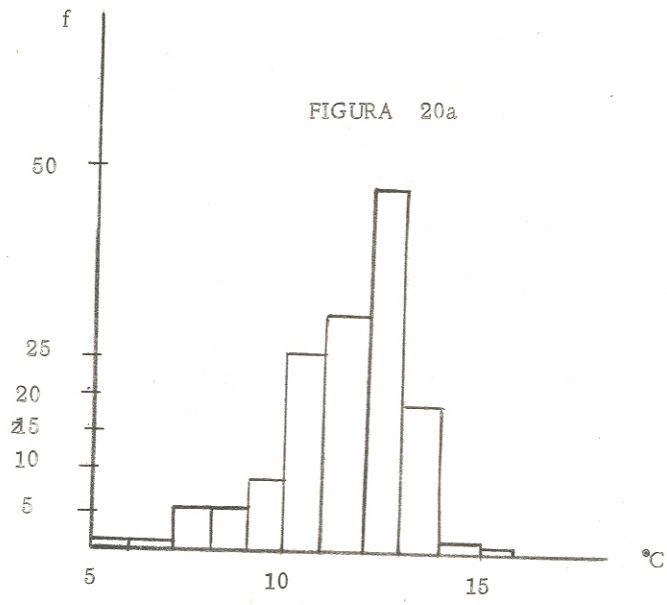
número de observaciones. Para nuestro caso $n=155$ y debieramos usar mas ó menos once clases lo que corresponde a un intervalo de 1°C .

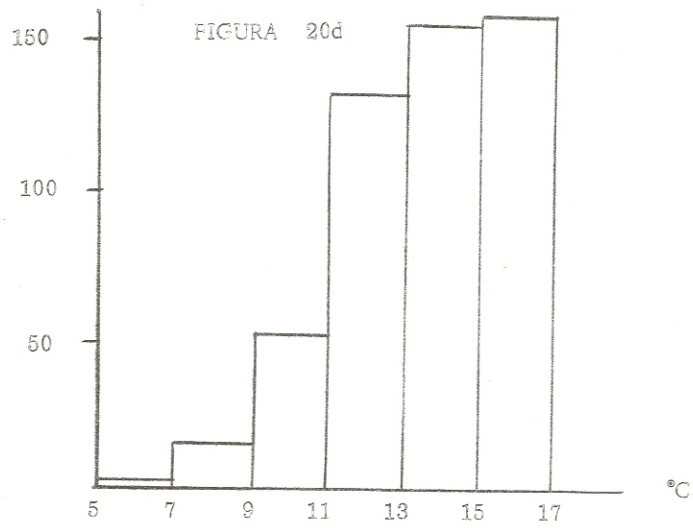
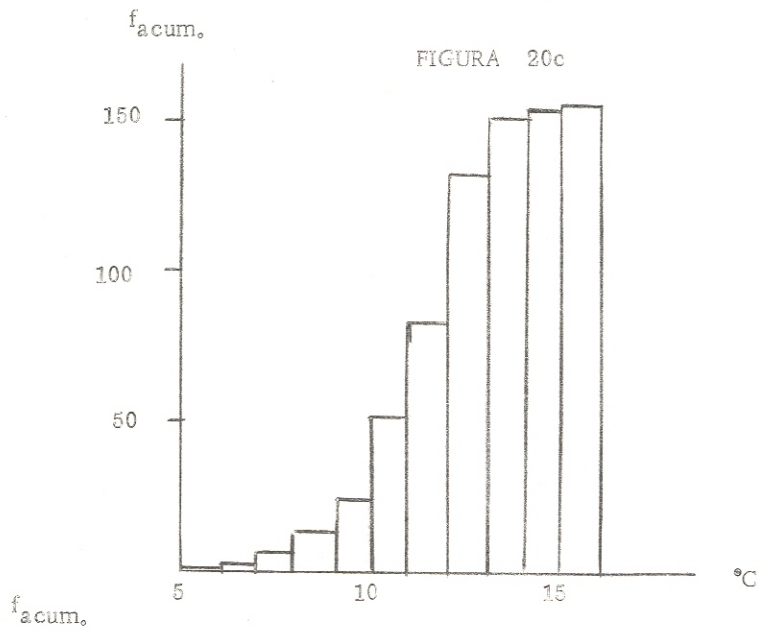
Si estuviéramos interesados en obtener estimados de frecuencia relativa ó probabilidad sobre varios intervalos de clase utilizaríamos tablas ó diagramas acumulados. Figura 20 c) y d) muestra los diagramas de frecuencias acumuladas correspondientes al cuadro 4

intervalo	f	frec acum.
5 - 5.9	1	1
6 - 6.9	1	2
7 - 7.9	6	8
8 - 8.9	6	14
9 - 9.9	10	24
10 -10.9	27	51
11 -11.9	32	83
12 -12.9	49	132
13 -13.9	20	152
14 -14.9	2	154
15 -15.9	1	155

intervalo	f	frec. acum.
5 - 6.9	2	2
7 - 8.9	12	14
9 -10.9	37	51
11 -12.9	81	132
13 -14.9	22	154
15 -16.9	1	155

CUADRO 4





La representación de cantidades vectoriales tal como el viento puede hacerse diagramáticamente en la forma de la figura 21 que corresponde al porcentaje promedio de ocurrencia de vientos calmados y las 16 direcciones de la rosa de los vientos, para la ciudad de Guatemala durante el período 1957-67 con una observación horaria. El grosor de barras indica la intensidad promedio en nudos. La tabulación está en el cuadro 5

CUADRO 5 Frecuencia de ocurrencia de vientos de Mayo

	N	NNE	NE	ENE	E	ESE	SE	SSE
% D	43.0	49.6	65.5	71.4	73.4	100.0	77.8	65.5
% M	55.3	49.2	33.6	28.6	20.0	0.0	22.2	34.5
% F	1.7	1.2	0.9	0.0	6.6	0.0	0.0	0.0
% f	26.8	20.4	3.7	0.4	0.2	0.1	0.4	1.1

	S	SSO	SO	OSO	O	ONO	NO	NNO
% D	57.9	62.0	84.6	74.1	100.0	00.0	85.7	55.7
% M	41.7	38.0	15.4	25.8	0.0	100.0	14.3	42.3
% F	0.4	0.0	0.0	0.0	0.0	0.0	0.0	2.0
% f	13.2	3.5	1.0	0.3	0.1	0.1	0.2	1.3

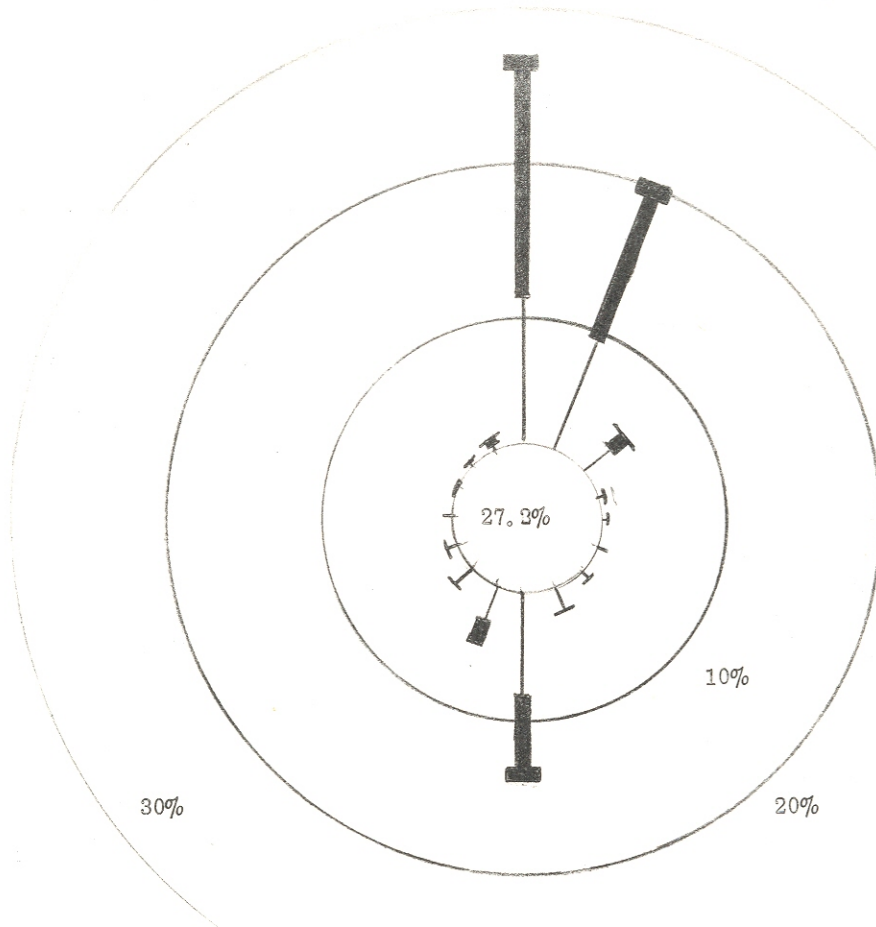


Figura 21

Leyenda

- | = Vientos débiles (D) son los menores de 10 nudos
- █ = Vientos moderados (M) son los que están entre 10 y 20 nudos
- █ = Vientos fuertes (F) son los mayores de 20 nudos

Nota: la cantidad en el centro del círculo representa el porcentaje de ocurrencia de viento en calma.

III.4 Uso de Parámetros como cantidades representativas

Para propósitos de concentración, los valores de una serie climatológica son a menudo sustituidos por un pequeño número de parámetros representativos de la serie ó "estadísticas" de la serie. La estadística quizás mas importante es aquella que dá un valor central al rededor del cual los diferentes valores de la serie están concentrados. De estas estadísticas la mas usada es la media aritmética, valor medio ó expectativa explicada en el artículo II.1.10. La mediana y la moda como medidas de tendencia central están someramente definidas al final de ese artículo. Otras dos medidas muy ligadas a la media aritmética son la media geométrica definida por

$$M_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \quad (3.1)$$

La media armónica M_h de $x_1, x_2 \dots x_n$ se define por

$$\frac{1}{M_h} = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i} \quad (3.2)$$

Ambos valores no son recomendados como medidas generales de tendencia central ya que pierden su sentido al tener la variable la posibilidad de tomar el valor cero. Sin embargo la media armónica es utilizada en relación con

la distancia entre isobaras ya que su media aritmética corresponde a la media aritmética del gradiente de presión.

Si la media aritmética se escoge como valor de tendencia central, una medida práctica de dispersión es la variancia definida en artículo II.1.11 aunque es más utilizada su raíz cuadrada (desviación estándar) ya que tiene la misma dimensional de la variable.

Algunas distribuciones en Meteorología son asimétricas y conviene en estos casos determinar el grado con que la distribución se desvía de otra considerada simétrica lo que se logra con el parámetro adimensional

$$c_1 = \frac{m_3}{s^3} \quad (3.3)$$

en donde m_3 es el tercer momento respecto a la media

$$m_3 = \frac{1}{n} \sum_1^n (x - \bar{x})^3 = \sum_1^n (x - \bar{x})^3 \cdot p(x) \quad (3.4)$$

y s la desviación estándar.

Por último para medir el grado de agudez ó kurtosis de una distribución se utiliza el parámetro adimensional

$$c_2 = \frac{m_4}{s^4} \quad (3.5)$$

en donde m_4 es el cuarto momento respecto a la media aritmética.

III.4.1 Regresión y Correlación

Hasta aquí se ha graficado en un sistema de coordenadas una variable en función del tiempo, sin embargo en un mismo instante de tiempo las medidas de varias variables climatológicas son tomadas y se trata ahora de encontrar la relación entre esas variables. Mas específicamente se trata de encontrar la relación matemática entre dos variables a partir de datos observados. La ecuación que define esta relación se llama "ecuación de la línea de regresión"

El método usado para ajuste de curvas es el de los mínimos cuadrados. Se trata de obtener $y = f(x)$ tal que la función $\sum (y - y_0)^2$ sea un mínimo. Se establece pues

$$\sum [f(x) - y_0]^2 = g(k_1 \dots k_n) \quad (3.6)$$

donde $y_0 = "y"$ observado y k_1, \dots, k_n son los parámetros $y = f(x)$ que se tratan de encontrar.

Hay que notar que debemos establecer primero la función general que se va a usar, es decir $y = f(x)$. Para obtener los n parámetros se necesitan n ecuaciones que se establecen derivando parcialmente respecto a cada uno de los parámetros e igualando a cero (en este ca-

so las variables son k_1, \dots, k_n y no "x" e "y". Para el caso particular en que la relación entre x e "y" sea lineal, la ecuación general es $y = mx + b$; los parámetros que hay que determinar son m (pendiente de la recta), y "b" (ordenada en el origen). Tendremos entonces de 3.6

$$\sum (mx + b - y_0)^2 = g(m, b) \quad (3.7)$$

las ecuaciones

$$\frac{\partial g(m, b)}{\partial b} = 0 \quad (3.8)$$

$$\frac{\partial g(m, b)}{\partial m} = 0 \quad (3.9)$$

permiten obtener m y b en función de los datos observados.

Al igual que con las distribuciones de frecuencia, una medida del grado de dispersión de los datos respecto a la línea de regresión está dada por el "error estándar de estimación"

$$S_{y.x} = \sqrt{\frac{\sum (y - y_{est})^2}{N}} \quad (3.10)$$

que tiene propiedades análogas a los de la desviación estándar. Para N pequeño un mejor estimado de $S_{y.x}$ se obtiene por

$$\hat{s}_{y.x} = \sqrt{\frac{\sum (y - y_{est})^2}{N-2}} \quad (3.11)$$

Por otro lado el grado de asociación de dos variables se representa por la covariancia y por el coeficiente de correlación explicados en el artículo II.1.14. Cabría agregar lo siguiente: si desviaciones grandes (positivas ó negativas) de x están asociadas con desviaciones grandes del mismo signo de y, la covariancia y el coeficiente de correlación son positivos. Aún más si desviaciones positivas de x están asociadas con desviaciones negativas de "y", y viceversa entonces la covariancia y el coeficiente de correlación son negativos. Por último si desviaciones positivas y negativas de x son igualmente frecuentes con desviaciones positivas y negativas de "y", entonces la covariancia y el coeficiente de correlación tenderán a cero.

III.4.2 Asociación y Contingencia

En el artículo II.1.13 se explicó en forma breve como agrupar dos variables aleatorias según una distribución específica. Sin embargo, las probabilidades (ó frecuencias relativas) esperadas usualmente no van a coincidir con las observadas, y en esta sección se esbozarán algunos aspectos sobre el grado de asociación entre las dos frecuencias.

Una medida del grado de discrepancia entre las frecuencias observadas y esperadas está dado por la estadística χ^2 (chi-cuadrado)

$$\chi^2 = \sum_1^n \frac{(f_o - f_e)^2}{f_e} \quad (3.12)$$

siendo f_o la frecuencia observada y f_e la frecuencia esperada, la distribución muestral de χ^2 se aproxima bastante a la distribución de chi-cuadrado que se verá mas adelante. Por otro lado, a una tabulación como la del cuadro 3 del artículo II.1.13 se le llama tabla de contingencia. $p(x_i \cap y_i)$ puede ser la probabilidad esperada ó la observada. El grado de asociación en una tabla de contingencia se mide por el coeficiente de contingencia definido por

$$c = \sqrt{\frac{\chi^2}{\chi^2 + N}} \quad (3.13)$$

cuyo valor es siempre menor de uno. Mientras mas cerca esté c de uno mayor será el grado de relación que proporciona la tabla.

III.5 La distribución de Poisson

Quando en la distribución binomial (II.1.7) n se vuelve muy grande y θ muy pequeño con expectativa $n\theta = \lambda$ constante, entonces la binomial tiende a la distribución de Poisson y en el límite se obtendrá

$$p_n(x) = \lambda^x \frac{e^{-\lambda}}{x!} \quad (3.14)$$

su forma acumulada es

$$F(x) = \sum_{t=0}^x \lambda^t \frac{e^{-\lambda}}{t!} \quad (3.15)$$

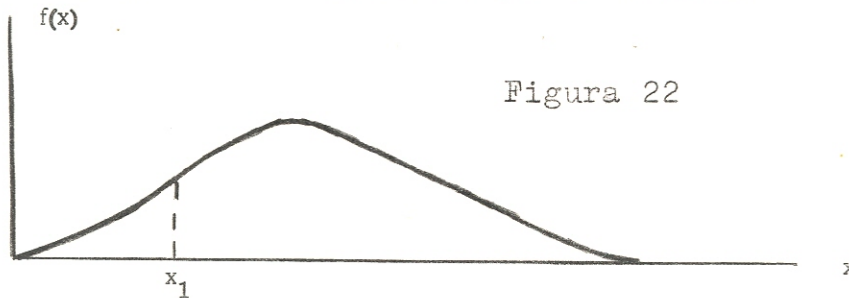
La distribución de Poisson se aplica entonces a eventos "raros" es decir con poca probabilidad por ejemplo la frecuencia de precipitación anual en forma de granizo, eventos de precipitación excesiva, frecuencia de tornados, etc.

III.6 La distribución gamma y chi-cuadrado

La distribución gamma es útil para ajustar variables meteorológicas restringidas a valores positivos tal como precipitación, presión del vapor, evaporación etc. Esta basada fundamentalmente en la función gamma definida por la integral

$$\Gamma(z) = \int_0^{\infty} x^{z-1} \cdot e^{-x} \cdot dx \quad (3.16)$$

y representa el área de la curva de la figura 22 entre los límites cero e infinito



la forma de la curva varía en función de z , por tanto $\Gamma(z)$ variará así como z varíe. En

la tabla VII del apéndice se presenta una tabla $\Gamma(z)$ para valores de z entre 1 y 2. Para valores de z mayores de 2 se usa la fórmula de recurrencia

$$\Gamma(n/z) = (n-1/z)(n-2/z) \dots (1/z) \Gamma(1/z) \quad (3.17)$$

cuya aplicación se verá luego. $\Gamma(z)$ puede expresarse mediante la notación

$$f(x)_z = x^{z-1} \cdot e^{-x} = \text{función de } x \text{ para un valor de } z.$$

$$\Gamma(z) = \int_0^{\infty} f(x)_z \cdot dx$$

la función

$$\Gamma(z) = k^z \int_0^{\infty} x^{z-1} \cdot e^{-kx} \cdot dx \quad (3.18)$$

donde k es constante. 3.18 es equivalente a 3.16. Establezcamos una nueva variable

$$y = kx \quad (3.19)$$

sustituyendo 3.19 en 3.18

$$\begin{aligned} \Gamma(z) &= k^z \int_0^{\infty} (y/k)^{z-1} \cdot e^{-y} \cdot d(y/k) \\ &= k^0 \int_0^{\infty} y^{z-1} \cdot e^{-y} \cdot dy \\ &= \int_0^{\infty} y^{z-1} \cdot e^{-y} \cdot dy \end{aligned} \quad (3.20)$$

Notar que los límites cero e infinito no varían y puesto que $\Gamma(z)$ depende solo de z , 3.20 y 3.16 son iguales

Si en 3.16 cambiamos el límite superior por uno finito " x_1 ", la función se llama gamma incompleta y se representa por

$$\gamma(z, x) = \int_0^{x_1} e^{-x} \cdot x^{z-1} \cdot dx \quad (3.21)$$

representando la parte del área de la curva de la figura 22 entre cero y x_1 . Notar que 3.21 depende ya no solo de z sino de la variable x . El valor de 3.21 se calcula por la relación

$$\gamma(z, x) = \Gamma(z) \cdot P(\chi^2/v) \quad (3.22)$$

siendo

$$\chi^2 = 2x$$

$$v = 2z \quad (3.23)$$

$P(\chi^2/v)$ se definirá luego.

La resolución de 3.22 requiere que la función acumulada esté en la forma 3.21, por tanto una función como la 3.18 habrá de ponerse primero en la forma 3.20 cambiando el límite superior por $y_1 = kx_1$.

Para el caso especial en que, en la función 3.18 $k=1/2$ y $z = v/2$ la variable x se representa por χ^2 (chi-cuadrado) y 3.18 se convierte en

$$\Gamma(v/2) = \int_0^{\infty} \frac{1}{2^{v/2}} (\chi^2)^{\frac{v}{2}-1} \cdot e^{-\chi^2/2} \cdot d(\chi^2) \quad (3.24)$$

y representa el área total de la curva de la figura 22 siendo el único cambio el hecho de que $x = \chi^2$ y $f(x) = f(\chi^2)$. La función 3.24 puede ponerse en la forma 3.20 con solo hacer $y = \chi^2/2$. La función

$$\gamma(v/2, \chi^2/2) = \gamma(v/2, y)$$

$$= \int_0^y e^{-y} \cdot (y)^{\frac{v}{2}-1} \cdot dy \quad (3.25)$$

tiene exactamente la misma interpretación que 3.21; la relación

$$P(\chi^2/v) = \frac{\gamma(v/2, \chi^2/2)}{\Gamma(v/2)} \quad (3.26)$$

representa la proporción del área total $\Gamma(v/2)$ que ocupa $\gamma(v/2, \chi^2/2)$; además

$$Q(\chi^2/v) = 1 - P(\chi^2/v) \quad (3.27)$$

es la proporción del área total que NO ocupa $\gamma(v/2, \chi^2/2)$. En tabla IV están los valores de $Q(\chi^2/v)$ para diferentes valores de χ^2 y v .

Puede verse pues que 3.25 (distribución a

cumulada de chi-cuadrado) no es mas que una forma especial de la función gamma incompleta (3.21). Antes de realizar un ejemplo numérico cabe agregar dos cosas: la estadística

$$\chi^2 = \frac{Ns^2}{\sigma^2} = \frac{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{\sigma^2} \quad (3.28)$$

sacada de x_1, \dots, x_n observaciones de una variable x tiene la distribución dada por 3.25. Por otro lado en tabla IV, valores de $Q(\chi^2/v)$ con $\chi^2 \geq 30$ pueden aproximarse por la relación

$$Q(\chi^2/v) \approx Q(x) \quad (3.29)$$

siendo

$$x = \frac{(\chi^2/v)^{1/3} - (1 - 2/9v)}{\sqrt{2/9v}} \quad (3.30)$$

siendo $Q(x) = 1 - \Phi(u)$; (ver ec. 2.32)

Vamos a aplicar lo visto en esta sección a las observaciones mensuales de precipitación del mes de Junio en el período 1931-1960 para la ciudad de Guatemala. Recordando el método de máxima verosimilitud (II.2.3) y aplicándolo a $f(x)$ de 3.18 $[f(x) = k^z x^{z-1} e^{-kx}]$, obtendremos los estimados de z y k :

$$\hat{z} = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad (3.31)$$

$$\hat{k} = \frac{\hat{z}}{\bar{x}} \quad (3.32)$$

donde A está dado por

$$A = \ln \bar{x} - \frac{\sum \ln x}{n} \quad (3.33)$$

En el cuadro siguiente están tabulados los valores de precipitación (x) y su logaritmo natural

JUNIO

año	(x)	ln x	año	(x)	ln x
1931	430.0	6.06379	1946	164.1	5.10048
1932	285.7	5.65494	1947	267.6	5.58949
1933	348.2	5.85277	1948	319.3	5.76613
1934	218.2	5.38541	1949	158.9	5.06827
1935	266.2	5.58425	1950	263.4	5.57367
1936	405.8	6.00586	1951	149.1	5.00462
1937	215.3	5.37203	1952	316.6	5.75764
1938	233.1	5.45147	1953	227.3	5.42759
1939	164.1	5.10048	1954	377.6	5.93383
1940	321.4	5.77268	1955	144.8	4.97535
1941	244.2	5.49799	1956	465.4	6.14290
1942	175.7	5.16879	1957	215.8	5.37435
1943	136.7	4.91768	1958	373.9	5.92399
1944	191.4	5.25436	1959	271.4	5.60359
1945	296.9	5.69339	1960	299.7	5.70344

CUADRO 6

tendremos entonces

$$A = 0.05531; \quad \bar{x} = 264.927$$

$$\hat{z} = 9.2061$$

$$\hat{k} = \frac{9.2061}{264.927} = 0.03475$$

con los parámetros de la distribución podemos contestar preguntas como cual es la probabilidad de que en Junio precipite menos de tantos mm?; por ej. que llueva menos de 100 mm?. La variable $x_1 = 100$ (3.18) se transforma en $y_1 = kx_1 = 3.475$; 3.21 será entonces

$$\gamma(9.2061, 3.475) = \int_0^{3.475} e^{-y} \cdot y^{8.206} \cdot dy$$

y de 3.17

$$\Gamma(z) = (8.206)(7.206) \dots (1.206) \Gamma(1.206)$$

$$= 68537.365 \times 0.9169 = 62841.91$$

$$\text{además } P(\chi^2/v) = 1 - Q(\chi^2/v) = 1 - 0.99176$$

$$= 0.00824$$

y la probabilidad de que en Junio llueva menos de 100 mm es apenas 0.0082 (ó 0.82%).

Hay que notar que $\Gamma(z)$ no es necesario calcularlo ya que la razón $\gamma(z, x) / \Gamma(z)$ es precisamente $P(\chi^2/v)$

III.7 La distribución "t" ó de "Student"

La distribución de "Student" es continua y tiene función de densidad de probabilidad

$$f(t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi v} \Gamma(v/2)} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \quad (3.34)$$

donde la variable "t" está definida por la estadística

$$t = \frac{\bar{x} - \mu}{s} \sqrt{N-1} \quad (3.35)$$

con $N-1$ grados de libertad

ó bien por la estadística

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{N_1 s_1^2 + N_2 s_2^2}} \sqrt{\frac{N_1 N_2 (N_1 + N_2 - 2)}{N_1 + N_2}}$$

ó

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \quad (3.36)$$

con $N_1 + N_2 - 2$ grados de libertad

donde

$$\sigma = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}} \quad (3.37)$$

siendo \bar{x}_1 , s_1^2 y N_1 la media, variancia y tamaño de la muestra uno respectivamente y lo mismo para \bar{x}_2 , s_2^2 y N_2 provenientes ambas muestras del mismo universo.

La distribución de "Student" tiende a la distribución normal estandarizada cuando v aumenta; para $v \geq 30$ prácticamente no hay diferencia entre las dos distribuciones. El parámetro " v " depende del número de observaciones hechas (es decir del tamaño de la muestra) y del número de parámetros que deben ser estimados a partir de la muestra. Mas específicamente es su diferencia, es decir

$$v = N - k \quad (3.38)$$

donde N = tamaño de la muestra y k = número de parámetros estimados. A " v " se le llama "número de grados de libertad" de la estadística. Por ejemplo para la primera estadística de " t " solo tenemos que estimar " μ ", por tanto el número de grados de libertad es $N-1$. Igualmente para la estadística χ^2 solo habremos de estimar σ^2 , entonces tendremos $v = N-1$.

La relación entre las distribuciones normal, chi-cuadrado y "Student" está dada por el siguiente teorema: "Si u tiene una distribución normal estandarizada, mientras v^2 , independiente de u tiene una distribución chi-cuadrado con " v " grados de libertad, entonces $u/v \cdot \sqrt{v}$

tiene una distribución "t" con v grados de libertad".

III.8 La distribución de Fisher ó distribución F.

La distribución con densidad de probabilidad

$$f(F) = \left(\frac{v_1}{v_2} \right)^{v_1/2} \frac{\Gamma\left(\frac{v_1 + v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right) \cdot \Gamma\left(\frac{v_2}{2}\right)}$$

$$\frac{F^{\frac{v_1-2}{2}}}{\left(1 + \frac{v_1}{v_2} \cdot F\right)^{\frac{v_1+v_2}{2}}} \quad (3.39)$$

es llamada distribución de Fisher con v_1 y v_2 grados de libertad. La variable F se define por la estadística

$$F = \frac{N_1 S_1^2 (N_2 - 1)}{N_2 S_2^2 (N_1 - 1)} \quad (3.40)$$

con $(N_1 - 1)$ y $(N_2 - 1)$ grados de libertad.

Esta distribución es muy útil en el análisis de variancias y temperaturas como se verá mas adelante.

III.9 Series de Tiempo

Una serie de tiempo es una secuencia de valores de una variable aleatoria tomados a través del tiempo. Al analizar una serie se trata de obtener información respecto a sus propiedades básicas tales como variaciones seculares (es decir su tendencia general) y oscilaciones. El objetivo primordial es, sin embargo pronosticar el comportamiento de la serie de tiempo en el futuro.

En general pueden clasificarse los movimientos de una serie en regulares ó periódicos e irregulares ó estocásticos. El análisis de los primeros se efectúa haciendo uso de las series de Fourier y es objeto de estudios mas avanzados; por ello no se estudiará en el presente trabajo. Unicamente se expondrán en forma breve algunos de los aspectos mas simples para determinar tendencias. En primer término definiremos un proceso estocástico como aquel que tiene covariancia cero (ver II.1.14) es decir que no existe una relación definida entre la variable y el tiempo. Estos procesos son originados por deficiencia de los instrumentos de medición ó por errores en las medidas y producen movimientos irregulares en la serie. Supongamos ahora que se tiene una secuencia de variables x_1, x_2, \dots, x_n . La secuencia formada por

$$\frac{x_1 + x_2 + \dots + x_n}{n}, \frac{x_2 + x_3 + \dots + x_{n/1}}{n}, \dots \quad (3.41)$$

se llama promedio móvil de orden n de la variable; si la variable está dada por día (por ejemplo la temperatura mínima diaria) podemos construir promedios móviles de 2, 3, etc. días

(por supuesto el promedio móvil de grado 1 es la secuencia misma); lo mismo puede hacerse si los datos son mensuales ó anuales. Notar que cada valor del promedio móvil se coloca equidistante de los valores extremos, por ejemplo el primer valor iría en el lugar $(x_n + x_1)/2$, el segundo en el $(x_{(n/1)} + x_2)/2$ y así sucesivamente.

Los promedios móviles se utilizan para de terminar la tendencia general de la serie así como sus variaciones debido a tener la propiedad de reducir las variaciones presentes en la secuencia. El orden del promedio móvil está pues limitado al tipo de variación que se desea representar. Por ejemplo si se desea representar la variación estacional de una variable dada mensualmente se computa primero un promedio móvil de orden 12; luego para que los valores del promedio móvil caigan en meses sucesivos en lugar de caer a mediados de cada mes se computa un promedio móvil de orden 2. La representación de este promedio dá la serie de tiempo sin las variaciones estacionales e irregulares, es decir representa la tendencia general y las variaciones cíclicas. Dividiendo los datos originales de la variable entre su respectivo promedio móvil y expresandolo como porcentaje llegamos a representar las variaciones estacionales en forma de índices. El promedio (ó en algunos casos la mediana) para cada mes se llama índice estacional.

III.10 Persistencia

La persistencia en Meteorología se refiere al grado de seguridad que se tiene de que un evento se realice en un intervalo de tiempo t (un día, un mes, etc) una vez que el evento

se ha llevado a cabo en el intervalo $t-1$. Una medida de la persistencia lo da el coeficiente de persistencia

$$r_p = 1 - \sqrt{\frac{1 - p_1}{1 - p}}^2 \quad (3.42)$$

donde p es la probabilidad de que el evento ocurra en el intervalo t y p_1 es la probabilidad de que el evento ocurra en el intervalo t si ya ha ocurrido en el intervalo $t-1$. Podríamos expresar estas probabilidades de acuerdo al artículo de probabilidad condicional (II.1.3) así

$$p = p(E_t) \quad (3.43)$$

y

$$p_1 = p(E_t/E_{t-1}) \quad (3.44)$$

r_p se expresaría entonces

$$r_p = 1 - \sqrt{\frac{1 - p(E_t/E_{t-1})}{1 - p(E_t)}} \quad (3.45)$$

r_p es cero si $p_1 = p$, es decir si la ocurrencia de E_t no nos dice nada respecto a la ocurrencia de E_{t-1} ; por el contrario $r_p = 1$ si $p_1 = 1$, es decir si la ocurrencia de E_t asegura la ocurrencia de E_{t-1} . Cuando $p_1 < p$, r_p se vuelve negativo teniendo el evento una tendencia a oscilar. En general r_p se usa para valores de p_1 mayores de p .

III.11 Interpretación de datos estadísticos

en Meteorología

En la introducción de este capítulo se explicó la forma en que deben agruparse los datos para que el análisis estadístico sea válido. En muchos casos no obstante las muestras son construidas a partir de datos sucesivos ó casi sucesivos con lo cual se introduce un grado de persistencia en los datos de la muestra contraviniendo los principios del análisis estadístico. Por ello siempre que sea posible deberán evitarse las variaciones condicionales y si esto no es posible su existencia debe ser reconocida y los errores que puedan tener las conclusiones sacadas de este tipo de muestras deberá tomarse en cuenta.

III.11.1 Pruebas respecto a la "bondad" del ajuste

Vamos a averiguar en esta sección que tan bueno fué el ajuste de la distribución gamma para los datos de precipitación del cuadro 6. Lo hacemos probando la hipótesis "el ajuste es bueno" contra la alterna "el ajuste no es tan bueno" por medio de la estadística chi-cuadrado (III.4.2). En el cuadro siguiente están las probabilidades observadas y las esperadas para intervalos de precipitación de 50 mm.

CUADRO 7

Pcpn.	150	200	250	300	350	400	450	450
P.Obs.	.10	.167	.20	.233	.133	.067	.067	.033
P.Esp.	.08	.170	.18	.260	.150	.080	.050	.029

La estadística χ^2 (ec. 3.12) es

$$\begin{aligned} \chi^2 = & \frac{(0.10-0.08)^2}{0.08} / \frac{(.167-0.170)^2}{0.17} / \\ & \frac{(0.20-0.18)^2}{0.18} / \frac{(.233-.260)^2}{0.26} / \frac{(.133-.150)^2}{0.15} \\ & / \frac{(0.067-0.080)^2}{0.08} / \frac{(0.067-0.05)^2}{0.05} / \\ & / \frac{(0.033-0.029)^2}{0.029} = 0.022 \end{aligned}$$

Dos parámetros fueron estimados en la distribución, el número de grados de libertad es entonces $8-1-2=5$. En tabla IV del apéndice puede verse que valores mayores de .022 son excedidos prácticamente en el 100% de los casos y podemos concluir que el ajuste es no solo "bueno" sino "muy bueno".

III.11.2 Análisis estadístico de temperaturas de la ciudad de Guatemala para un mes dado

El cuadro 8 presenta las temperaturas medias del mes de Mayo para el período 1928-67. Para los propósitos del siguiente análisis tomaremos el período 1932-56 como "normal" con promedio 19.5°C y desviación estándar 0.62°C

CUADRO 8

Temperaturas ambiente de Mayo para
la ciudad de Guatemala

	1	2	3	4	5
1920					
1930	19.3	20.8	20.2	19.7	20.0
1940	19.7	18.8	19.0	19.3	19.1
1950	18.7	19.1	20.0	18.7	19.3
1960	19.8	18.9	19.8	19.8	20.3
	6	7	8	9	10
1920			18.9	18.7	19.8
1930	20.2	19.6	20.3	20.3	19.9
1940	20.4	19.3	18.9	19.9	19.2
1950	18.5	19.7	20.6	19.7	20.7
1960	19.6	19.8			

asumiendo que las temperaturas ambiente de Mayo tienen una distribución $N(19.5, 0.62)$ podemos inferir lo siguiente

a) en $2/3$ de los casos la temperatura de Mayo está dentro del intervalo 20.1° - 18.9°

b) en un 95% de los casos la temperatura está dentro del intervalo 20.7° - 18.3° .

Puede verse que dentro del período "normal" un 60% de los casos está dentro del intervalo dado en a) y un 96% dentro del intervalo dado en b), estando bastante de acuerdo con la distribución teórica.

Las observaciones del intervalo 1928-31 con media 19.2° C y desviación estándar 0.42° C son dudosas y queremos determinar una prueba para saber si las incluimos ó no en el universo. Si la muestra fuera mayor, y dado que conocemos la media y la desviación estándar del universo, podríamos utilizar la estadística $(\bar{x} - \mu) / \sqrt{n} / \sigma$ que está normalmente distribuida; sin embargo como $n=4$ usaremos la estadística $(\bar{x} - \mu) / \sqrt{n-1} / s$ que tiene distribución "t" (ecuación 3.35) con $n-1$ grados de libertad (el único estimado es μ). Tendremos entonces

$$t = (19.2 - 19.5) / \sqrt{4-1} / 0.42 = -1.237$$

La probabilidad de obtener valores de $|t|$ mayores de 1.24 (tabla III del apéndice) es aproximadamente 0.3 y por tanto podemos aceptar la muestra como parte del universo.

Otra prueba para el mismo propósito se puede efectuar usando la estadística ns^2 / σ^2 (ecuación 3.28) en donde el único estimado que se necesita es σ^2 . La prueba puede plantearse así: si $s^2 = \sigma^2$ el valor de χ^2 es 4 y a un nivel α de 0.05 los valores extremos de χ^2 son 0.2 y 9.3. El valor de ns^2 / σ^2 con 3 grados de libertad es de 1.23 y nuevamente se acepta la muestra como parte del universo.

Supongamos ahora que deseamos investigar que tanta información respecto a la población puede proporcionar una muestra de 5 años, por ejemplo la del período 1928-32. El valor del promedio es 19.5 y la desviación estándar es $s = \sqrt{.564} = 0.75$. Usaremos nuevamente

$$t = \frac{\bar{x} - \mu}{s} \sqrt{n-1} = \frac{19.5 - \mu}{0.75} \sqrt{4} = \frac{19.5 - \mu}{0.375}$$

Utilizando un intervalo de confianza de 95% los límites de confianza son, de tabla III del apéndice, $t = \pm 2.776$; el intervalo para μ puede entonces expresarse así:

$$-2.776 \leq \frac{19.5 - \mu}{0.375} \leq 2.776$$

y resolviendo para μ tenemos el intervalo

$$18.5 \leq \mu \leq 20.5$$

Notar que, aunque casualmente el promedio de la muestra coincidió con el promedio de la población, nosotros solo disponíamos de la información de la muestra. De todas formas el intervalo resulta muy adecuado. El resto de promedios para intervalos sucesivos de 5 años hasta 1967 dá 19.9, 19.8, 19.4, 19.2, 19.2, 19.9 y 19.8. Puede verse que todos los valores están agrupados alrededor de 19.5 y cualquiera de ellos proporciona información respecto a μ .

La estadística $(\bar{x} - \mu) \sqrt{n} / \sigma$ funciona en pruebas de esta naturaleza cuando $n \geq 30$. Para $n < 30$ deja de estar normalmente distribuida. Además los intervalos que proporciona son mas angostos que los dados arriba; por ejemplo con

un intervalo de confianza de 95% los intervalos para μ en los periodos 1928-32 y 1933-37 son

$$19.2 \leq \mu \leq 19.8 \quad \text{y} \quad 19.4 \leq \mu \leq 20.4$$

Otro problema sería averiguar que información respecto a la variancia del universo puede darnos la variancia de la muestra s^2 . Para este propósito utilizamos la estadística ns^2/σ^2 ; con un intervalo de confianza de 95% los límites de confianza son 0.48 y 11.1 de tabla IV del apéndice, (aunque la curva de χ^2 no es simétrica, se han tomado áreas iguales en los dos extremos de la cola bilateral) el intervalo puede ponerse

$$0.48 \leq \frac{2.82}{\sigma^2} \leq 11.1$$

ó

$$0.254 \leq \sigma^2 \leq 5.87$$

y

$$0.5 \leq \sigma \leq 2.4$$

La desviación estándar de la población (0.62) está dentro del intervalo. Debe notarse que estos límites de confianza son muy abiertos por lo que no proporcionan mucha información respecto a la variancia del universo.

Complemento del análisis de temperaturas puede ser la respuesta a la pregunta ¿Pertenece dos muestras al mismo universo?. Tomamos

por ej. el período 1928-32 (promedio 19.5°C y desviación estándar $\sqrt{0.564} = 0.75$) y el de 1963-67 (promedio 19.8 y desviación estándar $\sqrt{.058} = 0.24$). Computemos la estadística dada por ecuación 3.40 considerando primero una prueba con solo las variancias:

$$F = \frac{n_1 s_1^2 (n_2 - 1)}{n_2 s_2^2 (n_1 - 1)} = 9.7$$

y tomando un nivel de significación de 1% en la cola de la derecha de la curva encontramos de tabla V del apéndice con $v_1 = v_2 = 4$ que $F_{1-\alpha} = 16$, luego aceptamos la hipótesis " las muestras pertenecen al universo " a un nivel de 1%. A un nivel de 5% obtendremos $F_{1-\alpha} = 6.4$ y entonces a este nivel rechazamos la hipótesis nula

La estadística dada en ecuación 3.36 puede ser usada para el mismo fin; haciendo $\mu_1 = \mu_2$ tendremos

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{-0.3}{0.624 \times 0.633} = -0.76$$

A un nivel de 5% con 8 grados de libertad los valores límites de t son ± 2.3 luego aceptamos la hipótesis nula.

III.11.3 Análisis de Variancia

El análisis de variancia responde a la pregunta ¿Pertenecen varias muestras al mismo universo?. Trata al mismo tiempo de determinar la forma en que las fuentes que producen variación contribuyen a la variancia total; además permite probar si el efecto de un factor particular es real ó es producto de errores aleatorios. Siendo el procedimiento aritmético se ilustrará con un ejemplo. El cuadro 9 presenta algunos de los datos del cuadro 8 separados en tres grupos, cada uno de cinco años sucesivos.

CUADRO 9 ANALISIS DE VARIANCIA

	1	2	3	
	1930	1940	1950	
1	19.3 (-0.7)	19.7 (0.5)	18.7 (-0.5)	
2	20.8 (0.8)	18.8 (-0.4)	19.1 (-0.1)	
3	20.2 (0.2)	19.0 (-0.2)	20.0 (0.8)	
4	19.7 (-0.3)	19.3 (0.1)	18.7 (-0.5)	
5	20.0 (0.0)	19.1 (-0.1)	19.3 (0.1)	
Σ	100.0	95.9	95.8	291.7
\bar{x}	20.0	19.2	19.2	19.5

Las cantidades entre paréntesis son las diferencias "d" de cada observación x de la media \bar{x} de su respectivo grupo.

Cualquier observación puede representarse en función de "d" y de la desviación "g" de la media del grupo \bar{x} respecto a la media general $\bar{\bar{x}}$ (19.5) en la siguiente forma

$$x = \bar{\bar{x}} + (\bar{x} - \bar{\bar{x}}) + (x - \bar{x})$$

$$x = \bar{x} + g + d \quad (3.46)$$

ó también

$$x' = x - \bar{x} = g + d$$

La variación de "x" puede dividirse en dos tipos: fluctuación aleatoria ó dentro de la muestra y fluctuación sistemática.

Las cantidades "g" son afectadas primordialmente por variaciones sistemáticas ya que las medias suavizan las fluctuaciones aleatorias de las observaciones originales; en menor grado son influenciadas por fluctuaciones aleatorias. Por otro lado, las cantidades "d" son producidas enteramente por fluctuaciones dentro de la muestra, esto es por variaciones aleatorias.

Elevando al cuadrado ambos miembros de 3.47 y sumando sobre toda la muestra tenemos:

$$\sum (x - \bar{x})^2 = \sum x'^2 = m \sum_g g^2 + \sum d^2 \quad (3.48)$$

en donde m es el número de observaciones por grupo y la g bajo una sumatoria significa suma sobre todos los grupos. En forma simbólica puede escribirse la ecuación 3.48 así

$$SC_T = SC_E + SC_D \quad (3.49)$$

SC significa "suma de cuadrados"; el subíndice T significa "total", E "entre grupos" y

D "dentro de grupos". Cada uno de los términos anteriores puede ponerse, para su cómputo de la siguiente forma:

$$\begin{aligned}
 SC_T &= \sum (x - \bar{x})^2 = \sum (x^2 - 2x\bar{x} + \bar{x}^2) \\
 &= \sum x^2 - 2\bar{x} \sum x + \sum \bar{x}^2 \\
 &= \sum x^2 - 2 \frac{(\sum x^2)}{N} + N \left(\frac{\sum x}{N} \right)^2 \\
 &= \sum x^2 - \frac{(\sum x)^2}{N} \quad (3.50)
 \end{aligned}$$

$$\begin{aligned}
 SC_E &= \sum (\bar{x} - \bar{x})^2 = m \sum_g (\bar{x} - \bar{x})^2 \\
 &= m \sum_g (\bar{x}^2 - 2\bar{x}\bar{x} + \bar{x}^2)
 \end{aligned}$$

poniendo $\bar{x} = \frac{\sum x}{m}$ donde el cero abajo de

la sumatoria significa suma dentro de un grupo tendremos

$$\begin{aligned}
 SC_E &= m \sum_g \left(\frac{\sum x}{m} \right)^2 - 2m \sum_g \left(\frac{\sum x}{m} \right) \left(\frac{\sum x}{N} \right) + \\
 &\quad m \sum_g \left(\frac{\sum x}{N} \right)^2 \\
 &= \frac{1}{m} \sum_g \left(\sum_0 x \right)^2 - 2 \frac{(\sum x)^2}{N} + \frac{(\sum x)^2}{N}
 \end{aligned}$$

$$SC_E = \frac{1}{m} \sum_g \left(\sum_0 x \right)^2 - \frac{\left(\sum_0 x \right)^2}{N} \quad (3.51)$$

por último

$$\begin{aligned} SC_D &= \sum (x - \bar{x})^2 = \sum (x^2 - 2x \bar{x} + \bar{x}^2) \\ &= \sum x^2 - 2 \sum x \left(\frac{\sum_0 x}{m} \right) + \sum \left(\frac{\sum_0 x}{m} \right)^2 \\ &= \sum_g \sum_0 x^2 - \frac{2}{m} \left(\sum_g \sum_0 x \right) \left(\sum_0 x \right) \frac{1}{m} \sum_g \left(\sum_0 x \right)^2 \\ &= \sum_g \left(\sum_0 x^2 \right) - \frac{1}{m} \sum_g \left(\sum_0 x \right)^2 \\ SC_D &= \sum_g \left(\sum_0 x^2 - \frac{\left(\sum_0 x \right)^2}{m} \right) \quad (3.52) \end{aligned}$$

Con cada "SC" está asociado un número de grados de libertad (gl). Las "x" tienen N-1 grados de libertad ya que hay N de ellas y su suma debe ser cero. Luego N-1 gl están asociados con SC_T .

En cada grupo hay m valores de d y su suma (en cada grupo) es cero, luego en cada grupo las "d" tienen m-1 gl. Si tenemos p grupos, el número de gl asociado con las d, y por tan-

to con SC_D es $p(m-1)$. Finalmente hay p valores de g que también suman cero. Por tanto los gl . asociados con SC_E son $p-1$.

La razón de cualquier SC y su respectivo gl (llamado "cuadrado medio" ó CM) tiene todo el aspecto de una variancia y es un estimado de un parámetro de la población. CM_T es un estimado de la variancia de la población. Para ver estimados de que parámetros son los otros CM se asume que todas las observaciones en la población pueden representarse por el modelo

$$x = G + E + \bar{x}$$

ó

$$x - \bar{x} = G + E$$

(3.53)

G representa las fluctuaciones sistemáticas únicamente y tiene variancia σ_G^2 mientras que E representa las variaciones aleatorias con variancia constante σ^2 .

CM_D está asociado con SC_D que depende de las fluctuaciones aleatorias E_D y es por tanto estimado de σ^2 . Por otra parte $CM_E = SC_E / p-1$ depende de las fluctuaciones aleatorias " E " y de las variaciones sistemáticas G , resulta ser un estimado de $\sigma^2 + m\sigma_G^2$.

Una vez calculado CM_D y CM_E puede llegarse a dos situaciones:

a) $CM_D > CM_E$ en cuyo caso las fluctuaciones sistemáticas no representan nada y po

demostramos decir que las variaciones de las variancias son debidas al azar.

b) $CM_E \searrow CM_D$ lo cual significa que σ_G^2 difiere significativamente de cero. Esto implica una variación sistemática que es probada por medio de la estadística

$$F = \frac{CM_E}{CM_D} \quad (3.54)$$

que tiene una distribución de Fisher (III. 8) con $v_1 = p-1$ y $v_2 = p(m-1)$ grados de libertad. La prueba se realiza como cualquiera de las pruebas de hipótesis vistas antes hacia una cola de la curva y a un nivel α dado (en la tabla V del apéndice están los valores de F en función de v_1 y v_2 para valores de α de 1% y 5%.)

Regresando al ejemplo presentamos en el cuadro 10 los cálculos de las SC, gl, etc.

CUADRO 10 COMPUTOS DEL ANALISIS DE VARIANCIA

Origen	SC	g	CM	Estimado de
Entre grupos (E)	2.30	2	1.15	$\sigma^2 + m\sigma_G^2$
Dentro de grupos (D)	2.98	12	0.25	σ^2
TOTAL	5.28	14		

Puede verse que $CM_E \searrow CM_D$ lo cual es indicativo de una variación sistemática apreciable.

La significación de la variación depende del nivel α escogido, por ejemplo la estadística CM_E/CM_D es 4.60 y para un nivel $\alpha=1\%$ F vale 6.93 ($v_1 = 2$ y $v_2 = 12$), luego diremos que a un nivel de 1% la variación no es significativa; sin embargo a un nivel de 5% la variación si es significativa. El concepto es mas bien relativo como puede verse y una conclusión definitiva no puede darse hasta no obtener mas información.

En el cuadro 11 están resumidas las distribuciones que se utilizan para las pruebas dependiendo del problema y de los parámetros de que se disponga

CUADRO 11

Problema	Parámetros Disponibles			
	μ y σ	solo μ	solo σ	Prueba general
Pertenece la muestra al universo	Normal	"t"	χ^2	
Puede representar se el universo por una muestra?	Normal	"t"	χ^2	
Pertenecen dos muestras al mismo universo?		"t"	F	
Pertenecen una serie de muestras al mismo universo?				F

III.12 Los pronósticos de tiempo

La utilidad verdaderamente aceptable que tiene el conocimiento de la información meteorológica pasada es que nos permite inferir, a partir de ella, sobre las condiciones de tiempo futuras. La importancia práctica que los pronósticos de tiempo tienen jamás debe ser subestimada, pero debe recordarse que la confianza que se tiene en ellos nunca es absoluta. Por esto toda persona ó institución que haga uso de un pronóstico de tiempo debe sopesarlo de acuerdo a sus intereses. Por ejemplo un constructor de obras de ingeniería va a hacer una fundición de concreto y solicita un pronóstico de lluvia para esa fecha. Se le dice que va a llover con un 85% de confianza. Como existe cierta probabilidad de que no llueva el constructor debe pensar si puede correr el riesgo y fundir ó si las pérdidas que le ocasionaría la lluvia son tales que es preferible esperar una mejoría del tiempo.

III.12.1 Predictores y Predictandos

Los métodos estadísticos para pronósticos de tiempo están basados en la relación del elemento que se desea pronosticar (llamado predictando) con las variables independientes que se utilizarán para pronosticar (predictores). La parte más difícil (y a la vez más importante) es la selección de los predictores. Debe hacerse (la selección) en base a un razonamiento físico más bien que a un carácter intuitivo por cuanto para que un fenómeno ocurra deben cumplirse ciertos requisitos los otros fenómenos;

si uno ó varios de los eventos no se lleva a cabo la variable que se desea pronosticar no ocurrirá u ocurrirá parcialmente. No hay una guía estricta sobre cuales deben ser los predictores para tal ó cual predictando, por ello, para cada caso y posición geográfica, es necesario graficar esquemas de las variables previas a la condición crítica del predictando. La cantidad de predictores es también importante; aunque teóricamente mientras mas predictores se tengan mejor será el pronóstico, esto presupondría el análisis de una gran cantidad de datos estadísticos. Además, al aumentar los predictores aumentamos simultáneamente en cada relación predictor-predictando las variaciones aleatorias que son, por su naturaleza, impredecibles. En otras palabras obtenemos relaciones "demasiado buenas" pero no aumentamos en mucho el éxito del pronóstico.

En la mayor parte de los casos se pronostica un elemento en un lugar en base a predictores escogidos en ese mismo lugar; sin embargo en algunas ocasiones es conveniente auxiliar se de estaciones que están en la ruta natural que sigue el viento en su movimiento de y hacia la estación de pronóstico (de hecho esto es lo que se hace, en mayor escala para predecir los movimientos de las grandes masas de aire).

III.12.2 Métodos de Regresión Gráfica

Los métodos de regresión gráfica consisten esencialmente en una simplificación sucesiva de varios predictores por otros mas simples y en menor número. Supongamos que tenemos n observaciones de tres variables x , y , z , en un

sistema de coordenadas como la de la figura siguiente; podemos graficar lo que se llama un diagrama de dispersión de x e y que no es mas que un asentamiento de los puntos (x_1, y_1) , el (x_2, y_2) , etc. en el sistema de ejes; además, se ha colocado al lado de cada punto el correspondiente valor de z . A ojo pueden dibujarse curvas que unan puntos de igual valor de z llamadas curvas equiescalares ó "isopletas". No deben seguir estas líneas los valores iguales de la variable en forma rigurosa sino mas bien deben tratar de suavizarse. Con esto se logra no solo facilitar su lectura sino también eliminar, hasta cierto punto, las variaciones aleatorias.

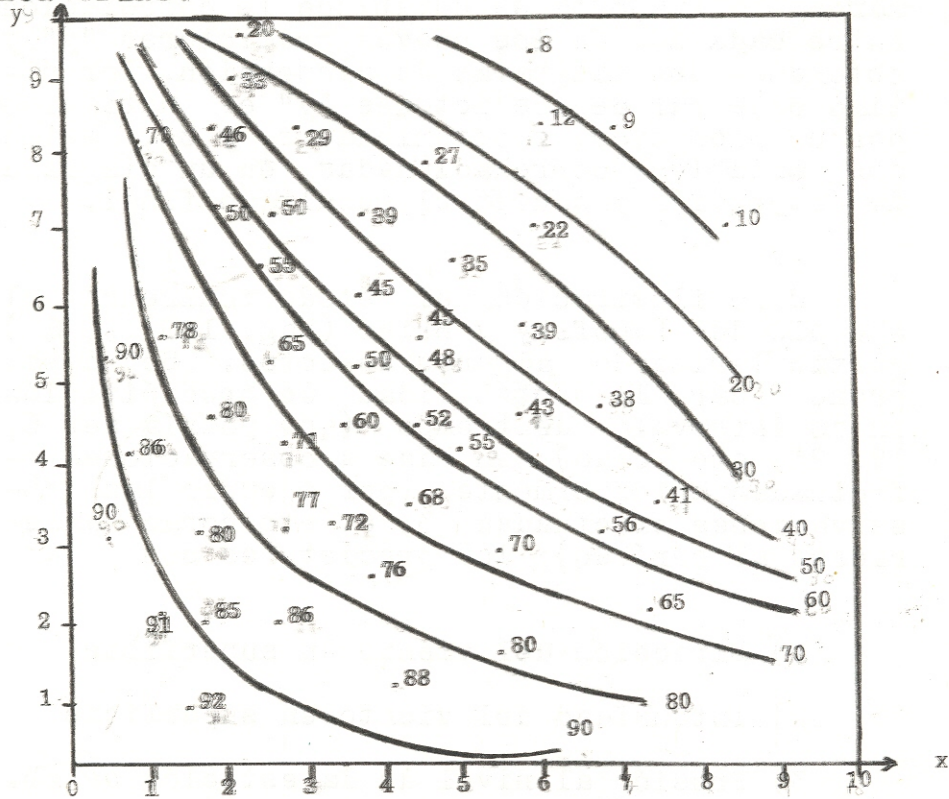


FIGURA 23

No es necesario explicar la importancia práctica que tienen los diagramas de dispersión en la determinación objetiva de la correlación entre dos variables. Asumamos volviendo a los pronósticos, que tenemos 8 predictores x_1, x_2, \dots, x_8 y un predictando P, y que conocemos la correlación entre cada par de predictores. Escogiendo los pares de predictores con la correlación mas definida podríamos graficar cuatro diagramas como el de la figura anterior, cada uno correspondiendo a un par de predictores y al predictando. Cada par de predictores puede ser sustituido por otro predictor "y" que es función de los predictores que lo originaron. Nuevamente se establece la correlación entre cada dos de los nuevos predictores "y" y obtenemos dos diagramas de dispersión. Por último cada par de predictores "y" es sustituido por un predictor z obteniendo z_1 y z_2 ; estas dos variables correlacionadas en un diagrama de dispersión producen el resultado final.

Como ilustración del método tomaremos el ejemplo de Panofsky y Brier (pag. 180) (6) particularizando algunos aspectos. Se desea pronosticar la probabilidad de precipitación en un intervalo de tiempo (que podría ser 6, 12, 24, etc horas) en base a observaciones efectuadas anteriormente (por ejemplo las observaciones efectuadas 3, 6, etc horas anteriores al período); los predictores son

- 1) Dirección del viento en superficie
- 2) Intensidad del viento en superficie
- 3) Presión al nivel de la estación en mb.

- 4) Tendencia de la presión en mb/3 horas
- 5) Componente N-S del viento a los 700 mb
- 6) Curvatura de los 700 mb.
- 7) Temperatura ambiente en superficie
- 8) Temperatura del punto de rocío en superficie

Los predictores de los cuales se espera una correlación lógica son x_1 con x_2 (dirección del viento con su intensidad), x_3 con x_4 , x_5 con x_6 y x_7 con x_8 . Una vez decididos los intervalos de pronóstico (los cuales están en función del uso que se le vaya a dar a la predicción) supondremos que tenemos una gran cantidad de observaciones a_1, \dots, a_n de los predictores y del predictando; el primer paso es entonces tabular para cada par de predictores la frecuencia relativa ó probabilidad empírica del predictando, debiendo anotar, en cuales observaciones se cumple la condición de cada par de predictores. Con estos cuadros (en nuestro ejemplo serían cuatro) pueden confeccionarse cuatro diagramas de dispersión con sus isopletas (en este caso las isopletas son curvas que unen puntos de igual probabilidad empírica) como en las figuras 24, 25, 26 y 27.

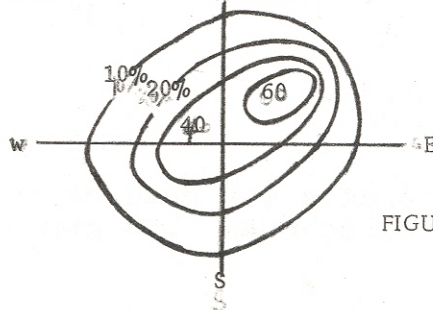
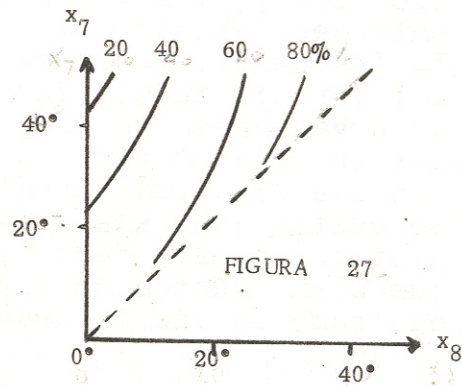
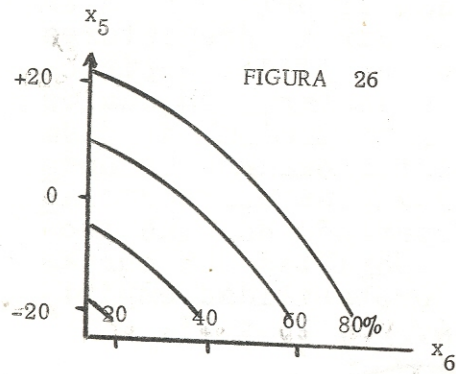
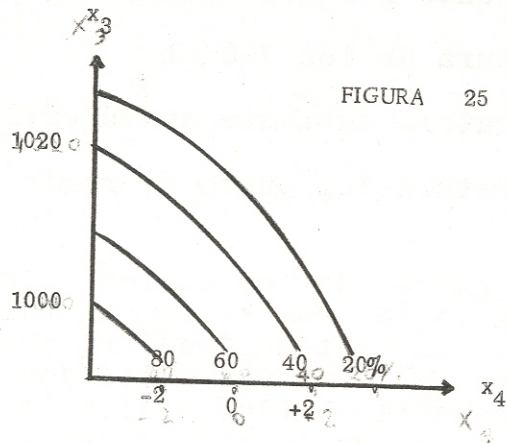
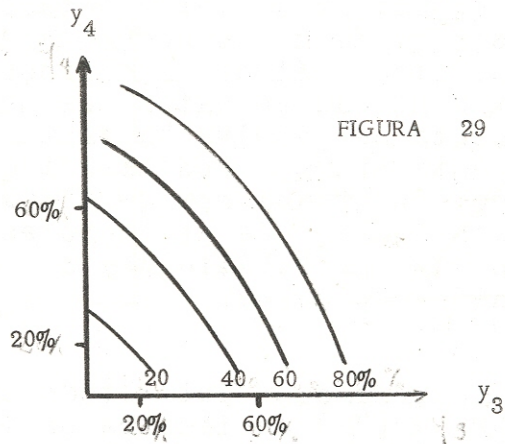
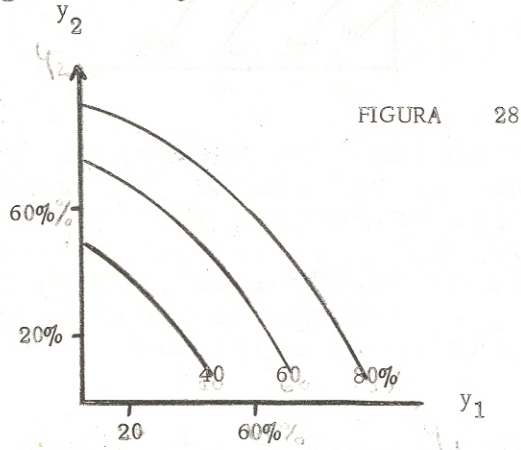


FIGURA 24

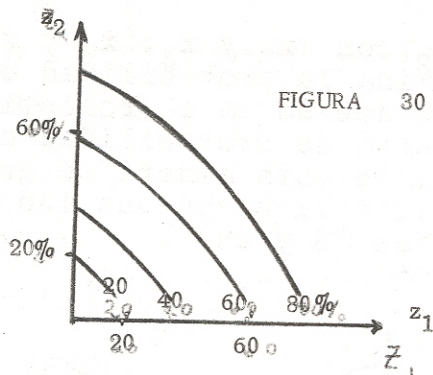


Volviendo a los cuadros de los predictores, se seleccionan dos grupos de pares (por

ejemplo x_1 y x_2 con x_3 y x_4 , x_5 y x_6 con x_7 y x_8) y se determina la probabilidad empírica de que el elemento que se va a pronosticar se realice para valores de probabilidad de cada par de predictores. De esta manera se grafican los diagramas de dispersión con sus isopletas respectivas (figuras 28 y 29)



Finalmente se traza el diagrama de dispersión resultante de agrupar los diagramas anteriores dando lugar al gráfico final (figura 30).

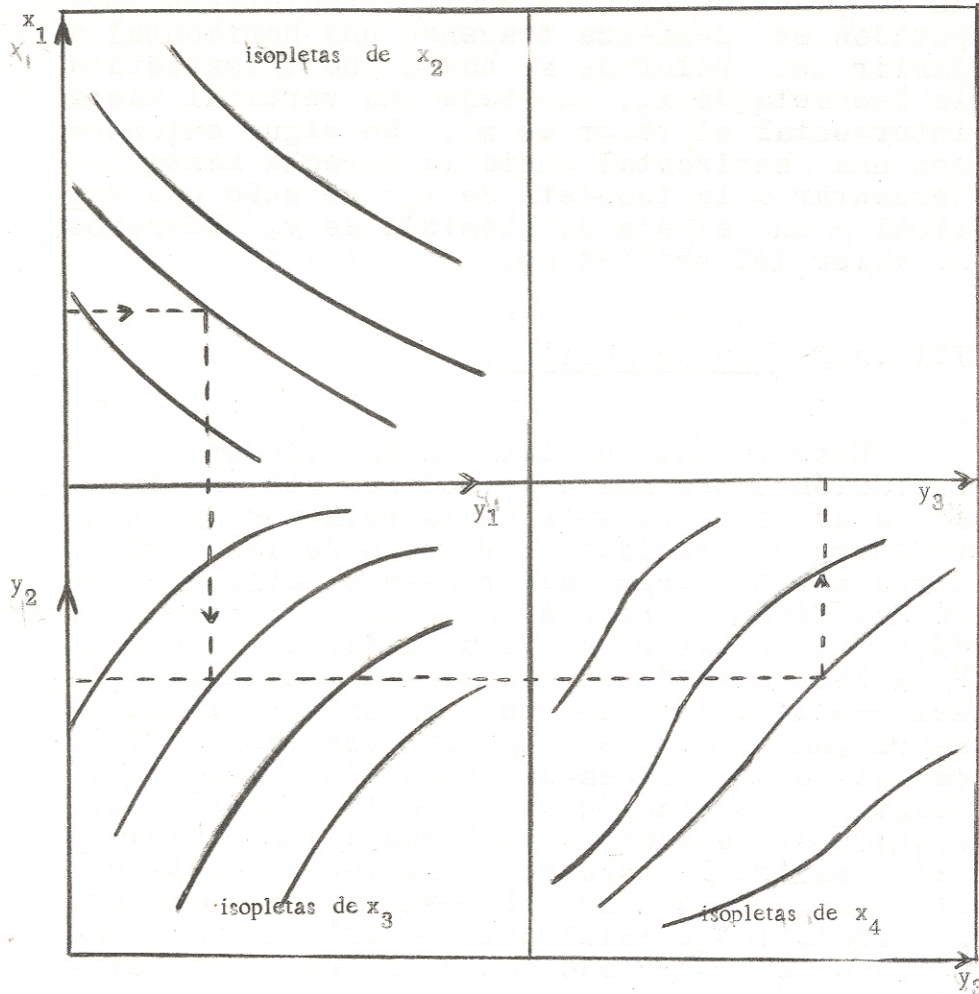


Al utilizar el método en la práctica se entra con los valores de los predictores x_1, \dots, x_8 a los primeros cuatro gráficos obteniendo los valores y_1, \dots, y_4 ; con estos pasamos a los gráficos semifinales y encontramos valores de z_1 y z_2 . Se llega al pronóstico final en el último gráfico de isopletas.

Como ventajas de este método pueden señalarse su facilidad de interpretación, en todas sus fases; además no se necesita de computadora como en otros métodos numéricos. Sin embargo si la cantidad de datos es muy grande, el procedimiento se vuelve bastante tedioso. El autor de este trabajo recomienda que para estas condiciones con observaciones continuas, se de lugar, en los cuadros de predictores, a modificaciones en las probabilidades a medida que vayan obteniéndose datos.

Una modificación del método propuesta por W. C. Spreen (7) se ilustra en la figura siguiente para el caso de cuatro predictores x_1, x_2, x_3 y x_4 y un predictando P .

FIGURA 31



Después de obtener los cuadros de los predictores, se configura el diagrama de isople²tas (x_1, y_1, x_2) (notar que las isople²tas son del predictor \bar{x}_2); abajo de este diagrama y usando el mismo eje de abscisas, trazamos el diagrama (y_2, y_1, x_3). A la derecha del diagrama anterior y utilizando el mismo eje de ordenadas, se confecciona el diagrama (y_2, y_3, x_4).

En la práctica con los cuatro predictores computados se comienza trazando una horizontal a partir del valor de x_1 hasta que intersecta a la isopleta de x_2 , se baja una vertical hasta intersectar el valor de x_3 , se sigue entonces con una horizontal hacia la derecha hasta intersectar a la isopleta de x_4 ; se sube una vertical y en el eje de abscisas de y_3 obtenemos el valor del pronóstico.

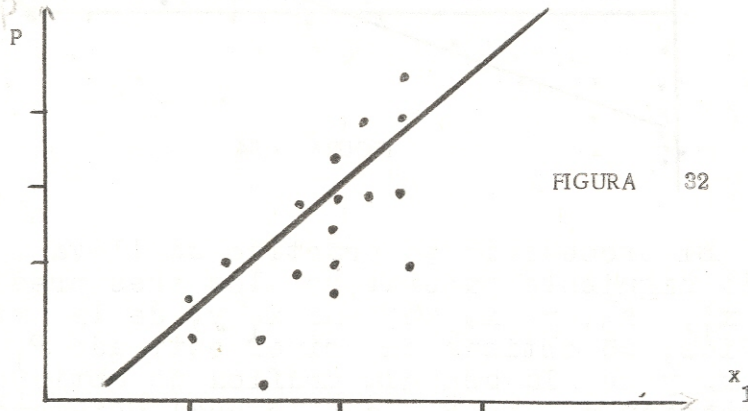
III.12.3 Método Residual

Este método consiste en establecer una correlación entre las desviaciones del pronóstico basado inicialmente en la relación existente entre el predictando P y uno de los predictores x_1 . En otras palabras se establece primero una línea de regresión de x_1 y P ; esto nos dá, para un valor de x_1 , un estimado inicial de P . A continuación se grafica la desviación d_1 del predictando respecto a la primera línea de regresión, contra el segundo predictor y se saca una segunda línea de regresión. Luego se asienta la desviación de d_2 de la primera desviación d_1 respecto a la segunda línea de regresión, contra la tercera variable. Se continúa el procedimiento para el resto de predictores; el pronóstico inicial del predictando va siendo entonces corregido con cada predictor sucesivo.

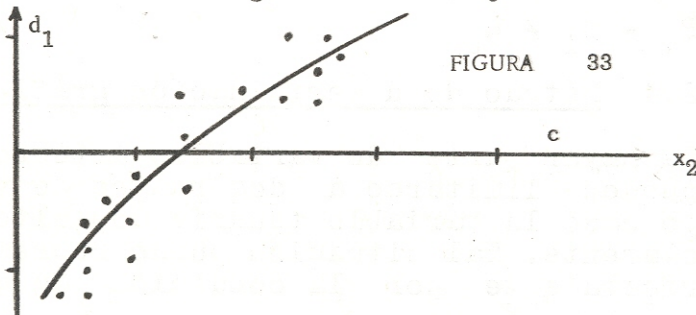
Se ilustrará el método con un ejemplo de Panofsky y Brier (6) (página 186) en el que se trata de pronosticar la temperatura mínima en base a tres predictores

- 1) Temperatura ambiente (x_1)
- 2) Velocidad del viento (x_2)
- 3) Cantidad de nubes bajas (x_3)

El primer paso es graficar un diagrama de dispersión de x_1 contra el predictando P y luego determinar una primera línea de regresión que podría quedar como la figura siguiente



Para cada valor de x_1 existe un valor simultáneo de x_2 y existe también una desviación d_1 del valor del predictando dado por la línea de regresión, respecto al verdadero valor del predictando. El conjunto de estas desviaciones es graficado contra x_2 lo que da lugar a una nueva línea de regresión dibujada a continuación



Nuevamente, asociado a cada valor de x_2 está un valor de x_3 y una desviación d_2 de la desviación d_1 . Se grafica entonces x_3 contra d_2 determinando una tercera línea de regresión como en la figura siguiente

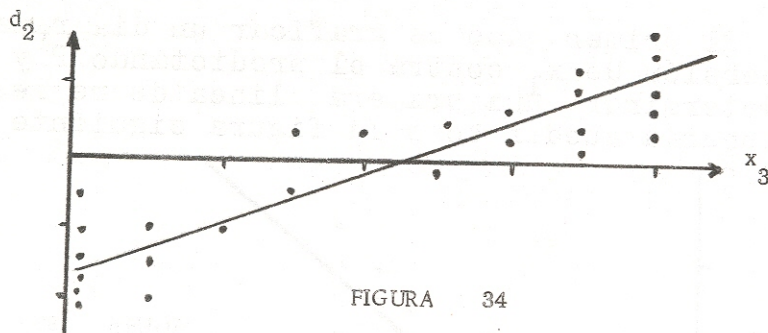


FIGURA 34

El procedimiento práctico se lleva a cabo de la siguiente manera: con los tres predictores x_1 , x_2 , y x_3 se toma x_1 y, de la primera gráfica, se obtiene un primer estimado P_1 de P ; con x_2 y de la segunda gráfica se obtiene una corrección d_1 de P_1 con lo cual obtenemos un segundo estimado P_2 tal que

$$P_2 = P_1 \neq d_1$$

con x_3 se obtiene una corrección para d_1 y consiguientemente el pronóstico deseado.

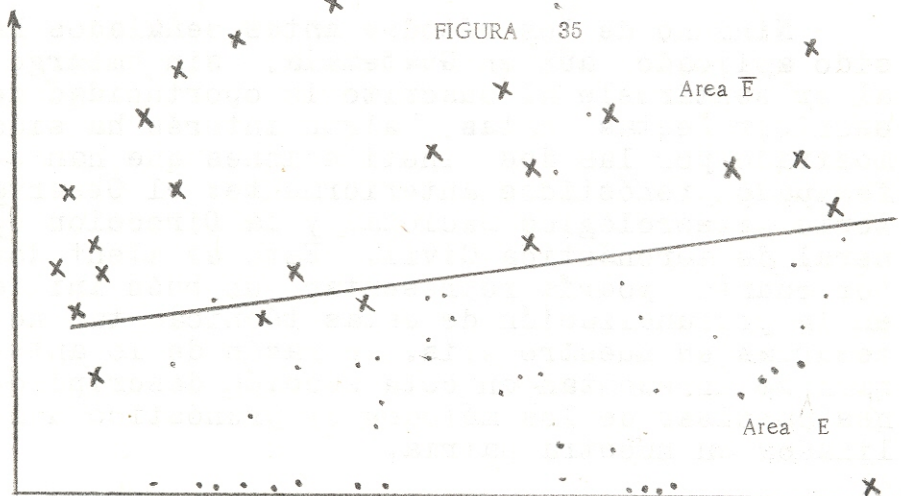
$$P_3 = P_1 \neq d_1 \neq d_2$$

III.12.4 Método de discriminación gráfica

La mayor parte de variables meteorológicas pueden limitarse a dos rangos de valores en cuyo caso la variable tendría dos alternativas solamente. Tal situación puede representarse perfectamente por la condición de que se

realice ó no el evento que nos interesa (ver sección II.1.1) y puede a su vez pronosticarse por el método de discriminación gráfica.

Básicamente el método consiste en lo siguiente: se tienen x_1, \dots, x_n predictores y se quiere predecir la ocurrencia ó no ocurrencia de un fenómeno E. Se agrupan los predictores por pares asignándole a cada par una cierta certidumbre en la verificación del pronóstico. Una vez agrupados se confeccionan diagramas de dispersión como el de la figura, en donde los puntos significan ocurrencia del evento (E) y la equis no ocurrencia del fenómeno (\bar{E})



Se traza entonces una línea tal que una parte del área corresponda a la mayor cantidad de ocurrencias del evento y la otra parte a las no ocurrencias del evento. Hecho esto con todos los predictores, el modo de efectuar el pronóstico es bien fácil: se entra a cada una de las gráficas con el correspondiente par de

predictores adicionando ó sustrayendo las respectivas certidumbres según sea el caso que de seamos pronosticar la ocurrencia ó la no ocurrencia de E. Naturalmente que la parte difícil de esto está en la elección de cada par de predictores y en el grado de certidumbre que se les proporcione. Como ya se dijo antes, no existe una regla fija en cuanto a esto; no obstante, deben escogerse siempre que sea posible obedeciendo a leyes naturales y razonamientos lógicos.

III.12.5 Métodos utilizados en Guatemala

Ninguno de los métodos antes señalados ha sido aplicado aún en Guatemala, sin embargo, al presentársele al suscrito la oportunidad de escribir estas notas, algún interés ha sido mostrado por las dos instituciones que han efectuado pronósticos anteriormente: el Observatorio Meteorológico Nacional y la Dirección General de Aeronáutica Civil. Esto es alentador por cuanto podría representar un buen inicio en la profundización de estas técnicas tan necesarias en nuestro país. En razón de lo anterior se presentan en esta sección descripciones concisas de los métodos de pronóstico utilizados en nuestra patria.

En el Observatorio Nacional se viene haciendo uso desde hace algún tiempo, de un método para pronósticos locales propuesto por M. de Gachons. Se basa el procedimiento en cierta correlación existente entre la tendencia general de la presión atmosférica y la temperatura ambiente en el período anterior al pronóstico.

Para poder determinar la tendencia de ambas curvas es necesario contar con un microbarógrafo y un termógrafo sensible, ambos colocados en un abrigo meteorológico estándar con objeto de uniformizar las medidas.

Una forma objetiva para determinar la tendencia es la siguiente: recordando que la marea barométrica a lo largo de un día tiene dos máximos (a las 10 de la mañana y a las 10 de la noche) y dos mínimos (a las 4 de la mañana y a las 4 de la tarde), se promedian los dos primeros valores de mínima y máxima y luego los dos segundos valores; el valor de presión sobre la curva correspondiente a cada promedio es trazado en la parte superior de un gráfico como el de la figura, en donde las abscisas representan el tiempo (debemos ser cuidadosos en graficar cada promedio en la hora verdadera en que ocurrió). En la parte inferior del gráfico se asientan, a una escala conveniente, las temperaturas que ocurrieron a la hora en que, simultáneamente, los promedios de presión estaban ocurriendo. Los puntos graficados en cada gráfica se unen por segmentos de recta. Dependiendo del período de ante-pronóstico que se tome se obtendrán uno ó varios segmentos de recta; por ejemplo si el período es de 24 horas se obtendrán (como en la figura) dos segmentos de recta, uno para presión y el otro de temperatura

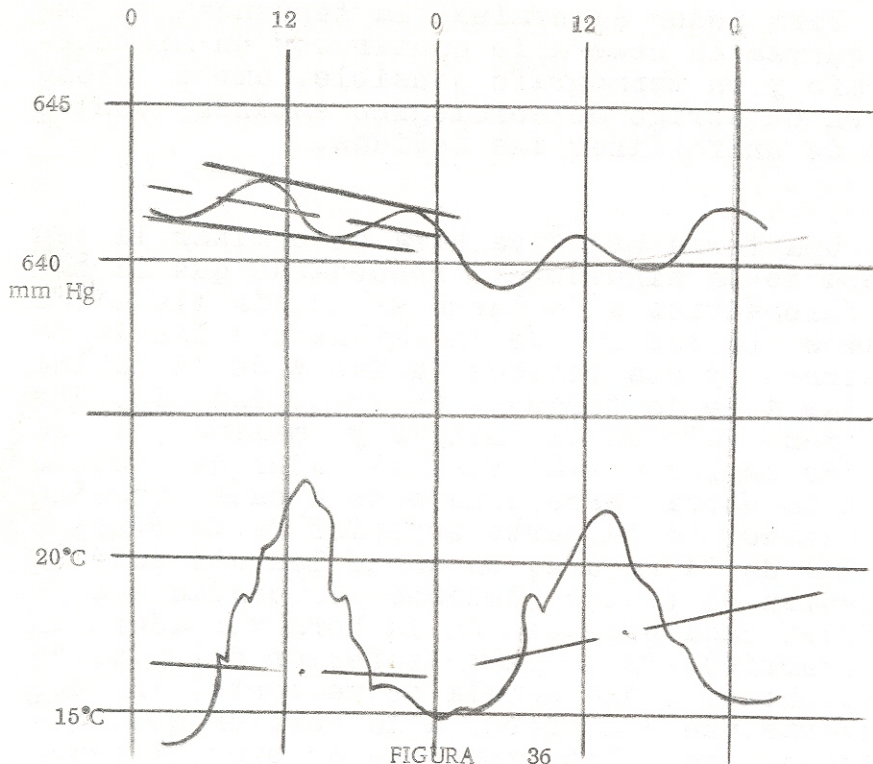


FIGURA 36

En cualquier caso, el pronóstico estará basado en consideraciones de convergencia ó divergencia relativa de las curvas, de acuerdo con las siguientes normas:

- A) Si las dos curvas se acercan "lentamente" hay indicios de mal tiempo
- B) Si las dos curvas se alejan "lentamente" hay indicios de mejora
- C) Si bruscamente se acercan hay indicios de condiciones borrascosas

- D) Si bruscamente se alejan existen indicios de mejora del tiempo, poco estable
- E) Si se acercan con oscilaciones hay indicios de un largo período de mal tiempo
- F) Si se alejan con oscilaciones poco uniformes existe probabilidad de buen tiempo mas ó menos estable
- G) Si las dos líneas van paralelas hay posibilidad de poco ó ningún cambio del tiempo

Puede notarse que el método no es muy explícito por cuanto no nos dice con que intensidad debe considerarse que la condición "lentamente" ó "bruscamente" se está verificando. Por otro lado la condición de paralelismo es bastante ambigua; las rectas pueden ser paralelas hacia arriba, horizontalmente ó hacia abajo y aunque las dos primeras condiciones si podrían ser indicativas de "poco ó ningún cambio de tiempo", es realmente poco probable que un paralelismo hacia abajo haga ver una condición no variable del tiempo. Sin embargo, a decir de los pronosticadores del Observatorio Nacional, los resultados mas ó menos satisfactorios obtenidos con este sistema de pronóstico, radican en el conocimiento preciso de las características de la región que se considera. En tal aspecto, el autor de este trabajo está de acuerdo y cree sinceramente que los resultados de estos pronósticos se deben a la experiencia y conocimiento que del valle de Guatemala tiene el pronosticador del Observatorio Nacional así como a la correcta interpretación de los aspectos que en cada caso presenta dicho método.

Así por ejemplo, se puntualiza que en ciertas épocas existe una determinada tendencia a acortar ó a alejar el momento de verificación de los fenómenos que se predicen, y que aparentemente estos aspectos reclaman un proceso cíclico que se ha venido comprobando y estudiando detenidamente.

En Aeronáutica Civil se vienen elaborando desde 1957 pronósticos cuya función primordial ha sido ayudar a mantener la seguridad en vuelo de aeronaves. Son elaborados en base a condiciones dinámicas de las masas de aire sobre grandes regiones y cumplen su cometido si son aplicados para predecir condiciones de tiempo, digamos promedio, sobre la ruta a cubrir por una aeronave; sin embargo pierden su eficacia cuando tratan de pronosticar condiciones locales (por ejemplo las del valle de Guatemala) ya que estas, aún estando influenciadas por movimientos regionales, lo están en mayor grado por características especiales del valle que crean condiciones de tiempo muy particulares y por tanto sumamente difíciles de pronosticar.

III.13 Resumen

En esta sección se tratarán de describir los elementos climatológicos que mas influencia tienen en las actividades humanas indicando someramente la forma en que deben ser tratados desde el ángulo estadístico (para los elementos ya estudiados en secciones anteriores simplemente se hará referencia a ellas).

III.13.1 La temperatura del aire

Probablemente la temperatura del aire es el factor mas importante -y también el mas común- en Climatología. El rango y nivel de la temperatura nos indica la posibilidad y el tipo de vida que puede existir en un determinado lugar. Las expresiones de clima "caluroso" ó "frío" son de por sí descripciones amplias del elemento y es indudable que la mayor parte de los procesos biológicos están íntimamente relacionados con él.

Entre la superficie terrestre sea esta líquida ó sólida, y el aire de la atmósfera se lleva a cabo un intercambio de energía calorífica de una magnitud tal, que hace a la temperatura de la superficie directamente responsable de la temperatura del aire. El calor es en parte transferido al ambiente por conducción a la capa adyacente a la superficie. Esta capa libera entonces parte de su energía por conducción y convección cediéndola a capas mas altas de la atmósfera.

Para propósitos estadísticos se utiliza el promedio aritmético de la temperatura ambiente que puede ser diario, mensual ó anual. El primero puede ser computado, en estaciones equipadas con termógrafos, de las lecturas hechas durante las 24 horas del día. Para estaciones que no tienen termógrafo, el promedio diario se computa a partir de observaciones hechas a termómetros. Si son efectuadas tres observaciones diarias a las 7, 14 y 21 horas local y las lecturas son respectivamente a, b y

c, Landsberg (14) dá como una buena aproximación de la media la fórmula

$$T_m = (a + b + 2c)/4 \quad (3.55)$$

Estaciones donde solo son hechas observaciones de temperaturas máxima y mínima (M y m respectivamente), obtienen la temperatura media por

$$T_m = (M + m)/2 \quad (3.56)$$

debiendo ser afectado el valor dado por la última fórmula por un factor que puede ser obtenido, para una estación particular, haciendo un registro simultaneo con termógrafo durante un año. El promedio mensual se computa dividiendo la suma de las medias diarias entre el número de días en el mes. Finalmente promedios anuales son obtenidos dividiendo la suma de los promedios mensuales entre 12. Cualquiera de estos promedios arreglado en forma de serie climatológica tiende a estar distribuido normalmente. Un análisis completo de este elemento, para un mes dado, se ha efectuado en las secciones III.11.2 y III.11.3 por lo que a ellas referimos al lector.

III.13.2 La precipitación

La precipitación es precedida en importancia únicamente por la temperatura. Las formas más variadas de vida están íntimamente vinculadas a las fuentes de agua, y donde no hay agua es sabido que la vida no puede existir. La disponibilidad de agua en una cuenca ó región es regulada fundamentalmente, por la cantidad que

pueda ser obtenida de la atmósfera en forma de precipitación.

Se define la precipitación como toda el agua que cae de la atmósfera a la superficie terrestre. Sus formas más importantes son la lluvia (que en su forma fina es llamada llovizna), la nieve y el granizo siendo las tres medibles por medio de pluviómetros.

Otras formas de precipitación menos accesibles a ser medidas tan fácilmente son el rocío y la escarcha y son debidas a condensación ó cristalización directa sobre superficies. El rocío aún cuando no tiene importancia en regiones con suficiente lluvia, gana una posición elevada como protector del reino vegetal en climas donde la lluvia es escasa. La escarcha por otra parte, para poder producirse sobre una superficie debe estar (la superficie) enfriada a bajo del nivel de congelación lo cual causa la sublimación en vez de la condensación que es el caso del rocío.

Volviendo a la precipitación pluvial, la cantidad caída en un año es a menudo tomada como base para definir el clima. Por ejemplo R de Ward (15) propone para definir el clima los siguientes intervalos: una precipitación anual de menos de 5 pulgadas representa un clima desértico; entre 5 y 10 pulgadas lo llama clima árido; de 10 a 20 pulgadas es un clima semiárido; de 20 a 40 pulgadas es un clima moderadamente húmedo; de 40 a 80 pulgadas es clima húmedo y con precipitación sobre las 80 pulgadas es un clima con lluvia excesiva. En climas ári

dos la Agricultura necesita de la irrigación para subsistir, en climas semi-áridos la irrigación es deseable pero en muchos casos los cultivos tienen bastante posibilidad de éxito dependiendo de un pequeño margen de lluvia. Por último en climas moderadamente húmedos y húmedos los recursos de agua son suficientes y abundantes y usualmente no es necesaria la irrigación.

En la sección III.6 se ha analizado el tipo de distribución que mas se acomoda a la precipitación; por ello no se insistirá aquí sobre este asunto, antes bien remitimos al lector a la sección antes mencionada.

III.13.3 El Viento

El viento es un elemento al que se le ha dado importancia desde tiempos inmemoriales y es el mejor indicativo del estado siempre variable de la atmósfera. Para el marino de tiempos pasados el éxito ó fracaso de cualquier empresa que emprendiera dependía casi exclusivamente del viento y aún actualmente todo programa de navegación marítima toma en cuenta las condiciones del viento. Tanto ó mas importante es para las operaciones aeronáuticas donde siendo la carga de combustible un renglón crítico, las corrientes de viento a niveles superiores son un factor vital.

El viento es el resultado de diferencias de presión; siempre que tal diferencia exista en la atmósfera se produce una corriente de ai

re que tiende a balancearla una vez que se ha superado la fuerza de fricción. Sin embargo la diferencia de presiones por si misma no es un hecho aislado sino una consecuencia de diferencias de temperatura que a su vez son el resultado del diferente grado de radiación que recibe la Tierra. La presión y sus diferencias son de gran importancia en Meteorología pero tienen un interés secundario en Climatología. Por el contrario el viento nos indica los cambios de presión y tiene una influencia en el clima muy parecida a la de la temperatura y precipitación.

Variaciones en la dirección del viento dan normalmente un mejor indicativo de los cambios de tiempo que variaciones en su intensidad; por esto algunas estadísticas de viento dan unicamente la dirección predominante mensual ó anual; sin embargo, siendo el viento un vector es necesario que sean incluidos su dirección y su magnitud para estar definido. Climatológicamente la intensidad del viento es tan importante como su dirección; debe tomarse en cuenta que los efectos que produce sobre las plantas, estructuras etc. son causados fundamentalmente por su intensidad. Parte de la comodidad humana puede decirse que depende, hasta cierto punto, de la velocidad del viento.

La variación diaria de la velocidad del viento se representa como una onda simple alcanzándose su valor máximo cerca de la superficie terrestre entre las 13 y 15 horas. Este intervalo es bastante uniforme no así el de los valores mínimos que es mas variable; sin embargo, usualmente se sucede en las primeras horas

de la madrugada. La variación diaria de la dirección de viento depende del lugar en que esté situada la estación y de los accidentesográficos que tenga a su alrededor; portezuelos actúan en muchos casos como túneles de viento prácticamente guiando al viento en una dirección determinada. En estaciones costeras y de montaña es usual encontrar predominancias en direcciones diferentes durante el día y durante la noche. Una forma en que podría graficarse la variación diaria de la dirección e intensidad del viento en estaciones que no cuentan con indicador de viento con registro pero si realizan observaciones horarias (tal el caso de las estaciones de Aeronáutica Civil), sería colocando en una pequeña rosa de vientos la dirección, siendo la intensidad representada por la distancia al centro de la rosa. La forma más común de presentar este vector, no obstante, es la que se dió en la sección III.3. Debe comprenderse que desde el punto de vista de la Climatología es mas interesante la presentación dada en esa sección que una relación de las variaciones diarias; además, acortar los intervalos de intensidades y tratar de establecer una distribución que se acomode a las frecuencias de intensidades para cada dirección resultaría inoperante y costoso; por ello hemos de conformarnos con presentar el viento en su forma tradicional.

III.13.4 La Evaporación

La evaporación es el cuarto elemento utilizado mas comunmente para describir el clima y es el fenómeno físico por medio del cual el agua que cae a la superficie terrestre regresa

a la atmósfera. Ríos, lagos, suelos húmedos y principalmente los océanos son las fuentes más importantes para proveer de agua a la atmósfera cuya cantidad depende de tres factores a saber: del grado de saturación del ambiente, de la disponibilidad, en volumen, de agua y de la intensidad del viento. Los métodos usuales para medir la evaporación asumen que se dispone de una cantidad ilimitada de agua lo que es cierto únicamente en el caso de lagos y mares; por consiguiente al hacer medidas de evaporación debe tomarse en cuenta esta suposición. Para estaciones sin evaporímetro han sido desarrolladas fórmulas empíricas que expresan la evaporación por día en función de las condiciones de humedad y magnitud de viento predominantes. A este respecto existe bastante variación entre unas fórmulas y otras debido esencialmente, a que cada investigador ha trabajado en condiciones diferentes y bajo suposiciones también diferentes. Una fórmula conveniente siempre que sea posible determinar la tensión del vapor y la intensidad del viento a dos niveles diferentes (utilizando por ej. un mástil de 2 mts. de altura) es la derivada por Thornwaite y Holzman que es:

$$E = 17.1(e_1 - e_2)(u_2 - u_1) / (t/459.4) \quad (3.57)$$

siendo

- E la evaporación en pulgadas por día
- e la tensión del vapor en pulgadas de mercurio
- u la velocidad del viento en millas por hora
- t la temperatura en grados Fahrenheit

y los subíndices 1 y 2 se refieren a los niveles 1 y 2

Landsberg (14) dá como fórmula práctica para determinar la evaporación la siguiente

$$E = (0.4 + 0.004v)(e - e') \quad (3.58)$$

donde

E es la evaporación en pulgadas por día

v es la velocidad del viento en millas por hora

e es la presión de saturación del vapor de agua a la temperatura ambiente

e' es la presión de saturación del vapor de agua a la temperatura del punto de rocío.

fórmula mas accesible a países como los nuestros en que el tipo de medidas que se hacen son limitadas.

La variación diaria de la evaporación es facilmente perceptible y bastante pronunciada; normalmente entre el 75% y el 90% de la evaporación total diaria ocurre entre las 6 de la mañana y las 6 de la tarde. Esto es debido a que la ocurrencia de radiación a esas horas produce un aumento en la temperatura ambiente y una disminución en la humedad relativa haciendo que la diferencia (e-e') suba durante el día. Durante la noche por el contrario la temperatura baja y la humedad relativa sube de

tal manera que $(e-e')$ baja y con ella la evaporación. La variación estacional de evaporación muestra sus valores máximos durante la época seca y sus mínimos en la época lluviosa causados principalmente por las temperaturas altas en verano y bajas en invierno que permiten altos y bajos valores de $(e-e')$ respectivamente.

No es usual presentar los datos de evaporación en forma de distribución de frecuencias por la dificultad que existe en acomodar una curva a los registros de evaporación. Este problema se hace evidente en el histograma presentado a continuación. Puede observarse lo difícil que resultaría el acomodo de una distribución conocida a los datos a menos que se desee sacrificar gran parte de ellos.

CUADRO 12 EVAPORACION DE MARZO EN mms.

año	evaporación	año	evaporación
1936	224.8	1952	216.8
1937	236.4	1953	154.6
1938	229.1	1954	118.9
1939	250.5	1955	59.1
1940	259.4	1956	153.5
1941	257.2	1957	122.6
1942	264.2	1958	62.2
1943	262.3	1959	190.7
1944	186.1	1960	156.7
1945	188.8	1961	65.3
1946	225.5	1962	181.2
1947	183.3	1963	174.6
1948	199.5	1964	125.8
1949	146.0	1965	71.1
1950	184.6	1966	203.7
1951	153.0	1967	217.0

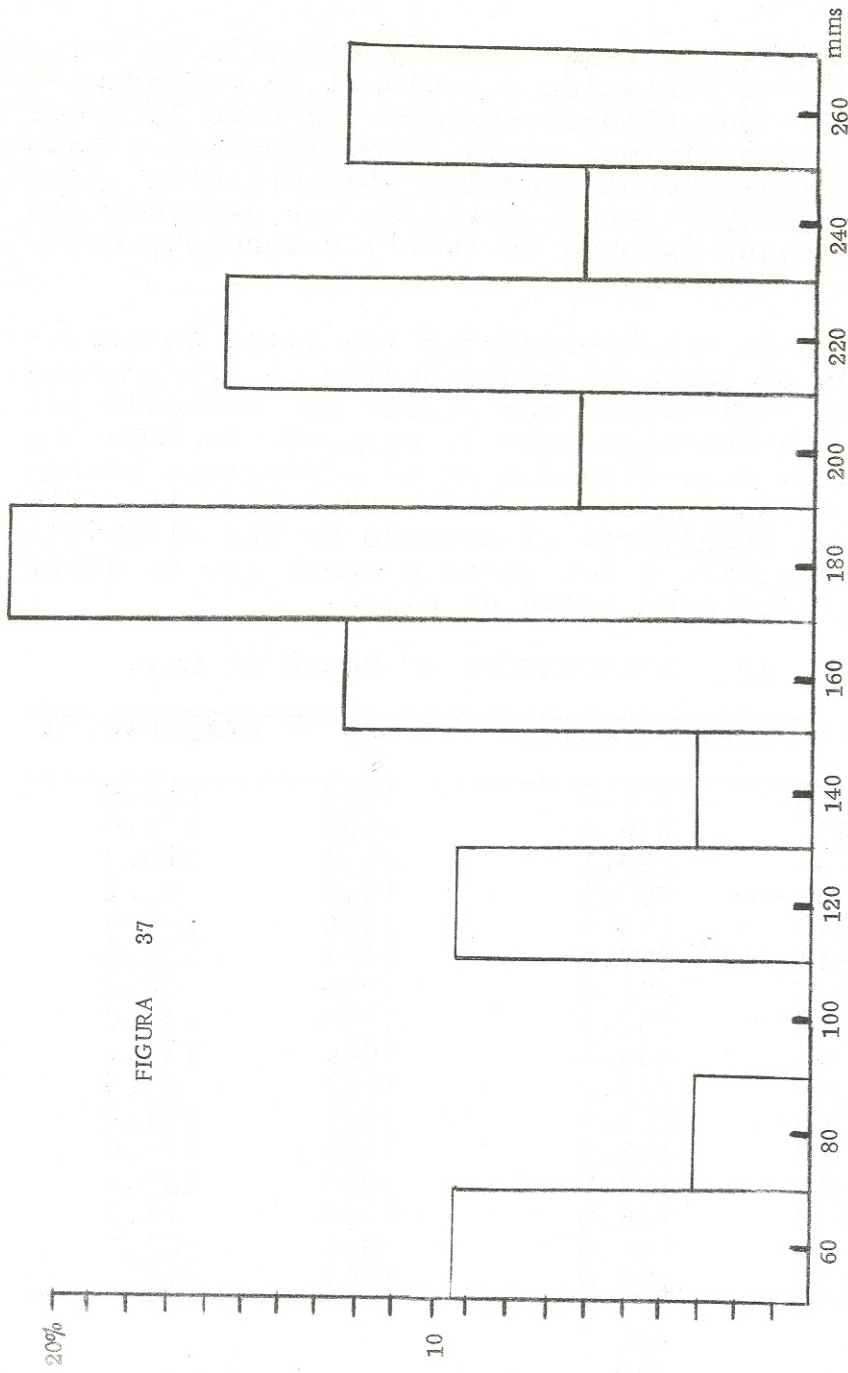


FIGURA 37

III.14 Recomendaciones

1. Darle mas importancia a la Estadística dentro del personal que labora en el campo de la Meteorología, si posible impartiendo cursillos de orientación tendientes a destacar los usos prácticos que pueden obtenerse de ella.
2. Tratar de reunir los registros de observaciones Meteorológicas efectuadas por instituciones diferentes con estaciones diseminadas por todo el país, a efecto de tener congregados, bajo un mismo patrón, todos los datos existentes y futuros.
3. Tratar de normalizar los métodos de observaciones Meteorológicas efectuadas con diferentes fines, lo cual incluiría la normalización del instrumental a usar así como de las unidades de medida de los diferentes elementos, dándole preferencia al sistema métrico.
4. Motivar el interés por adquirir un mayor conocimiento de los aspectos fundamentales del clima en el público mediante la difusión permanente de pronósticos de tiempo, estados actuales de la atmósfera y artículos relacionados con el tema en todos los órganos de divulgación posibles.

5. Aplicar en nuestra patria los métodos de pronósticos explicados en las secciones III.12.2, III.12.3 y III.12.4 los cuales tienen la ventaja de no necesitar de muchos conocimientos matemáticos para ser usados reduciéndose al final de cuentas al establecimiento mecánico de una cierta correlación de fenómenos.

6. Evitar el estudio de fenómenos Meteorológicos de segunda importancia; antes bien, dedicarle mayor atención a elementos que, como los cuatro indicados en el resumen, definen completamente el clima y en general el estado del tiempo.

Arturo Zea Córdón

V° B°

Ing. Enrique Murillo D.
Asesor

V° B°

Ing. Jorge Arias B.
Jefe Departamento

IMPRIMASE

Ing. Amando Vides T.
Decano

TABLA I

Ordenadas de la curva normal estandarizada en "z".

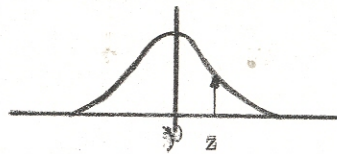


TABLA II

Areas bajo la curva normal estandarizada de cero a "z".

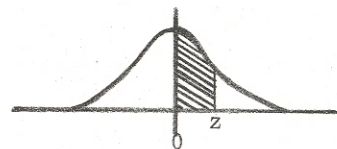


TABLA III

Valores de t_p para la distribución "t" con "v" grados de libertad. El área sombreada es igual a "p".

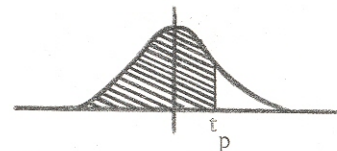


TABLA IV

Areas bajo la curva de Chi-Cuadrado, entre χ^2 e infinito con "v" grados de libertad.

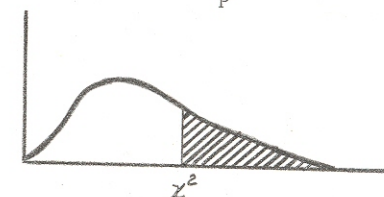


TABLA V

Valores de x_p para la distribución de Fisher con v_1 y v_2 grados de libertad. El valor superior corresponde a $p=0.05$ y el inferior a $p=0.01$

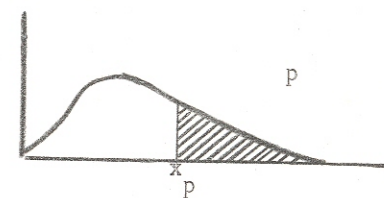


TABLA VI

Dá valores de $e^{-\lambda} (6 e^{-\lambda})$ en la distribución de Poisson.

TABLA VII

Dá valores de la ec. 3.16 (función gamma) para "z" entre 1 y 2

TABLA III : Distribución "t"

V	V										
1	1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	.727	.325	.158
	2	9.92	6.96	4.30	2.92	1.89	1.061	.816	.617	.889	.142
	3	5.84	4.54	3.18	2.35	1.64	.978	.765	.584	.277	.137
	4	4.60	3.75	2.78	2.13	1.53	.941	.741	.569	.271	.134
	5	4.03	3.36	2.57	2.02	1.48	.920	.727	.559	.267	.132
	6	3.71	3.14	2.45	1.94	1.44	.906	.717	.553	.265	.131
	7	3.50	3.00	2.36	1.90	1.42	.896	.711	.549	.263	.130
	8	3.36	2.90	2.31	1.86	1.40	.889	.706	.546	.262	.130
	9	3.25	2.82	2.26	1.83	1.38	.883	.703	.543	.261	.129
	10	3.17	2.76	2.23	1.81	1.37	.879	.700	.542	.260	.129
	11	3.11	2.72	2.20	1.80	1.36	.876	.697	.540	.260	.129
	12	3.06	2.68	2.18	1.78	1.36	.873	.695	.539	.259	.128
	13	3.01	2.65	2.16	1.77	1.35	.870	.694	.538	.259	.128
	14	2.98	2.62	2.14	1.76	1.34	.868	.692	.537	.258	.128
	15	2.95	2.60	2.13	1.75	1.34	.866	.691	.536	.258	.128
	16	2.92	2.58	2.12	1.75	1.34	.865	.690	.535	.258	.128
	17	2.90	2.57	2.11	1.74	1.33	.863	.689	.534	.257	.128
	18	2.88	2.55	2.10	1.73	1.33	.862	.688	.534	.257	.127
	19	2.86	2.54	2.09	1.73	1.33	.861	.688	.533	.257	.127
	20	2.84	2.53	2.09	1.72	1.32	.860	.687	.533	.257	.127
	21	2.83	2.52	2.08	1.72	1.32	.859	.686	.532	.257	.127
	22	2.82	2.51	2.07	1.72	1.32	.858	.686	.532	.256	.127
	23	2.81	2.50	2.07	1.71	1.32	.858	.685	.532	.256	.127
	24	2.80	2.49	2.06	1.71	1.32	.857	.685	.531	.256	.127
	25	2.79	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
	26	2.78	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
	27	2.77	2.47	2.05	1.70	1.31	.855	.684	.531	.256	.127
	28	2.76	2.47	2.05	1.70	1.31	.855	.683	.530	.256	.127
	29	2.76	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
	30	2.75	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
	31	2.70	2.42	2.02	1.68	1.30	.851	.681	.529	.255	.126
	60	2.66	2.39	2.00	1.67	1.30	.848	.679	.527	.254	.126
	120	2.62	2.36	1.98	1.66	1.29	.845	.677	.526	.254	.126
	∞	2.58	2.33	1.96	1.64	1.28	.842	.674	.524	.253	.126

TABLA IV: Chi-Cuadrado $Q(\chi^2/v)$

v	$\chi^2=0.001$	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010
1	0.97477	0.96433	0.95632	0.94957	0.94363	0.93826	0.93332	0.92873	0.92442	0.92034
2	0.99950	0.99900	0.99850	0.99800	0.99750	0.99700	0.99651	0.99601	0.99551	0.99501
3	0.99999	0.99998	0.99996	0.99993	0.99991	0.99988	0.99984	0.99981	0.99977	0.99973
4	0.						0.99999	0.99999	0.99999	0.99999
v	$\chi^2=0.01$	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
1	0.92034	0.88754	0.86249	0.84148	0.82306	0.80650	0.79134	0.77730	0.76418	0.75183
2	0.99501	0.99005	0.98511	0.98020	0.97531	0.97045	0.96661	0.96079	0.95600	0.95123
3	0.99973	0.99925	0.99863	0.99790	0.99707	0.99616	0.99518	0.99412	0.99301	0.99184
4	0.99999	0.99995	0.99989	0.99980	0.99969	0.99956	0.99940	0.99922	0.99902	0.99879
5			0.99999	0.99998	0.99997	0.99995	0.99993	0.99991	0.99987	0.99984
6							0.99999	0.99999	0.99999	0.99998
v	$\chi^2=0.1$	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.75183	0.65472	0.58388	0.52709	0.47950	0.43858	0.40278	0.37109	0.34278	0.31731
2	0.95123	0.90484	0.86071	0.81873	0.77880	0.74082	0.70469	0.67032	0.63763	0.60653
3	0.99184	0.97759	0.96003	0.95024	0.91889	0.89643	0.87320	0.84947	0.82543	0.80125
4	0.99879	0.99532	0.98981	0.98248	0.97350	0.96306	0.95133	0.93845	0.92456	0.90980
5	0.99984	0.99911	0.99764	0.99533	0.99212	0.98800	0.98297	0.97703	0.97022	0.96257
6	0.99998	0.99985	0.99950	0.99885	0.99784	0.99640	0.99449	0.99207	0.98912	0.98561
7		0.99997	0.99990	0.99974	0.99945	0.99899	0.99834	0.99744	0.99628	0.99483
8			0.99998	0.99994	0.99987	0.99973	0.99953	0.99922	0.99880	0.99825
9				0.99999	0.99997	0.99993	0.99987	0.99978	0.99964	0.99944
10					0.99999	0.99998	0.99997	0.99994	0.99989	0.99983
11							0.99999	0.99998	0.99997	0.99995
12								0.99999	0.99999	0.99999
v	$\chi^2=1.1$	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
1	0.29427	0.27332	0.25421	0.23672	0.22067	0.20590	0.19229	0.17971	0.16808	0.15730
2	0.57695	0.54881	0.52205	0.49659	0.47237	0.44933	0.42741	0.40657	0.38674	0.36788
3	0.77707	0.75300	0.72913	0.70553	0.68227	0.65939	0.63693	0.61493	0.59342	0.57241
4	0.89427	0.87810	0.86138	0.84420	0.82664	0.80879	0.79072	0.77248	0.75414	0.73576
5	0.95410	0.94488	0.93493	0.92431	0.91307	0.90125	0.88890	0.87607	0.86280	0.84915
6	0.98154	0.97689	0.97166	0.96586	0.95949	0.95258	0.94512	0.93714	0.92866	0.91970
7	0.99305	0.99093	0.98848	0.98557	0.98231	0.97864	0.97457	0.97008	0.96517	0.95984
8	0.99753	0.99664	0.99555	0.99425	0.99271	0.99092	0.98887	0.98654	0.98393	0.98101
9	0.99917	0.99882	0.99828	0.99732	0.99715	0.99633	0.99537	0.99425	0.99295	0.99147
10	0.99973	0.99961	0.99944	0.99921	0.99894	0.99859	0.99817	0.99766	0.99705	0.99634
11	0.99992	0.99987	0.99981	0.99973	0.99962	0.99948	0.99930	0.99908	0.99882	0.99850
12	0.99998	0.99996	0.99994	0.99991	0.99987	0.99982	0.99975	0.99966	0.99954	0.99941
13	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99991	0.99988	0.99983	0.99977
14			0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99992
15						0.99999	0.99999	0.99999	0.99998	0.99997
16								0.99999	0.99999	0.99999
v	$\chi^2=2.2$	2.4	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0
1	0.13801	0.12134	0.10686	0.09426	0.08327	0.07364	0.06520	0.05778	0.05125	0.04550
2	0.03287	0.30119	0.27253	0.24660	0.22313	0.20190	0.18268	0.16530	0.14957	0.13534
3	0.53195	0.49363	0.45749	0.42350	0.38163	0.36181	0.33397	0.30802	0.28389	0.26146
4	0.69903	0.66263	0.62682	0.59183	0.55783	0.52493	0.49325	0.46284	0.43375	0.40601

TABLA IV : Chi-Cuadrado $Q(\chi^2/v)$

5	0.82084	0.79147	0.76137	0.73079	0.69999	0.66918	0.63857	0.60381	0.57856	0.54942
6	0.90042	0.87949	0.85711	0.83350	0.80885	0.78336	0.75722	0.73062	0.70372	0.67668
7	0.94795	0.93444	0.91938	0.90287	0.88500	0.86590	0.84570	0.82452	0.80250	0.77978
8	0.97426	0.96623	0.95691	0.94628	0.93436	0.92119	0.90681	0.89129	0.87470	0.85712
9	0.98790	0.98345	0.97807	0.97170	0.96430	0.95583	0.94631	0.93572	0.92408	0.91141
10	0.99457	0.99225	0.98934	0.98575	0.98142	0.97632	0.97039	0.96359	0.95592	0.94735
11	0.99766	0.99652	0.99503	0.99311	0.99073	0.98781	0.98431	0.98019	0.97541	0.96992
12	0.99903	0.99850	0.99777	0.99680	0.99554	0.99396	0.99200	0.98962	0.98678	0.98344
13	0.99961	0.99938	0.99903	0.99856	0.99793	0.99711	0.99606	0.99475	0.99314	0.99119
14	0.99985	0.99975	0.99960	0.99938	0.99907	0.99866	0.99813	0.99743	0.99655	0.99547
15	0.99994	0.99990	0.99984	0.99974	0.99960	0.99940	0.99913	0.99878	0.99832	0.99774
16	0.99998	0.99996	0.99994	0.99989	0.99983	0.99974	0.99961	0.99944	0.99921	0.99890
17	0.99999	0.99999	0.99998	0.99996	0.99993	0.99989	0.99983	0.99975	0.99964	0.99948
18			0.99999	0.99998	0.99997	0.99995	0.99993	0.99989	0.99984	0.99976
19				0.99999	0.99999	0.99998	0.99997	0.99995	0.99993	0.99989
20						0.99999	0.99999	0.99998	0.99997	0.99995
21							0.99999	0.99998	0.99997	0.99995
22								0.99999	0.99999	0.99998
									0.99999	0.99999

v	$\chi^2=4.2$	4.4	4.6	4.8	5.0	5.2	5.4	5.6	5.8	6.0
1	0.04042	0.03594	0.03197	0.02846	0.02535	0.02259	0.02014	0.01796	0.01603	0.01431
2	0.12246	0.11080	0.10026	0.09072	0.08209	0.07427	0.06721	0.06081	0.05502	0.04979
3	0.24066	0.22139	0.20354	0.18704	0.17180	0.15772	0.14474	0.13278	0.12176	0.11166
4	0.37962	0.35457	0.33085	0.30844	0.28730	0.26739	0.24866	0.23108	0.21459	0.19915
5	0.52099	0.49337	0.46662	0.44077	0.41688	0.39396	0.36904	0.34711	0.32617	0.30622
6	0.64963	0.62271	0.59604	0.56971	0.54381	0.51843	0.49363	0.46945	0.44596	0.42319
7	0.75647	0.73272	0.70864	0.68435	0.65996x	0.63557	0.61127	0.58715	0.56329	0.53975
8	0.83864	0.81935	0.79935	0.77872	0.75758	0.73600	0.71409	0.69194	0.66962	0.64723
9	0.89776	0.88317	0.86769	0.85138	0.83431	0.81654	0.79814	0.77919	0.75976	0.73992
10	0.93727	0.92750	0.91625	0.90418	0.89118	0.87742	0.86291	0.84768	0.83178	0.81526
11	0.96370	0.95672	0.94898	0.94046	0.93117	0.92109	0.91026	0.89868	0.88637	0.87337
12	0.97955	0.97509	0.97002	0.96433	0.95798	0.95096	0.94327	0.93489	0.92583	0.91608
13	0.98887	0.98614	0.98298	0.97934	0.97519	0.97052	0.96530	0.95951	0.95313	0.94615
14	0.99414	0.99254	0.99064	0.98841	0.98581	0.98283	0.97943	0.97559	0.97128	0.96649
15	0.99701	0.99610	0.99501	0.99369	0.99213	0.99029	0.98816	0.98571	0.98291	0.97975
16	0.99851	0.99802	0.99741	0.99666	0.99575	0.99467	0.99338	0.99187	0.99012	0.98810
17	0.99923	0.99902	0.99869	0.99823	0.99777	0.99715	0.99639	0.99550	0.99443	0.99319
18	0.99966	0.99953	0.99936	0.99914	0.99886	0.99851	0.99809	0.99757	0.99694	0.99620
19	0.99985	0.99978	0.99969	0.99958	0.99943	0.99924	0.99901	0.99872	0.99836	0.99793
20	0.99993	0.99990	0.99986	0.99980	0.99972	0.99962	0.99950	0.99934	0.99914	0.99890
21	0.99997	0.99995	0.99993	0.99991	0.99987	0.99982	0.99975	0.99967	0.99956	0.99943
22	0.99999	0.99998	0.99997	0.99996	0.99994	0.99991	0.99988	0.99984	0.99978	0.99971
23	0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99992	0.99989	0.99986
24			0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99995	0.99993
25					0.99999	0.99999	0.99999	0.99998	0.99998	0.99997
26							0.99999	0.99999	0.99999	0.99998
27								0.99999	0.99999	0.99999

v	$\chi^2=6.2$	6.4	6.6	6.8	7.0	7.2	7.4x	7.6	7.8	8.0
1	0.01278	0.01141	0.01020	0.00912	0.00815	0.00729	0.00652	0.00584	0.00522	0.00468
2	0.04505	0.04076	0.03688	0.03337	0.03020	0.02732	0.02472	0.02237	0.02024	0.01832
3	0.10223	0.09369	0.08580	0.07855	0.07190	0.06579	0.06013	0.05504	0.05033	0.04601
4	0.18470	0.17120	0.15860	0.14684	0.13589	0.12569	0.11620	0.10738	0.09919	0.09158
5	0.28724	0.26922	0.25213	0.23595	0.22064	0.20619	0.19255	0.17970	0.16761	0.15624
6	0.40116	0.37990	0.35943	0.33974	0.32085	0.30275	0.28543	0.26890	0.25313	0.23810
7	0.51660	0.49390	0.47168	0.45000	0.42888	0.40836	0.38845	0.36918	0.35056	0.33259
8	0.62484	0.60252	0.58034	0.55936	0.53863	0.51822	0.49815	0.47849	0.45925	0.44047
9	0.71975	0.69931	0.67869	0.65793	0.63712	0.61631	0.59555	0.57490	0.55442	0.53415
10	0.79819	0.78061	0.76259	0.74418	0.72544	0.70644	0.68722	0.66784	0.64837	0.62884
11	0.85869	0.84539	0.83049	0.81504	0.79908	0.78266	0.76583	0.74862	0.73110x	0.71330
12	0.90567	0.89459	0.88238	0.87054	0.85761	0.84412	0.83009	0.81556	0.80056	0.78513
13	0.93857	0.93038	0.92157	0.91216	0.90215	0.89155	0.88038	0.86865	0.85638	0.84360
14	0.96120	0.95538	0.94903	0.94215	0.93471	0.92673	0.91819	0.90911	0.89948	0.88933
15	0.97619	0.97222	0.96782	0.96296	0.95765	0.95196	0.94599	0.93981	0.93345	0.92695

TABLA IV : Chi-Cuadrado $\alpha(\chi^2/v)$

16	0.98579	0.98317	0.98022	0.97693	0.97328	0.96921	0.96476	0.95989	0.95460	0.94887
17	0.99174	0.99007	0.98816	0.98599	0.98355	0.98081	0.97775	0.97437	0.97064	0.96655
18	0.99532	0.99429	0.99309	0.99171	0.99013	0.98833	0.98630	0.98402	0.98147	0.97864
19	0.99741	0.99679	0.99606	0.99521	0.99421	0.99307	0.99176	0.99026	0.98857	0.98667
20	0.99860	0.99824	0.99781	0.99729	0.99669	0.99598	0.99515	0.99420	0.99311	0.99187
21	0.99926	0.99905	0.99880	0.99850	0.99814	0.99771	0.99721	0.99662	0.99594	0.99514
22	0.99962	0.99950	0.99936	0.99919	0.99898	0.99873	0.99843	0.99807	0.99765	0.99716
23	0.99981	0.99974	0.99967	0.99957	0.99945	0.99931	0.99913	0.99892	0.99867	0.99837
24	0.99990	0.99987	0.99983	0.99978	0.99971	0.99963	0.99953	0.99941	0.99926	0.99908
25	0.99995	0.99994	0.99991	0.99989	0.99985	0.99981	0.99975	0.99968	0.99960	0.99949
26	0.99998	0.99997	0.99996	0.99994	0.99992	0.99990	0.99987	0.99983	0.99978	0.99973
27	0.99999	0.99999	0.99998	0.99997	0.99996	0.99995	0.99993	0.99991	0.99989	0.99985
28		0.99999	0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99992
29				0.99999	0.99999	0.99999	0.99998	0.99998	0.99997	0.99996
30						0.99999	0.99998	0.99998	0.99997	0.99996

v	$\chi^2=8.2$	8.4	8.6	8.8	9.0	9.2	9.4	9.6	9.8	10.0
1	0.00419	0.00375	0.00336	0.003010	0.00270	0.00242	0.00217	0.00195	0.00175	0.00157
2	0.01657	0.01500	0.01357	0.01228	0.01111	0.01005	0.00910	0.00823	0.00745	0.00674
3	0.04205	0.03843	0.03511	0.03207	0.02929	0.02675	0.02442	0.02229	0.02034	0.01857
4	0.08452	0.07798	0.07191	0.06630	0.06110	0.05629	0.05184	0.04773	0.04394	0.04043
5	0.14555	0.13553	0.12612	0.11731	0.10906	0.10135	0.09413	0.08740	0.08110	0.07524
6	0.22381	0.21024	0.19736	0.18514	0.17358	0.16264	0.15230	0.14254	0.13333	0.12465
7	0.31529	0.29865	0.28266	0.26734	0.25266	0.23861	0.22520	0.21240	0.20019	0.18857
8	0.41418	0.39540	0.37715	0.35945	0.34230	0.32571	0.30968	0.29423	0.27935	0.26503
9	0.51412	0.49439	0.47499	0.45594	0.43727	0.41902	0.40120	0.38383	0.36692	0.35049
10	0.60931	0.58983	0.57044	0.55118	0.53210	0.51323	0.49461	0.47626	0.45821	0.44049
11 x	0.69528	0.67709	0.65876	0.64035	0.62189	0.60344	0.58502	0.56669	0.54846	0.53039
12	0.76931	0.75314	0.73666	0.71991	0.70293	0.68576	0.66844	0.65101	0.63350	0.61596
13	0.83033	0.81660	0.80244	0.78788	0.77294	0.75768	0.74211	0.72627	0.71020	0.69393
14	0.87865	0.86746	0.85579	0.84365	0.83105	0.81803	0.80461	0.79081	0.77666	0.76218
15	0.91551	0.90675	0.89749	0.88774	0.87752	0.86683	0.85569	0.84412	0.83213	0.81974
16	0.94269	0.93606	0.92897	0.92142	0.91341	0.90495	0.89603	0.88667	0.87686	0.86663
17	0.96208	0.95723	0.95198	0.94633	0.94026	0.93373	0.92687	0.91954	0.91179	0.90361
18	0.97551	0.97207	0.96830	0.96420	0.95974	0.95493	0.94974	0.94418	0.93824	0.93191
19	0.98454	0.98217	0.97955	0.97666	0.97343	0.97001	0.96623	0.96213	0.95771	0.95295
20	0.99046	0.98887	0.98709	0.98511	0.98291	0.98047	0.97779	0.97486	0.97166	0.96817
21	0.99424	0.99320	0.99203	0.99070	0.98921	0.98755	0.98570	0.98365	0.98139	0.97891
22	0.99659	0.99593	0.99518	0.99431	0.99333	0.99222	0.99098	0.98958	0.98803	0.98630
23	0.99802	0.99761	0.99714	0.99659	0.99596	0.99524	0.99442	0.99349	0.99245	0.99128
24	0.99888	0.99863	0.99833	0.99799	0.99760	0.99714	0.99661	0.99601	0.99532	0.99455
25	0.99937	0.99922	0.99905	0.99884	0.99860	0.99831	0.99798	0.99760	0.99716	0.99665
26	0.99966c	0.99957	0.99947	0.99934	0.99919	0.99902	0.99882	0.99858	0.99830	0.99798
27	0.99981	0.99977	0.99971	0.99963	0.99955	0.99944	0.99932	0.99917	0.99900	0.99880
28	0.99990	0.99987	0.99984	0.99980	0.99975	0.99969	0.99962	0.99953	0.99942	0.99930
29	0.99995	0.99993	0.99991	0.99989	0.99986	0.99983	0.99979	0.99973	0.99967	0.99960
30	0.99997	0.99997	0.99996	0.99994	0.99993	0.99991	0.99988	0.99985	0.99982	0.99977

v	$\chi^2=10.5$	11.0	11.5	12.0	12.5	13.0	13.5	14.0	14.5	15.0
1	0.00119	0.00091	0.00070	0.00053	0.00041	0.00031	0.00024	0.00018	0.00014	0.00011
2	0.00525	0.00409	0.00318	0.00248	0.00193	0.00150	0.00117	0.00091	0.00071	0.00055
3	0.01476	0.01173	0.00931	0.00738	0.00585	0.00464	0.00367	0.00291	0.00230	0.00182
4	0.03280	0.02656	0.02148	0.01735	0.01400	0.01128	0.00907	0.00730	0.00586	0.00470
5	0.06225	0.05138	0.04232	0.03479	0.02854	0.02338	0.01912	0.01631	0.12233	0.01036
6	0.10511	0.08838	0.07410	0.06197	0.05170	0.04304	0.03575	0.02964	0.02452	0.02026
7	0.16196	0.13862	0.11825	0.10056	0.08527	0.07211	0.06082	0.05118	0.04297	0.03600
8	0.23167	0.20170	0.17495	0.15120	0.13025	0.11185	0.09577	0.08177	0.06963	0.05915
9	0.31154	0.27571	0.24299	0.21331	0.18657	0.16261	0.14126	0.12233	0.10562	0.09094
10	0.39777	0.35752	0.31991	0.28506	0.25299	0.22367	0.19704	0.17299	0.15138	0.13206
11	0.48605	0.44326	0.40237	0.36364	0.32726	0.29333	0.26190	0.23299	0.20655	0.18250
12	0.57218	0.52892	0.48662	0.44568	0.40640	0.36904	0.33377	0.30071	0.26992	0.24144
13	0.65263	0.61082	0.56901	0.52764	0.48713	0.44781	0.40997	0.37384	0.33960	0.30735
14	0.72479	0.68604	0.64639	0.60680	0.56622	0.52652	0.48759	0.44971	0.41316	0.37815
15	0.78717	0.75259	0.71641	0.67903	0.64086	0.60230	0.56374	0.52553	0.48900	0.45142

TABLA IV : Chi-Cuadrado $Q(\chi^2/v)$

16	0.83925	0.80949	0.77762	0.74398	0.70890	0.67276	0.63591	0.59871	0.56152	0.52464
17	0.88135	0.85656	0.82942	0.80014	0.76896	0.73619	0.70212	0.66710	0.63145	0.59548
18	0.91436	0.89436	0.87195	0.84724	0.82038	0.79157	0.76106	0.72909	0.69596	0.66197
19	0.93952	0.92384	0.90587	0.88562	0.86316	0.83857	0.81202	0.78639	0.75380	0.72260
20	0.95817	0.94622	0.93221	0.91608	0.89779	0.87738	0.85492	0.83050	0.80427	0.77641
21	0.97166	0.96279	0.95214	0.93962	0.92513	0.90862	0.89010	0.86960	0.84718	0.82295
22	0.98118	0.97475	0.96686	0.95738	0.94618	0.93316	0.91827	0.90148	0.88279	0.86224
23	0.98773	0.98319	0.97748	0.97047	0.96201	0.95199	0.94030	0.92687	0.91165	0.89463
24	0.99216	0.98901	0.98498	0.97991	0.97367	0.96612	0.95715	0.94665	0.93454	0.92076
25	0.99507	0.99295	0.99015	0.98657	0.98206	0.97650	0.96976	0.96173	0.95230	0.94138
26	0.99696	0.99555	0.99366	0.99117	0.98798	0.98397	0.97902	0.97300	0.96581	0.95733
27	0.99815	0.99724	0.99598	0.99429	0.99208	0.98925	0.98567	0.98125	0.97588	0.96943
28	0.99890	0.99831	0.99749	0.99637	0.99487	0.99290	0.99037	0.98719	0.98324	0.97844
29	0.99935	0.99899	0.99846	0.99773	0.99672	0.99538	0.99363	0.99138	0.98854	0.98502
30	0.99963	0.99940	0.99907	0.99860	0.99794	0.99704	0.99585	0.99428	0.99227	0.98974

v	$\chi^2=15.5$	16.0	16.5	17.0	17.5	18.0	18.5	19.0	19.5	20.0
1	0.00008	0.00006	0.00005	0.00004	0.00003	0.00002	0.00002	0.00001	0.00001	0.00001
2	0.00043	0.00034	0.00026	0.00020	0.00016	0.00012	0.00010	0.00008	0.00006	0.00005
3	0.00144	0.00113	0.00090	0.00071	0.00056	0.00044	0.00035	0.00027	0.00022	0.00017
4	0.00377	0.00302	0.00242	0.00193	0.00154	0.00123	0.00099	0.00079	0.00063	0.00050
5	0.00843	0.00684	0.00555	0.00450	0.00364	0.00295	0.00238	0.00192	0.00155	0.00125
6	0.01670	0.01375	0.01131	0.00928	0.00761	0.00623	0.00510	0.00416	0.00340	0.00277
7	0.03010	0.02512	0.02092	0.01640	0.01444	0.01197	0.00991	0.00819	0.00676	0.00557
8	0.05012	0.04238	0.03576	0.03011	0.02530	0.02123	0.01777	0.01486	0.01240	0.01034
9	0.07809	0.06688	0.05715	0.04872	0.04144	0.03517	0.02980	0.02519	0.02126	0.01791
10	0.11487	0.09963	0.08619	0.07436	0.06401	0.05496	0.04709	0.04026	0.03435	0.02925
11	0.16073	0.14113	0.12356	0.10788	0.09393	0.08158	0.07068	0.06109	0.05269	0.04534
12	0.21522	0.19124	0.16939	0.14960	0.13174	0.11569	0.10133	0.08853	0.07716	0.06709
13	0.27719	0.24913	0.22318	0.19930	0.17744	0.15752	0.13944	0.12310	0.10840	0.09521
14	0.34485	0.31337	0.28380	0.25618	0.23051	0.20678	0.18495	0.16495	0.14671	0.13014
15	0.41604	0.38205	0.34962	0.31886	0.28986	0.26267	0.23729	0.21373	0.19196	0.17193
16	0.48837	0.45296	0.41864	0.38560	0.35398	0.32390	0.29544	0.26866	0.24359	0.22022
17	0.55951	0.52383	0.48871	0.45437	0.42102	0.38884	0.35797	0.32853	0.30060	0.27423
18	0.62740	0.59255	0.55770	0.52311	0.48902	0.45565	0.42320	0.39182	0.36166	0.33282
19	0.69033	0.65728	0.62370	0.58987	0.55603	0.52244	0.48931	0.45684	0.42521	0.39458
20	0.74712	0.71662	0.68516	0.65297	0.62031	0.58741	0.55451	0.52183	0.48957	0.45793
21	0.79705	0.76965	0.74093	0.71111	0.68039	0.64900	0.61718	0.58514	0.55310	0.52126
22	0.83990	0.81589	0.79032	0.76336	0.73519	0.70599	0.67597	0.64533	0.61428	0.58304
23	0.87582	0.85527	0.83304	0.80925	0.78402	0.75749	0.72983	0.70122	0.67185	0.64191
24	0.90527	0.88808	0.86919	0.84866	0.82657	0.80302	0.77801	0.75199	0.72483	0.69678
25	0.92891	0.91483	0.89912	0.88179	0.86287	0.84239	0.82044	0.79712	0.77254	0.74683
26	0.94749	0.93620	0.92341	0.90908	0.89320	0.87577	0.85683	0.83643	0.81464	0.79156
27	0.96182	0.95295	0.94274	0.93112	0.91806	0.90352	0.88750	0.87000	0.85107	0.83076
28	0.97266	0.96582	0.95782	0.94859	0.93805	0.92615	0.91285	0.89814	0.88200	0.86446
29	0.98071	0.97554	0.96939	0.96218	0.95383	0.94427	0.93344	0.92129	0.90779	0.89293
30	0.98659	0.98274	0.97810	0.97258	0.96608	0.95853	0.94986	0.94001	0.92891	0.91654

v	$\chi^2=21$	22	23	24	25	26	27	28	29	30
1	0.00001									
2	0.00003	0.00002	0.00001	0.00001						
3	0.00011	0.00007	0.00004	0.00003	0.00002	0.00001	0.00001	0.		
4	0.00032	0.00020	0.00013	0.00008	0.00005	0.00003	0.00002	0.00001	0.00001	0.00001
5	0.00081	0.00052	0.00034	0.00022	0.00014	0.00009	0.00006	0.00004	0.00002	0.00002
6	0.00184	0.00121	0.00080	0.00052	0.00034	0.00022	0.00015	0.00009	0.00006	0.00004
7	0.00377	0.00254	0.00171	0.00114	0.00076	0.00050	0.00033	0.00022	0.00015	0.00010
8	0.00715	0.00492	0.00336	0.00229	0.00155	0.00105	0.00071	0.00047	0.00032	0.00021
9	0.01265	0.00888	0.00620	0.00430	0.00297	0.00204	0.00140	0.00095	0.00065	0.00044
10	0.02109	0.01511	0.01075	0.00760	0.00535	0.00374	0.00260	0.00181	0.00125	0.00086
11	0.03337	0.02437	0.01768	0.01273	0.00912	0.00649	0.00460	0.00324	0.00227	0.00159
12	0.05038	0.03752	0.02773	0.02034	0.01482	0.01073	0.00773	0.00553	0.00394	0.00279
13	0.07293	0.05536	0.04168	0.03113	0.02308	0.01700	0.01244	0.00905	0.00655	0.00471
14	0.10163	0.07861	0.06027	0.04582	0.03457	0.02589	0.01925	0.01423	0.01045	0.00763
15	0.13683	0.10780	0.08414	0.06509	0.04994	0.03802	0.02874	0.02157	0.01609	0.01192

TABLA IV : Chi-Cuadrado $Q(\chi^2/v)$

16	0.17851	0.14319	0.11374	0.08950	0.06982	0.05403	0.04148	0.03162	0.02394	0.01800
17	0.22629	0.18472	0.14925	0.11944	0.09471	0.07446	0.05807	0.04494	0.03453	0.02635
18	0.27941	0.23199	0.19059	0.15503	0.12492	0.09976	0.07900	0.06206	0.04838	0.03745
19	0.33680	0.28426	0.23734	0.19615	0.16054	0.13019	0.10465	0.08343	0.06599	0.05180
20	0.39713	0.34051	0.28880	0.24239	0.20143	0.16581	0.13526	0.10940	0.08776	0.06985
21	0.45894	0.39951	0.34398	0.29306	0.24716	0.20645	0.17085	0.14015	0.11400	0.09199
22	0.52074	0.45989	0.40173	0.34723	0.29707	0.25168	0.21123	0.17568	0.14486	0.11846
23	0.58109	0.52025	0.46077	0.40381	0.35029	0.30087	0.25597	0.21508	0.18031	0.14940
24	0.63873	0.57927	0.51980	0.46160	0.40576	0.35317	0.30445	0.26004	0.22013	0.18475
25	0.69261	0.63574	0.57756	0.51937	0.46237	0.40760	0.35588	0.30785	0.26392	0.22429
26	0.74196	0.68870	0.63295	0.57597	0.51898	0.46311	0.40993	0.35846	0.31108	0.26761
27	0.78629	0.73738	0.68501	0.63032	0.57446	0.51860	0.46379	0.41097	0.36090	0.31415
28	0.82535	0.78129	0.73304	0.68154	0.62784	0.57305	0.51825	0.46445	0.41253	0.36322
29	0.85915	0.82019	0.77654	0.72893	0.67825	0.62549	0.57171	0.51791	0.46507	0.41400
30	0.88789	0.85404	0.81526	0.77203	0.72503	0.67513	0.62326	0.57044	0.51760	0.46565

v	$\chi^2=31$	32	33	34	35	36	37	38	39	40
5	0.00001	0.00001								
6	0.00003	0.00002	0.00001	0.00001	0.					
7	0.00006	0.00004	0.00003	0.00002	0.00001	0.00001				
8	0.00014	0.00009	0.00006	0.00004	0.00003	0.00002	0.00001	0.00001		
9	0.00030	0.00020	0.00013	0.00009	0.00006	0.00004	0.00003	0.00002	0.00001	0.00001
10	0.00059	0.00040	0.00027	0.00019	0.00012	0.00008	0.00006	0.00004	0.00003	0.00002
11	0.00110	0.00076	0.00053	0.00036	0.00025	0.00017	0.00012	0.00008	0.00005	0.00004
12	0.00197	0.00138	0.00097	0.00068	0.00047	0.00032	0.00022	0.00014	0.00011	0.00007
13	0.00337	0.00240	0.00170	0.00120	0.00085	0.00059	0.00041	0.00029	0.00020	0.00014
14	0.00554	0.00401	0.00288	0.00206	0.00147	0.00104	0.00074	0.00052	0.00036	0.00026
15	0.00878	0.00644	0.00469	0.00341	0.00246	0.00177	0.00127	0.00090	0.00064	0.00045
16	0.01346	0.01000	0.00739	0.00543	0.00397	0.00289	0.00210	0.00151	0.00109	0.00078
17	0.01997	0.01505	0.01127	0.00840	0.00622	0.00459	0.00337	0.00246	0.00179	0.00129
18	0.02879	0.02199	0.01669	0.01260	0.00945	0.00706	0.00524	0.00387	0.00285	0.00209
19	0.04037	0.03125	0.02404	0.01838	0.01397	0.01056	0.00793	0.00593	0.00442	0.00327
20	0.05519	0.04330	0.03374	0.02613	0.02010	0.01538	0.01170	0.00836	0.00667	0.00500
21	0.07366	0.05855	0.04622	0.03624	0.02824	0.02187	0.01683	0.01289	0.00981	0.00744
22	0.09612	0.07740	0.06187	0.04912	0.03875	0.03037	0.02366	0.01832	0.01411	0.01081
23	0.12279	0.10014	0.08107	0.06516	0.05202	0.04125	0.03251	0.02540	0.01984	0.01437
24	0.15378	0.12699	0.10407	0.08467	0.06840	0.05489	0.04376	0.03467	0.02731	0.02139
25	0.18902	0.15801	0.13107	0.10791	0.08820	0.07160	0.05774	0.04626	0.03684	0.02916
26	0.22827	0.19312	0.16210	0.13502	0.11165	0.09167	0.07475	0.06056	0.04875	0.03901
27	0.27114	0.23208	0.19707	0.16605	0.13887	0.11530	0.09507	0.07786	0.06336	0.05124
28	0.31708	0.27451	0.23574	0.20087	0.16987	0.14260	0.11886	0.09840	0.08092	0.06613
29	0.36542	0.31987	0.27774	0.23926	0.20454	0.17356	0.14622	0.12234	0.10166	0.08394
30	0.41541	0.36753	0.32254	0.28083	0.24264	0.20808	0.17714	0.14975	0.12573	0.10486

v	$\chi^2=42$	44	46	48	50	52	54	56	58	60
10	0.00001									
11	0.00002	0.00001								
12	0.00003	0.00002	0.00001							
13	0.00006	0.00003	0.00001	0.00001						
14	0.00012	0.00006	0.00003	0.00001	0.00001					
15	0.00023	0.00011	0.00005	0.00003	0.00001	0.00001				
16	0.00040	0.00020	0.00010	0.00005	0.00002	0.00001	0.00001			
17	0.00067	0.00034	0.00017	0.00009	0.00004	0.00002	0.00001	0.00001		
18	0.00111	0.00058	0.00030	0.00015	0.00008	0.00004	0.00002	0.00001		
19	0.00177	0.00094	0.00050	0.00026	0.00013	0.00007	0.00003	0.00002	0.00001	
20	0.00277	0.00151	0.00081	0.00043	0.00022	0.00011	0.00006	0.00003	0.00001	0.00001
21	0.00421	0.00234	0.00128	0.00069	0.00036	0.00019	0.00010	0.00005	0.00003	0.00001
22	0.00625	0.00355	0.00198	0.00109	0.00059	0.00031	0.00016	0.00009	0.00004	0.00002
23	0.00908	0.00526	0.00299	0.00167	0.00092	0.00050	0.00027	0.00014	0.00007	0.00004
24	0.01291	0.00763	0.00443	0.00252	0.00142	0.00078	0.00043	0.00023	0.00012	0.00006
25	0.01797	0.01085	0.00642	0.00373	0.00213	0.00120	0.00066	0.00036	0.00020	0.00011
26	0.02455	0.01512	0.00912	0.00540	0.00314	0.00180	0.00102	0.00056	0.00031	0.00017
27	0.03292	0.02068	0.01272	0.00768	0.00455	0.00265	0.00152	0.00086	0.00048	0.00026
28	0.04336	0.02779	0.01743	0.01072	0.00647	0.00384	0.00234	0.00139	0.00072	0.00041

TABLA IV : Chi-Cuadrado $Q(\chi^2/v)$

v	$\chi^2=62$	64	66	68	70	72	74	76
29	0.05616	0.03670	0.02346	0.01470	0.00903	0.00545	0.00324	0.00189
30	0.07157	0.04769	0.03107	0.01983	0.01240	0.00762	0.00460	0.00273
21	0.00001	0.						
22	0.00001	0.00001						
23	0.00002	0.00001	0.00001					
24	0.00003	0.00002	0.00001					
25	0.00006	0.00003	0.00002	0.00001				
26	0.00009	0.00005	0.00003	0.00001	0.00001			
27	0.00014	0.00008	0.00004	0.00002	0.00001	0.00001		
28	0.00023	0.00012	0.00007	0.00004	0.00002	0.00001	0.00001	
29	0.00035	0.00019	0.00011	0.00006	0.00003	0.00002	0.00001	
30	0.00052	0.00029	0.00016	0.00009	0.00005	0.00003	0.00001	0.00001

TABLA V : Distribución de Fisher

v_1	1	2	3	4	5	6	7	8	9	10	12	16	20	30	50	100	500	∞
v_2										10								
1	161 4,052	200 4,999	216 5,403	225 5,625	230 5,764	234 5,859	237 5,928	239 5,981	241 6,022	242 6,056	244 6,106	246 6,169	248 6,208	250 6,258	252 6,302	253 6,334	254 6,361	254 6,366
2	18.51 98.49	19.00 99.00	19.16 99.17	19.25 99.25	19.30 99.30	19.33 99.33	19.36 99.34	19.37 99.36	19.38 99.38	19.39 99.40	19.41 99.42	19.43 99.44	19.44 99.45	19.46 99.47	19.47 99.48	19.49 99.49	19.50 99.50	19.50 99.50
3	10.13 34.12	9.55 30.82	9.28 29.46	9.12 28.71	9.01 28.24	8.94 27.91	8.88 27.67	8.84 27.49	8.81 27.34	8.78 27.23	8.74 27.05	8.69 26.83	8.66 26.69	8.62 26.50	8.58 26.35	8.56 26.23	8.54 26.14	8.53 26.12
4	7.71 21.20	6.94 18.00	6.59 16.69	6.39 15.98	6.26 15.52	6.16 15.21	6.09 14.98	6.04 14.80	6.00 14.66	5.96 14.54	5.91 14.37	5.84 14.15	5.80 14.02	5.74 13.83	5.70 13.69	5.66 13.57	5.64 13.48	5.63 13.46
5	6.61 16.26	5.79 13.27	5.41 12.06	5.19 11.39	5.05 10.97	4.95 10.67	4.88 10.45	4.82 10.27	4.78 10.15	4.74 10.05	4.68 9.89	4.60 9.68	4.56 9.55	4.50 9.38	4.44 9.24	4.40 9.13	4.37 9.04	4.36 9.02
6	5.99 13.74	5.14 10.92	4.76 9.78	4.53 9.15	4.39 8.75	4.28 8.47	4.21 8.26	4.15 8.10	4.10 7.98	4.06 7.87	4.00 7.72	3.92 7.52	3.87 7.39	3.81 7.23	3.75 7.09	3.71 6.99	3.68 6.90	3.67 6.88
7	5.59 12.25	4.74 9.55	4.35 8.45	4.12 7.85	3.97 7.46	3.87 7.19	3.79 7.00	3.73 6.84	3.68 6.71	3.63 6.62	3.57 6.47	3.49 6.27	3.44 6.15	3.38 5.98	3.32 5.85	3.28 5.75	3.24 5.67	3.23 5.65
8	5.32 11.26	4.46 8.65	4.07 7.59	3.84 7.01	3.69 6.63	3.58 6.37	3.50 6.19	3.44 6.03	3.39 5.91	3.34 5.82	3.28 5.67	3.20 5.48	3.15 5.36	3.08 5.20	3.03 5.06	2.98 4.96	2.94 4.88	2.93 4.86
9	5.12 10.56	4.26 8.02	3.86 6.99	3.63 6.42	3.48 6.06	3.37 5.80	3.29 5.62	3.23 5.47	3.18 5.35	3.13 5.26	3.07 5.11	2.98 4.92	2.93 4.80	2.86 4.64	2.80 4.51	2.76 4.41	2.72 4.33	2.71 4.31
10	4.96 10.04	4.10 7.56	3.71 6.55	3.48 5.99	3.33 5.64	3.22 6.39	3.14 5.21	3.07 5.06	3.02 4.95	2.97 4.85	2.91 4.71	2.82 4.52	2.77 4.41	2.70 4.25	2.64 4.12	2.59 4.01	2.55 3.93	2.54 3.91
11	4.84 9.65	3.98 7.20	3.59 6.22	3.36 5.67	3.20 5.32	3.09 5.07	3.01 4.88	2.95 4.74	2.90 4.63	2.86 4.54	2.79 4.40	2.70 4.21	2.65 4.10	2.57 3.94	2.50 3.80	2.45 3.70	2.41 3.62	2.40 3.60
12	4.75 9.33	3.88 6.93	3.49 5.95	3.26 5.41	3.11 5.06	3.00 4.82	2.92 4.65	2.85 4.50	2.80 4.39	2.76 4.30	2.69 4.16	2.60 3.98	2.54 3.86	2.46 3.70	2.40 3.56	2.35 3.46	2.31 3.38	2.30 3.36
13	4.67 9.07	3.80 6.70	3.41 5.74	3.18 5.20	3.02 4.86	2.92 4.62	2.84 4.44	2.77 4.30	2.72 4.19	2.67 4.10	2.60 3.96	2.51 3.78	2.46 3.67	2.38 3.51	2.32 3.37	2.25 3.27	2.22 3.18	2.21 3.16
14	4.60 8.86	3.74 6.51	3.34 5.56	3.11 5.03	2.96 4.69	2.85 4.46	2.77 4.28	2.70 4.14	2.65 4.03	2.60 3.94	2.53 3.80	2.44 3.62	2.39 3.51	2.31 3.34	2.24 3.21	2.19 3.11	2.14 3.02	2.13 3.00
15	4.54 8.68	3.68 6.36	3.29 5.42	3.06 4.89	2.90 4.56	2.79 4.32	2.70 4.14	2.64 4.00	2.59 3.89	2.55 3.80	2.48 3.67	2.39 3.48	2.33 3.36	2.25 3.20	2.18 3.07	2.12 2.97	2.08 2.89	2.07 2.87
16	4.49 8.53	3.63 6.23	3.24 5.29	3.01 4.77	2.85 4.44	2.74 4.20	2.66 4.03	2.59 3.89	2.54 3.78	2.49 3.69	2.42 3.55	2.33 3.37	2.28 3.25	2.20 3.10	2.13 2.96	2.07 2.86	2.02 2.77	2.01 2.75
17	4.45 8.40	3.59 6.11	3.20 5.18	2.96 4.67	2.81 4.34	2.70 4.10	2.62 3.93	2.55 3.79	2.50 3.68	2.45 3.59	2.38 3.45	2.29 3.27	2.23 3.16	2.15 3.00	2.08 2.86	2.02 2.76	1.97 2.67	1.96 2.65
18	4.41 8.28	3.55 6.01	3.16 5.09	2.93 4.58	2.77 4.25	2.66 4.01	2.58 3.85	2.51 3.71	2.46 3.60	2.41 3.51	2.34 3.37	2.25 3.19	2.19 3.07	2.11 2.91	2.04 2.78	1.98 2.68	1.93 2.59	1.92 2.57
19	4.38 8.18	3.52 5.93	3.13 5.01	2.90 4.50	2.74 4.17	2.63 3.94	2.55 3.77	2.48 3.63	2.43 3.52	2.38 3.43	2.31 3.30	2.21 3.12	2.15 3.00	2.07 2.84	2.00 2.70	1.94 2.60	1.90 2.51	1.88 2.49
20	4.35 8.10	3.49 5.85	3.10 4.94	2.87 4.43	2.71 4.10	2.60 3.87	2.52 3.71	2.45 3.56	2.40 3.45	2.35 3.37	2.28 3.23	2.18 3.05	2.12 2.94	2.04 2.77	1.96 2.63	1.90 2.53	1.85 2.44	1.84 2.42
21	4.32 8.02	3.47 5.78	3.07 4.87	2.84 4.37	2.68 4.04	2.57 3.81	2.49 3.65	2.42 3.51	2.37 3.40	2.32 3.31	2.25 3.17	2.15 3.09	2.09 2.88	2.00 2.72	1.93 2.58	1.87 2.47	1.82 2.38	1.81 2.86
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.23	2.13	2.07	1.99	1.91	1.84	1.80	1.79

v_1	1	2	3	4	5	6	7	8	9	10	12	16	20	30	50	100	500	∞
v_2	1																	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.20	2.10	2.04	1.96	1.88	1.82	1.77	1.06
	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.89	2.78	2.62	2.48	2.37	2.28	2.26
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.18	2.09	2.02	1.94	1.86	1.80	1.74	1.73
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.25	3.17	3.03	2.85	2.74	2.58	2.44	2.33	2.23	2.21
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.16	2.06	2.00	1.92	1.84	1.77	1.72	1.71
	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.21	3.13	2.99	2.81	2.70	2.54	2.40	2.29	2.19	2.17
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.05	1.99	1.90	1.82	1.76	1.70	1.69
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	2.96	2.77	2.66	2.50	2.36	2.25	2.15	2.13
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.13	2.03	1.97	1.88	1.80	1.74	1.68	1.67
	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.14	3.06	2.93	2.74	2.63	2.47	2.33	2.21	2.12	2.10
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.12	2.02	1.96	1.87	1.78	1.72	1.67	1.65
	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11	3.03	2.90	2.71	2.60	2.44	2.30	2.18	2.09	2.06
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.10	2.00	1.94	1.85	1.77	1.71	1.65	1.64
	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.08	3.00	2.87	2.68	2.57	2.41	2.27	2.15	2.06	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.09	1.99	1.93	1.84	1.76	1.69	1.64	1.62
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06	2.98	2.84	2.66	2.55	2.38	2.24	2.13	2.03	2.01
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.07	1.97	1.91	1.82	1.74	1.67	1.61	1.59
	7.50	5.34	4.46	3.97	3.66	3.42	3.25	3.12	3.01	2.94	2.80	2.62	2.51	2.34	2.20	2.08	1.98	1.96
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.05	1.95	1.89	1.80	1.71	1.64	1.59	1.57
	7.44	5.29	4.42	3.93	3.61	3.38	3.21	3.08	2.97	2.89	2.76	2.58	2.47	2.30	2.15	2.04	1.94	1.91
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.03	1.93	1.87	1.78	1.69	1.62	1.56	1.55
	7.39	5.25	4.38	3.89	3.58	3.35	3.18	3.04	2.94	2.86	2.72	2.54	2.43	2.26	2.12	2.00	1.90	1.87
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.02	1.92	1.85	1.76	1.67	1.60	1.54	1.53
	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.91	2.82	2.69	2.51	2.40	2.22	2.08	1.97	1.86	1.84
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.00	1.90	1.84	1.74	1.66	1.59	1.53	1.51
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88	2.80	2.66	2.49	2.37	2.20	2.05	1.94	1.84	1.81
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	1.99	1.89	1.82	1.73	1.64	1.57	1.51	1.49
	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86	2.77	2.64	2.46	2.35	2.17	2.02	1.91	1.80	1.78
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	1.98	1.88	1.81	1.72	1.63	1.56	1.50	1.48
	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84	2.75	2.62	2.44	2.32	2.15	2.00	1.88	1.78	1.75
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.15	2.09	2.04	1.97	1.87	1.80	1.71	1.62	1.54	1.48	1.46
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.60	2.42	2.30	2.13	1.98	1.86	1.76	1.72
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.96	1.86	1.79	1.70	1.61	1.53	1.47	1.45
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.58	2.40	2.28	2.11	1.96	1.84	1.73	1.70
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.95	1.85	1.78	1.69	1.60	1.62	1.46	1.44
	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.56	2.39	2.26	2.10	1.94	1.82	1.71	1.68
55	4.02	3.17	2.88	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.93	1.83	1.76	1.67	1.58	1.50	1.43	1.41
	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66	2.53	2.35	2.23	2.06	1.90	1.78	1.66	1.64
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.81	1.75	1.65	1.56	1.48	1.41	1.39
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.32	2.20	2.03	1.87	1.74	1.63	1.60
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.90	1.80	1.73	1.63	1.54	1.46	1.39	1.37
	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.79	2.70	2.61	2.47	2.30	2.18	2.00	1.84	1.71	1.60	1.56
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.89	1.79	1.72	1.62	1.53	1.45	1.37	1.35
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.59	2.45	2.28	2.15	1.98	1.82	1.69	1.56	1.53

TABLA V : Distribución de Fisher

iv-12

v_2	1	2	3	4	5	6	7	8	9	10	12	16	20	30	50	100	500	∞	
v_2																			
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.88	1.77	1.70	1.60	1.51	1.42	1.35	1.32	
	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.41	2.24	2.11	1.94	1.78	1.65	1.52	1.49	
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.85	1.75	1.68	1.57	1.48	1.39	1.30	1.28	
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.36	2.19	2.06	1.89	1.73	1.59	1.46	1.43	
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.83	1.72	1.65	1.55	1.45	1.36	1.27	1.25	
	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.65	2.56	2.47	2.33	2.15	2.03	1.85	1.68	1.54	1.40	1.37	
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.82	1.71	1.64	1.54	1.44	1.34	1.25	1.22	
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.30	2.12	2.00	1.83	1.66	1.51	1.37	1.33	
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.80	1.69	1.62	1.52	1.42	1.32	1.22	1.19	
	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.28	2.09	1.97	1.79	1.62	1.48	1.33	1.28	
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.78	1.67	1.60	1.49	1.38	1.28	1.16	1.13	
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46	2.37	2.23	2.04	1.92	1.74	1.57	1.42	1.24	1.19	
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.76	1.65	1.58	1.47	1.36	1.26	1.13	1.08	
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.20	2.01	1.89	1.71	1.54	1.38	1.19	1.11	
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.75	1.64	1.57	1.46	1.35	1.24	1.11	1.00	
	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	1.99	1.87	1.69	1.52	1.36	1.15	1.00	

TABLA VI : Valores de $e^{-\lambda}$

λ	0	1	2	3	4	5	6	7	8	9
0.0	1.0000	.9900	.9802	.9704	.9608	.9512	.9418	.9324	.9231	.9139
0.1	.9048	.8958	.8869	.8781	.8694	.8607	.8521	.8437	.8353	.8270
0.2	.8187	.8106	.8025	.7945	.7866	.7788	.7711	.7634	.7558	.7483
0.3	.7408	.7334	.7261	.7189	.7118	.7047	.6977	.6907	.6839	.6771
0.4	.6703	.6636	.6570	.6505	.6440	.6376	.6313	.6250	.6188	.6126
0.5	.6065	.6005	.5945	.5886	.5827	.5770	.5712	.5655	.5599	.5543
0.6	.5488	.5434	.5379	.5326	.5273	.5220	.5169	.5117	.5066	.5016
0.7	.4966	.4916	.4868	.4819	.4771	.4724	.4677	.4630	.4584	.4538
0.8	.4493	.4449	.4404	.4360	.4317	.4274	.4232	.4190	.4148	.4107
0.9	.4066	.4025	.3985	.3946	.3906	.3867	.3829	.3791	.3753	.3716
$e^{-\lambda}$										
	1	2	3	4	5	6	7	8	9	10
	.367880	.135340	.049790	.018320	.006738	.002479	.000912	.000335	.000123	.000045

TABLA VII : Función Gamma

z	$\Gamma(z)$	$\ln \Gamma(z)$	z	$\Gamma(z)$	$\ln \Gamma(z)$
1.500	0.88622	-0.12078	1.750	0.91906	-0.08440
1.505	0.88639	-0.12058	1.755	0.92020	-0.08315
1.510	0.88659	-0.12037	1.760	0.92137	-0.08188
1.515	0.88680	-0.12013	1.765	0.92255	-0.08060
1.520	0.88703	-0.11986	1.770	0.92376	-0.07929
1.525	0.88729	-0.11957	1.775	0.92498	-0.07797
1.530	0.88756	-0.11927	1.780	0.92622	-0.07663
1.535	0.88786	-0.11893	1.785	0.92748	-0.07527
1.540	0.88817	-0.11858	1.790	0.92876	-0.07389
1.545	0.88851	-0.11820	1.795	0.93006	-0.07249
1.550	0.88886	-0.11780	1.8000	0.93138	-0.07108
1.555	0.88924	-0.11738	1.805	0.93272	-0.06964
1.560	0.88963	-0.11693	1.810	0.93407	-0.06819
1.565	0.89005	-0.11647	1.815	0.93545	-0.06672
1.570	0.89048	-0.11598	1.820	0.93684	-0.06523
1.575	0.89094	-0.11547	1.825	0.93825	-0.06373
1.580	0.89141	-0.11494	1.830	0.93969	-0.06220
1.585	0.89191	-0.11438	1.835	0.94114	-0.06066
1.590	0.89242	-0.11380	1.840	0.94261	-0.05910
1.595	0.89296	-0.11321	1.845	0.94410	-0.05752
1.600	0.89351	-0.11259	1.850	0.94561	-0.05592
1.605	0.89408	-0.11195	1.855	0.94713	-0.05430
1.610	0.89468	-0.11128	1.860	0.94868	-0.05267
1.615	0.89529	-0.11060	1.865	0.95025	-0.05102
1.620	0.89592	-0.10990	1.870	0.95184	-0.04935
1.625	0.89657	-0.10917	1.875	0.95344	-0.04767
1.630	0.89724	-0.10842	1.880	0.95507	-0.04596
1.635	0.89793	-0.10765	1.885	0.95671	-0.04424
1.640	0.89864	-0.10687	1.890	0.95837	-0.04251
1.645	0.89936	-0.10606	1.895	0.96006	-0.04075
1.650	0.90011	-0.10523	1.900	0.96176	-0.03898
1.655	0.90088	-0.10437	1.905	0.96348	-0.03719
1.660	0.90166	-0.10350	1.910	0.96523	-0.03538
1.665	0.90247	-0.10261	1.915	0.96699	-0.03356
1.670	0.90329	-0.10170	1.920	0.96877	-0.03172
1.675	0.90413	-0.10077	1.925	0.97057	-0.02986
1.680	0.90500	-0.09981	1.930	0.97239	-0.02799
1.685	0.90588	-0.09884	1.935	0.97423	-0.02609
1.690	0.90678	-0.09785	1.940	0.97609	-0.02419
1.695	0.90770	-0.09684	1.945	0.97797	-0.02226
1.700	0.90863	-0.09580	1.950	0.97988	-0.02032
1.705	0.90959	-0.09475	1.955	0.98180	-0.01836
1.710	0.91057	-0.09368	1.960	0.98374	-0.01639
1.715	0.91156	-0.09259	1.965	0.98570	-0.01439
1.720	0.91258	-0.09147	1.970	0.98768	-0.01239
1.725	0.91361	-0.09034	1.975	0.98968	-0.01036
1.730	0.91466	-0.08919	1.980	0.99170	-0.00832
1.735	0.91573	-0.08802	1.985	0.99375	-0.00626
1.740	0.91682	-0.08683	1.990	0.99581	-0.00419
1.745	0.91793	-0.08562	1.995	0.99789	-0.00210

TABLA VII : Función Gamma

z	$\Gamma(z)$	$\ln \Gamma(z)$	z	$\Gamma(z)$	$\ln \Gamma(z)$
1.000	1.00000	0.00000	1.250	0.90640	-0.09827
1.005	0.99713	-0.00286	1.255	0.90538	-0.09939
1.010	0.99432	-0.00569	1.260	0.90439	-0.10048
1.015	0.99156	-0.00847	1.265	0.90343	-0.10154
1.020	0.98884	-0.01121	1.270	0.90250	-0.10258
1.025	0.98617	-0.01392	1.275	0.90159	-0.10358
1.030	0.98354	-0.01658	1.280	0.90071	-0.10456
1.035	0.98097	-0.01921	1.285	0.89986	-0.10550
1.040	0.97843	-0.02179	1.290	0.89904	-0.10642
1.045	0.97594	-0.02434	1.295	0.89824	-0.10731
1.050	0.97350	-0.02685	1.300	0.89747	-0.10817
1.055	0.97110	-0.02932	1.305	0.89672	-0.10900
1.060	0.96874	-0.03175	1.310	0.89600	-0.10981
1.065	0.96642	-0.03414	1.315	0.89530	-0.11058
1.070	0.96415	-0.03650	1.320	0.89464	-0.11133
1.075	0.96191	-0.03882	1.325	0.89399	-0.11205
1.080	0.95972	-0.04110	1.330	0.89337	-0.11274
1.085	0.95757	-0.04335	1.335	0.89278	-0.11341
1.090	0.95545	-0.04556	1.340	0.89221	-0.11404
1.095	0.95338	-0.04773	1.345	0.89167	-0.11465
1.100	0.95135	-0.04987	1.350	0.89115	-0.11524
1.105	0.94935	-0.05197	1.355	0.89065	-0.11579
1.110	0.94739	-0.05403	1.360	0.89018	-0.11632
1.115	0.94547	-0.05606	1.365	0.88973	-0.11682
1.120	0.94359	-0.05806	1.370	0.88931	-0.11730
1.125	0.94174	-0.06002	1.375	0.88891	-0.11775
1.130	0.93993	-0.06194	1.380	0.88853	-0.11817
1.135	0.93815	-0.06383	1.385	0.88818	-0.11857
1.140	0.93641	-0.06569	1.390	0.88785	-0.11894
1.145	0.93471	-0.06751	1.395	0.88754	-0.11929
1.150	0.93304	-0.06930	1.400	0.88726	-0.11961
1.155	0.93140	-0.07106	1.405	0.88700	-0.11990
1.160	0.92980	-0.07278	1.410	0.88676	-0.12017
1.165	0.92823	-0.07447	1.415	0.88654	-0.12041
1.170	0.92669	-0.07612	1.420	0.88635	-0.12063
1.175	0.92519	-0.07774	1.425	0.88618	-0.12082
1.180	0.92372	-0.07933	1.430	0.88603	-0.12099
1.185	0.92229	-0.08089	1.435	0.88590	-0.12114
1.190	0.92088	-0.08242	1.440	0.88580	-0.12125
1.195	0.91951	-0.08391	1.445	0.88572	-0.12135
1.200	0.91816	-0.08537	1.450	0.88566	-0.12142
1.205	0.91685	-0.08680	1.455	0.88562	-0.12146
1.210	0.91557	-0.08820	1.460	0.88560	-0.12148
1.215	0.91432	-0.08956	1.465	0.88560	-0.12148
1.220	0.91310	-0.09090	1.470	0.88563	-0.12145
1.225	0.91191	-0.09220	1.475	0.88567	-0.12140
1.230	0.91075	-0.09348	1.480	0.88574	-0.12132
1.235	0.90962	-0.09472	1.485	0.88583	-0.12122
1.240	0.90852	-0.09593	1.490	0.88594	-0.12110
1.245	0.90744	-0.09711	1.495	0.88607	-0.12095

B I B L I O G R A F I A

1. M. H. Quenouille-Fundamentals of Statistical Reasoning. Hafner Publishing Company, New York.
2. Worthing Archie & Geffner Joseph-Treatment of experimental data. John Wiley & Sons, New York.
3. Brownlee K. A.-Statistical Theory and Methodology in Science and Engineering. John Wiley & Sons, New York.
4. Meyer, Paul L.-Introductory Probability and Statistical Reasoning. Addison-Wesley Publishing Company Inc.
5. Chernoff H. & Moses L.-Elementary Decision Theory. John Wiley & Sons, New York.
6. Panofsky Hans & Brier Glenn-Some Applications of Statistics to Meteorology. The Pennsylvania State University.
7. Nota técnica No 71 de la Organización Meteorológica Mundial.
8. Burchinal J. C & Dickerson W. H-Rainfall Probability and its applications. West Virginia University.
9. Nota técnica No 81 de la Organización Meteorológica Mundial.
10. Arkin Herbert & Colton Raymond-Tables for Statisticians. Barnes & Noble Inc. New York.
11. Spiegel Murray- Theory and Problems of Statistics. Schaum Publishing Co. New York.
12. Arias Jorge-Apuntes de Probabilidades y Estadística. Editorial A.E.I. Facultad de Ingeniería. Guatemala.

13. Byers Norace-General Meteorology. Mc Graw Hill Book Co. Inc.
14. Landsberg Helmut-Physical Climatology. Gray Printing Co. Inc. Pennsylvania.