

24
33

UNIVERSIDAD NACIONAL AUTONOMA DE
MÉXICO



Facultad de Ciencias

Algunas Funciones de Probabilidad de
Distancias Entre Puntos Distribuidos al Azar.

Tesis Profesional
Que para obtener el Título de
MATEMÁTICO

Presenta

Ernesto Maqueda Flores Roux

México, D.F

1989.

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

I. Introducción	1
II. Solución analítica	3
III. Solución numérica en el caso de dos dimensiones	6
IV. Análisis de la existencia de borde	11
V. Simplex vs. mínima χ^2	15
VI. Cambio de d a $d - 1$ dimensiones	17
VII. Conclusiones y resultados	22
VIII. Bibliografía	24
IX. Apéndice	25

I. INTRODUCCIÓN

Partamos del siguiente enunciado : Si un cierto número de personas se encuentra aleatoriamente distribuido en un recinto, y en un momento dado cada persona contamina a su vecino más cercano, ¿cuál es la probabilidad de no ser contaminado?

Podemos generalizar este problema de la siguiente manera : Considerense N personas dispuestas aleatoriamente en un espacio acotado en d dimensiones. En un cierto momento, cada persona contamina a su vecino más cercano con una probabilidad p . ¿Cuál es la proporción de personas que no son afectadas cuando N tiende a infinito, despreciando las correcciones de frontera?

Este problema, llamado del vecino maldito (*the vicious neighbour problem*) fue planteado por Abilock en 1967,⁽¹⁾ y resuelto para d y p generales por R. Tao y F. Y. Wu.⁽²⁾ La probabilidad buscada se expresa como un polinomio finito en p .

El enunciado de este problema sugiere un problema afín : Si las N personas están colocadas al azar siguiendo una distribución uniforme* en un espacio acotado en d dimensiones, ¿cómo se distribuirían las distancias entre cada persona y su vecino más cercano? Cabe hacer notar que existen personas que no son el vecino más cercano de nadie (y por lo tanto no son contaminadas con una probabilidad de uno), así como personas que son mutuamente vecinos más cercanos.

En las primeras secciones de este trabajo trataremos de atacar este último problema y sugerir una expresión analítica para la distribución de distancias entre vecinos más cercanos en espacios euclidianos en d dimensiones. Buscamos la solución de dos maneras totalmente distintas :

1. Considerando que el espacio no es acotado (es decir, no existe frontera), proponemos una demostración de la expresión analítica de la función de distribución.
2. Consideramos que el espacio es acotado, es decir, existe una frontera. Típicamente, tomaremos para dos dimensiones un cuadrado de área uno, y para tres dimensiones un cubo de volumen uno. Variando el número de puntos (la densidad), utilizando dos métodos de ajuste de curvas a datos experimentales, y empleando dos generadores de números al azar, compararemos las diferentes opciones, para encontrar cuál da el mejor resultado (lo cual es posible saberlo, dado que conocemos la respuesta al problema *a priori*); después hacemos uso de la técnica desarrollada para atacar un problema semejante del cual ignoramos la respuesta.

Todo el presente trabajo está sujeto de una manera u otra a la generación de números al azar. Podríamos definir una secuencia generada por computadora como aleatoria si el programa determinista que produce la secuencia es diferente y estadísticamente no correlacionado con el programa que utiliza la serie producida. Es decir, cualesquiera dos generadores de casuales deben producir estadísticamente los mismos resultados cuando

* Se dice que una variable aleatoria se distribuye uniformemente si su función de densidad está dada por $f_X(x) = \frac{1}{b-a} I_{(a,b)}(x)$. Tomaremos de ahora en adelante que $a = 0$ y $b = 1$, es decir, una distribución uniforme en el intervalo $(0,1)$.

son aplicados a un problema particular. Si no sucede así, entonces (al menos desde el punto de vista del usuario en particular), cuando menos uno de ellos no es un buen generador.⁽³⁾

II. SOLUCIÓN ANALÍTICA

Teorema : La función de densidad de la distribución* de las distancias entre vecinos más cercanos cuando los puntos se distribuyen uniformemente en un espacio en d dimensiones está dada por:

$$\omega(s) = d \left[\Gamma\left(\frac{d+1}{d}\right) \right]^d s^{d-1} e^{-\left[\Gamma\left(\frac{d+1}{d}\right)\right]^d s^d}$$

Demostración :

Sea p la probabilidad de encontrar un punto en un volumen de tamaño uno en un espacio de d dimensiones. Entonces p es proporcional a la densidad de puntos ρ . Tomemos un volumen de radio r y dividámoslo en n cáscaras de radio r_i y de ancho Δr_i . Este volumen está dado por $V_d = c_d r^d$, donde c_d es una constante que depende de la dimensionalidad.

La probabilidad de que no ocurra un punto en la i -ésima cáscara es

$$1 - p c_d r_i^{d-1} \Delta r_i.$$

* Esta distribución coincide con la propuesta por T. A. Brody en 1973 en su trabajo *A statistical measure for the repulsion of energy levels*,⁽⁴⁾ por lo que la llamaremos de ahora en adelante *Distribución de Brody*, como se hace en la literatura.⁽⁵⁾

Por lo tanto, la probabilidad de que el primer vecino al punto en el centro de la esfera en d dimensiones esté en la cáscara n -ésima estaría dada por :

$$\omega(r)\Delta r = \prod_{i=1}^{n-1} (1 - p c_{d-1} r_i^{d-1} \Delta r_i) p c_{d-1} r^{d-1} \Delta r$$

Ahora, si $n \gg 1$, entonces

$$\begin{aligned} \omega(r)\Delta r &= p c_{d-1} r^{d-1} e^{-\sum_{i=1}^{n-1} p c_{d-1} r_i^{d-1} \Delta r_i} \Delta r \\ &= p c_{d-1} r^{d-1} e^{-p c_d r^d} \Delta r \end{aligned}$$

Si denotamos por s a la distancia mínima, como tenemos que $p = (s/r)^d$, $\forall j$, entonces $s^j = p r^j$, $\forall j$, por lo que

$$\omega(s)\Delta s = A s^{d-1} e^{-c_d s^d} \Delta s$$

Para encontrar A y c_d , partimos de que se toma la media de la distribución (por definición) como uno, y normalizamos. Por lo tanto, se obtiene

$$\omega(s) = d \left[\Gamma\left(\frac{d+1}{d}\right) \right]^d s^{d-1} e^{-[r(\frac{s+1}{d})]^d}$$

que es el resultado buscado. ■

Propiedades de la distribución de Brody :

Los momentos de esta función de probabilidad (que están bien definidos sólo para $d > 0$) están dados por :

$$\mu_k' = \int_0^\infty s^k \omega(s) ds = \Gamma\left(\frac{d+k}{d}\right) / \left[\Gamma\left(\frac{d+1}{d}\right) \right]^k$$

Si $d = 1$, obtenemos lo siguiente : $\omega(s) = \Gamma(2)s^0 e^{-\Gamma(2)s} = e^{-s}$ que es un caso particular de la función de densidad exponencial* donde $\lambda = 1$.

Si $d = 2$, nos encontramos con la distribución de Wigner (que es la medida estadística comúnmente utilizada para representar la repulsión de niveles de energía⁽¹⁾):

$$\omega(s) = 2\left[\Gamma\left(\frac{3}{2}\right)\right]^2 s e^{-\Gamma(3/2)s^2} = \frac{\pi}{2} s e^{-\frac{\pi}{2}s^2}$$

Si $d \rightarrow \infty$, entonces $\mu'_k \rightarrow 1 \quad \forall k$, de donde $Var[X] = E[X^2] - (E[X])^2 = \mu'_2 - (\mu'_1)^2 = 0$, lo que significa que tenemos la distribución delta de Dirac[†]. Esto puede apreciarse en la Figura 1 (donde se grafica la densidad de la distribución de Brody para distintos valores del parámetro d), ya que al aumentar d , el valor máximo de la función aumenta, y tiende a localizarse en $x = 1$.

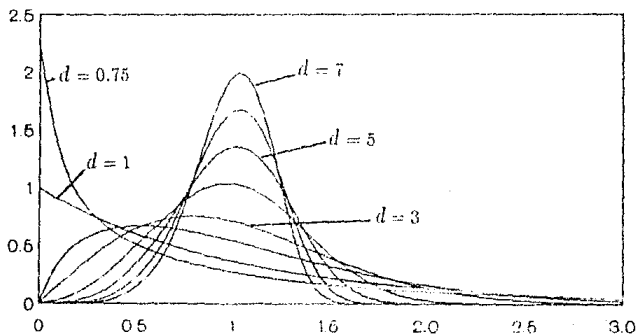


FIGURA 1 : Distribución de Brody

* Se dice que X tiene una distribución exponencial si su función de densidad está dada por $f_X(x) = \lambda e^{-\lambda x} I_{[0, \infty)}(x)$, con $\lambda > 0$. La esperanza de X está dada por $\frac{1}{\lambda}$ y la varianza por $\frac{1}{\lambda^2}$.

† La distribución delta de Dirac $\delta(t - t_0)$ queda definida por las siguientes propiedades: (1) $\delta(t - t_0) = 0$ si $t \neq t_0$, (2) $\int_{-\infty}^{\infty} \delta(t - t_0) dt = 1$, (3) la transformada de Laplace de $\delta(t - t_0)$ es $e^{-t_0 s}$. Para la distribución de Brody cuando $d \rightarrow \infty$, se tiene que $t_0 = 1$.

III. SOLUCIÓN NUMÉRICA EN EL CASO DE DOS DIMENSIONES

Preliminares

Para encontrar numéricamente la solución del problema, se utilizaron varias técnicas alternativas, de tal manera que se pudiera ver cuál resulta la más eficiente para atacar este tipo de problema :

1. **Generadores de números al azar :** Se emplearon dos generadores de casuales: el de la programoteca del IMSL ⁽⁷⁾ y el generador propuesto por Brody. ⁽⁶⁾ Se hace un análisis comparativo de los resultados que arroja cada uno de estos generadores.
2. **Ajuste de curvas :** Se comparan dos técnicas de ajuste de curvas a histogramas: El método *Simplex* ⁽⁸⁾ y el método de mínima χ^2 .
3. **Pruebas de hipótesis :** Se utiliza como prueba de bondad de ajuste la prueba χ^2 . Para comparar los dos métodos de ajuste de curvas se emplea una prueba paramétrica (la estadística *t* para diferencias apareadas) y una no paramétrica (la prueba del signo).

Método

Se tomó un cuadrado de lado uno, y utilizando ambos generadores de números al azar, se lanzaron N puntos al azar (donde cada coordenada se distribuye uniformemente). En un principio, se hicieron estadísticas para N igual a 10,000 , 25,000 , 50,000 y 100,000. Para poder hacer uso del teorema del límite central, ^(*) ya que no se conoce la distribución de los parámetros estimados, se hicieron cálculos para 30 cuadrados en cada caso. Se prueba la hipótesis nula $H_0 : d = 2$ contra la hipótesis alternativa

* El teorema del límite central afirma que si Y_1, Y_2, \dots, Y_n son variables aleatorias independientes idénticamente distribuidas con media μ y varianza $\sigma^2 < \infty$, entonces $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ converge a una función de distribución normal con media μ y varianza $\frac{\sigma^2}{n}$. Es decir, $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$. Si σ^2 es desconocida y se estima con $s^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$, entonces $t = \frac{\sqrt{n}(\bar{Y} - \mu)}{s}$ se distribuye como una *t* de Student con $n - 1$ grados de libertad. Para las aplicaciones prácticas del teorema del límite central, la aproximación es bastante buena si $n \geq 30$.

$H_a : d \neq 2$, calculando la estadística t . Los valores críticos de la estadística en cuestión son ± 2.045 para un nivel de significancia $\alpha = 5\%$; es decir, si la t calculada es mayor que 2.045 o menor que -2.045, se rechaza la hipótesis nula. Los resultados obtenidos se muestran a continuación en la *Tabla 1*.

Tabla 1

Número de puntos	10,000	25,000	50,000	100,000
<i>Generador de Brody</i>				
$d(\text{Simplex})$	1.999 725	2.002 019	1.999 155	2.002 008
desviación estándar	0.023 132	0.0190 99	0.022 538	0.009 242
$t_{\text{calculada}}$	-0.064 744	0.579 012	-0.205 133	1.190 236
$d(\chi^2)$	1.983 971	1.993 623	1.997 609	1.999 615
desviación estándar	0.017 942	0.013 135	0.009 550	0.006 661
$t_{\text{calculada}}$	-4.592 89	-2.658 71	-1.371 23	-0.316 10
<i>Generador IMSL</i>				
$d(\text{Simplex})$	1.986 183	1.993 300	1.991 697	1.996 866
desviación estándar	0.029 516	0.018 075	0.013 282	0.008 427
$t_{\text{calculada}}$	-2.563 85	-2.029 97	-3.423 77	-2.036 82
$d(\chi^2)$	1.975 320	1.987 138	1.991 232	1.996 866
desviación estándar	0.017 670	0.014 568	0.009 921	0.006 507
$t_{\text{calculada}}$	-7.619 73	-4.835 51	-4.840 30	-3.775 48

Interpretación de la Tabla 1

Basándonos en la tabla anterior, podemos hacer las siguientes observaciones y sacar algunas conclusiones :

- Con el generador de casuales de Brody, ajustando una curva con el método *Simplex*, nunca se rechaza la hipótesis nula.
- Con el generador de casuales de Brody, haciendo el ajuste minimizando la χ^2 asociada, se nota una clara disminución del valor de la estadística t conforme aumenta el número de puntos. Sólo se rechaza la hipótesis nula para la d obtenida para 10,000

y 25,000 puntos por cuadrado. Esto se debe al efecto de borde, lo cual analizamos más tarde.

- Con el generador IMSL y el método de ajuste de mínima χ^2 , es claro el rechazo de la hipótesis nula en todos los casos, dado que los cuatro valores t obtenidos son mayores que 3.

- Con el generador IMSL y el método de ajuste *Simplex*, no se rechaza la hipótesis nula para 25,000 y 100,000 puntos por cuadrado, pero el valor de la estadística t en ambos casos es tan cercano al valor crítico en cuestión, que cualquier afirmación sería peligrosa. Sin embargo, dadas las estadísticas t obtenidas para 10,000 y 50,000 puntos, y por la cercanía al valor crítico, nos inclinamos a pensar que si creciera el tamaño de la muestra (i.e., si n fuera mayor a 30), se rechazaría la hipótesis nula. Si este fuera el caso, rechazaríamos la hipótesis nula en todos los cálculos efectuados haciendo uso del generador de casuales de la programoteca del IMSL, a pesar de que sabemos que la hipótesis es cierta.

- Las Tablas A, B, C y D del Apéndice contienen los datos que fueron utilizados para construir la *Tabla 1*. Las primeras dos columnas muestran los valores de d obtenidos con el *Simplex* y su error respectivo; las últimas dos columnas muestran los valores de d obtenidos con el método de mínima χ^2 y el valor de ésta, habiendo partido de histogramas con 32 clases. El valor crítico de la χ^2 tomando un nivel de significancia $\alpha = 5\%$ y 30 grados de libertad (32 clases menos un parámetro estimado menos uno) es 43.7730. Con un asterisco están marcados los valores de la χ^2 mínima que exceden el valor crítico. Una simple inspección de las tablas permite ver que en un gran número de casos el valor de la mínima χ^2 es mayor que el valor crítico en cuestión, pero es notorio que invariablemente este número es mayor cuando se utiliza el generador IMSL. En la *Tabla 2* se resume esto.

Tabla 2 : Valores de χ^2 sobre el valor crítico

	Generador de Brody	Generador IMSL
10,000 puntos	10 de 30	15 de 30
25,000 puntos	15 de 30	16 de 30
50,000 puntos	12 de 30	14 de 30
100,000 puntos	7 de 30	27 de 30

Es claro que un generador es mejor si al tomar un número grande de números generados, no se presentan problemas inminentes; en este caso, con el generador IMSL, los resultados empeoran notablemente cuando tenemos 100,000 puntos por cuadrado, ya que 27 de los 30 valores de la mínima χ^2 exceden el valor crítico. Debería suceder (como efectivamente sucede con el generador de Brody) que al aumentar el número de puntos por cuadrado, el número de veces que el valor mínimo de la χ^2 excede al valor crítico disminuye.

- Estas dos últimas observaciones nos permiten concluir lo siguiente : el generador IMSL es menos bueno que el generador de Brody. Todos los cálculos basados en el generador IMSL, sobre todo si la serie empleada consta de varias decenas de miles, involucran de una u otra manera un error que aparece continuamente y que no parece existir si se utiliza el generador de Brody. Esto es un indicador de posibles correlaciones y ciclos en las series generadas por el IMSL. Como las consecuencias que pueden traer las correlaciones y ciclos pueden ser bastantes graves, ⁽⁶⁾ en el resto del presente trabajo emplearemos únicamente el generador de casuales de Brody.

Corrección de Fluctuaciones

Una buena manera de aminorar las fluctuaciones estadísticas que presentan todos los histogramas (lo cual puede observarse en la Figura 2, donde se grafica un histograma calculado y la curva ajustada) hasta ahora obtenidos es, siguiendo el espíritu de las matrices estocásticas, ⁽⁷⁾ superponer los resultados obtenidos al haber variado la semilla del generador. Para lograrlo, en cada uno de los cuatro casos (10,000 , 25,000 , 50,000

y 100,000 puntos), se hizo lo siguiente. Para cada intervalo de distancias, se sumaron las treinta barras que se obtuvieron (una por cada histograma), y se dividieron entre 30 (i.e., se promediaron). De esta manera, el histograma nuevo es mucho más suave que los treinta histogramas que le dieron origen, como se podrá ver en la *Tabla 3*.

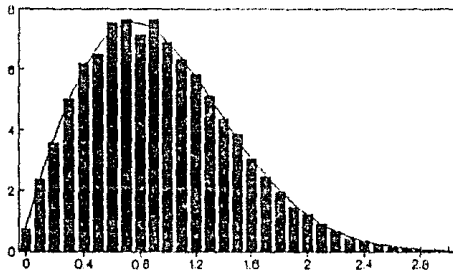


FIGURA 2 : Un Histograma Típico

TABLA 3 : Fluctuaciones Corregidas

Número de puntos	10,000	25,000	50,000	100,000
$d(\text{Simplex})$	1.999 513	2.001 930	2.002 201	2.001 976
error	256	1,802	2,043	11,981
χ^2 asociada	2.175	2.013	2.381	2.976
$d(\chi^2)$	2.001 715	2.002 476	2.002 998	2.001 795
χ^2	2.154	2.009	2.366	2.974
χ^2 si $\bar{d} = 2$	2.167	2.075	2.568	3.117

En todos los casos el valor de la χ^2 asociada a la d obtenida con el *Simplex* y a la d obtenida por el método de mínima χ^2 es totalmente admisible (el valor crítico es 43.7730), además de que nunca se rechaza la hipótesis nula $H_0 : d = 2$.

Conclusión

Los resultados numéricos permiten afirmar que el parámetro buscado, la d , es efectivamente 2, y por lo tanto, la función de distribución de la distancia entre cada punto y su vecino más cercano coincide con la distribución de Wigner.

IV. ANÁLISIS DE LA EXISTENCIA DE BORDE

El trabajar sobre un cuadrado de dimensiones fijas y totalmente acotado puede traer consigo la existencia de un cierto efecto de borde. Así por ejemplo, existe un cierto número de puntos (que llamaremos *puntos peligrosos*) cuyos vecinos más cercanos muy probablemente pueden estar fuera del cuadrado. Sin embargo, como sólo consideramos los puntos dentro del cuadrado, resulta que no tomamos la distancia mínima, sino la distancia mínima *dentro* del cuadrado. Esto perturba los cálculos para encontrar la d , por lo que tiene sentido ver qué tan grande es este efecto. Es necesario por lo tanto ver cómo es la relación de *puntos peligrosos* con respecto al número total de puntos, y ver cómo ésta cambia al variar N (número total de puntos). Entonces, buscaremos una relación entre N , N_p (número de *puntos peligrosos*) y la distancia mínima promedio entre los puntos (que llamaremos δ). Esta relación puede verse en la *Tabla 4*, y se gráfica en las *Figuras 3 y 4*.

Tabla 4 : Distancia Mínima Promedio

Puntos por cuadrado (N)	Distancia Mínima Promedio (δ)
2	0.520 841
3	0.389 119
4	0.322 230
5	0.280 976
6	0.248 982
7	0.226 914
8	0.210 399
9	0.195 397
10	0.182 799
15	0.147 031
20	0.124 827
50	0.074 971
100	0.052 111

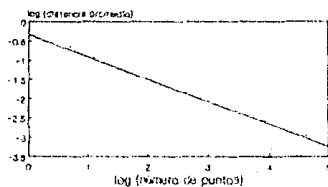


FIGURA 3 : Número de Puntos vs. Distancia Promedio

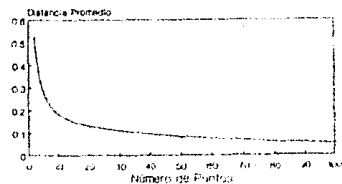


FIGURA 4 : Número de Puntos vs. Distancia Promedio

A la relación mostrada en la Tabla 4 podemos ajustarle una curva de la forma $\delta = N^{-\alpha}$, donde N es el número de puntos por cuadrado, y α el parámetro a ajustar. Por mínimos cuadrados* obtenemos que $\alpha = 0.70$ (con una desviación estándar de 0.016). Esto significa entonces que la distancia promedio disminuye exponencialmente conforme aumenta el número de puntos en el cuadrado.

Utilizaremos este resultado para encontrar la relación de puntos peligrosos con respecto al total de puntos en el cuadrado. Si dividimos cada lado del cuadrado en $n = 1/\delta$ partes, entonces la franja de ancho δ alrededor de la frontera consta de $4(n-1)$ cuadrados de área δ^2 . El número total de puntos en el cuadrado, N , es entonces igual a $\rho\delta^2n^2$, donde ρ es la densidad de puntos. El número de *puntos peligrosos* está dado por $4(n-1)\delta^2\rho$. La razón buscada, N_p/N , por lo tanto, es igual a

$$\begin{aligned}\frac{N_p}{N} &= \frac{16(n-1)\delta^2\rho}{4\rho\delta^2n^2} = \frac{4(n-1)}{n^2} = \frac{4}{n} - \frac{4}{n^2} \\ &= 4\delta - 4\delta^2 \approx 4N^{-0.7} - 4N^{-1.4}\end{aligned}$$

Entonces, N_p/N disminuye conforme aumenta N , es decir, la proporción de *puntos peligrosos* tiende a cero si el número de puntos por cuadrado tiende a infinito. Utilizando lo anterior podemos encontrar la proporción de *puntos peligrosos* para ocho casos, cinco de los cuales atacaremos numéricamente para encontrar cuál es el efecto que tiene la frontera sobre la d estimada anteriormente.

* Se ajusta una recta de la forma $\log(\delta) = -\alpha\log(N)$, de tal manera que se minimice la suma de los residuales al cuadrado.

Tabla 5 : N y la Proporción de Puntos Peligrosos

N	N_p/N	N_p
50	0.2420	12.1
100	0.1529	15.3
500	0.0509	25.5
1,000	0.0315	31.5
1,500	0.0238	35.7
10,000	0.0063	63.3
25,000	0.0033	82.5
50,000	0.0021	102.7
100,000	0.0013	126.5

En las Tablas E, F, G, H e I del Apéndice se encuentran los resultados obtenidos al hacer quince estadísticas promedio (para la corrección de fluctuaciones) con 100 histogramas cada una, para 50, 100, 500, 1,000 y 1,500 puntos por cuadrado. Se dan las estimaciones del parámetro d por el método *Simplex* y por mínima χ^2 , así como sus respectivos errores. Las últimas dos columnas (denominadas *límite inferior* y *límite superior*) forman el intervalo donde el parámetro d da un valor de χ^2 menor que el valor crítico. Aunque aceptaríamos la hipótesis nula siempre para $d = 2$, es claro que ambos métodos de ajuste invariablemente dan una d menor que 2. Además, conforme aumenta el número de puntos, el valor de d optimizado por ambos métodos se acerca más a 2. Por lo tanto, es posible afirmar que el efecto de frontera afecta el parámetro d haciéndolo valer menos que 2, como se muestra en la Tabla 6.

Esto puede entenderse de la siguiente manera. Cuando N_p/N es grande, el número de puntos para los cuales no se toma la distancia mínima no es despreciable. Entonces, en estos casos se toma como distancia mínima la del segundo vecino (y en algunos casos la del tercero, cuarto, quinto, etc.). Esto hace que el promedio de distancias aumente, lo cual es relevante ya que todas las distancias mínimas obtenidas se dividen entre el promedio para normalizar y lograr de esta manera una media de uno. Como el promedio de distancias es mayor que el que se tendría sin borde, entonces las distancias

que efectivamente corresponden a primeros vecinos se ven disminuidas con respecto al promedio total, por lo que la gráfica de la función de densidad se inclina más hacia la izquierda, que es lo que sucede cuando el parámetro d disminuye.

Tabla 6 : El efecto de frontera y la d óptima

N	$\bar{d}(\text{Simplex})$	<i>desv. std.</i>	$\bar{d}(\chi^2_{\min})$	<i>desv. std.</i>
50	1.968 330	0.035 588	1.899 108	0.027 606
100	1.971 533	0.026 708	1.913 753	0.020 983
500	1.980 254	0.007 479	1.923 189	0.008 225
1,000	1.981 647	0.006 788	1.974 322	0.007 478
1,500	1.982 675	0.005 714	1.976 850	0.006 005

V. SIMPLEX vs. MÍNIMA χ^2

A continuación se hacen dos pruebas estadísticas que comparan el método *Simplex* y el método de mínima χ^2 para el ajuste de curvas a histogramas dados para la obtención de parámetros poblacionales. La hipótesis nula H_0 es que ambos métodos dan estadísticamente el mismo resultado. Se utilizan dos pruebas (la prueba t de diferencias apareadas y la prueba del signo) sobre los datos de las Tablas A, B, C y D del Apéndice.

Prueba de diferencias apareadas*: En los cuatro casos en cuestión (10,000 , 25,000 , 50,000 y 100,000 puntos) calculamos $d(\text{Simplex}) - d(\chi^2)$. En la Tabla 7 puede apreciarse que siempre la t calculada es positiva (es decir, el promedio de las diferencias es positivo); para un nivel de significancia de $\alpha = 5\%$, los valores críticos para una t con 29 grados de libertad (30 elementos de la muestra menos uno) haciendo

* La prueba de diferencias apareadas consiste en comparar los métodos y el número de los (x_i, y_i) cuando las observaciones en ambas aparecen por pares. Tomaremos como hipótesis nula $H_0: \mu_1 - \mu_2 = 0$ y como hipótesis alternativa $H_a: \mu_1 - \mu_2 \neq 0$. Como las observaciones están apareadas, pueden escribirse como $(X_i, Y_i), \forall i = 1, \dots, n$, donde X_i es una observación de la población 1 y Y_i es una observación de la población 2. Se calcula $d_i = X_i - Y_i, \bar{d} = \frac{\sum_{i=1}^n d_i}{n}$, y su variancia muestral, s_d^2 . Entonces, para n suficientemente grande $t = \frac{\bar{d} - (\mu_1 - \mu_2)}{s_d / \sqrt{n}}$ se distribuye como una t de Student con $n - 1$ grados de libertad.

una prueba de dos colas son ± 2.045 , por lo que en 3 de 4 casos se rechaza la hipótesis nula H_0 : las dos pruebas son iguales.

Prueba del signo: Para un nivel de significancia de $\alpha = 5\%$ y una prueba de dos colas, los valores críticos son ± 1.96 . En los cuatro casos, el valor de la Z calculada es positivo y mayor que el valor crítico, por lo que esta prueba no paramétrica nos indica que el valor del parámetro estimado empleando el método *Simplex* es mayor que empleando el método de mínima χ^2 .

Tabla 7 : Simplex vs. Mínima χ^2

Número de puntos	10,000	25,000	50,000	100,000
<i>Prueba de Diferencias Apareadas</i>				
$d(S) - d(\chi^2)$	0.015 753	0.008 335	0.001 516	0.002 393
desviación estándar	0.012 100	0.009 259	0.018 219	0.004 022
$t_{calculada}$	5.042	3.511	0.329	2.304
<i>Prueba del Signo</i>				
Signos Positivos	28	25	24	22
Signos Negativos	2	5	6	8
$Z_{calculada}$	4.747	3.651	3.286	2.556

Conclusión : La estimación de un parámetro utilizando el método *Simplex* da en general valores por arriba de la estimación del parámetro usando el método de mínima χ^2 . Como consecuencia, si se utiliza como prueba de bondad de ajuste la χ^2 , la probabilidad de rechazar la hipótesis nula planteada aumenta si se utiliza la estimación del método *Simplex*. A lo largo del presente trabajo se seguirán presentando las estimaciones de parámetros por ambos métodos.

* La prueba del signo, al igual que la prueba de diferencias apareadas, consiste en comparar observaciones que aparecen por pares, es decir, pueden escribirse de la forma $(X_i, Y_i), Y_i = 1, \dots, n$. Omitimos todas las observaciones donde $X_i = Y_i$. Sea $p = P(X_i - Y_i > 0)$ y sea $q = 1 - p = P(X_i - Y_i < 0)$. Si nos interesa ver si las poblaciones son iguales, entonces la hipótesis nula será $H_0: p = 0.5$ y la alternativa $H_a: p \neq 0.5$. Si J es el número de veces que $X_i - Y_i > 0$, entonces $Z = \frac{J - n/2}{\sqrt{n/4}}$ se distribuye como una normal con media cero y varianza uno, si tenemos diez o más observaciones, por lo que dado el valor de Z y habiendo escogido un nivel de significancia, puede o no rechazarse H_0 .

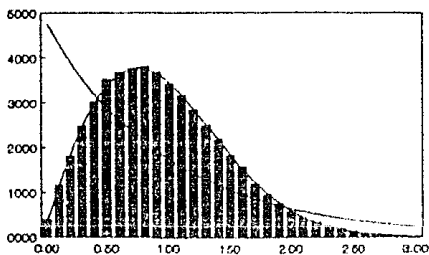
VI. CAMBIO DE d A $d-1$ DIMENSIONES

En esta parte aplicaremos los resultados obtenidos en las secciones anteriores para atacar un problema del cual no conocemos la respuesta. Sabemos ya que la distribución de Brody representa la distribución de distancias entre vecinos más cercanos en un espacio euclidiano de d dimensiones. Esto es válido para d igual a un entero positivo. Además sabemos que valores cercanos a d representan la distribución buscada cuando existen correcciones de frontera. Nos preguntamos: ¿Qué representa la distribución de Brody cuando d no es un entero ni un valor que represente alguna corrección de frontera? ¿Es posible que represente lo mismo cuando el espacio no es de dimensión entera (*un fractal*)? No sabiendo si la distribución de Brody será útil, abordaremos el siguiente problema.

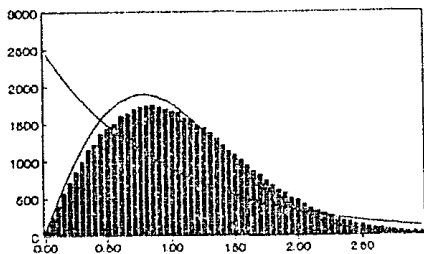
¿Qué sucede si en vez de tomar un cuadrado de lado uno, tomamos un rectángulo con un lado fijo (la base, de longitud uno) y el otro (la altura) lo vamos disminuyendo de manera que su longitud tienda a cero? Si la densidad de puntos es lo suficientemente grande, sabemos que en el cuadrado la distribución de Wigner (i.e., la distribución de Brody cuando $d = 2$) es la respuesta al problema planteado; si la altura es cero, entonces sabemos que la distribución exponencial con $\lambda = 1$ representa el fenómeno estudiado (la distribución de Brody cuando $d = 1$). ¿Qué sucede en los pasos intermedios? ¿La distribución utilizada con $d \in [1, 2]$ es una buena aproximación a lo observado experimentalmente?

Comenzaremos por hacer un estudio gráfico del problema. A continuación se presentan los histogramas obtenidos haciendo variar la altura del rectángulo.

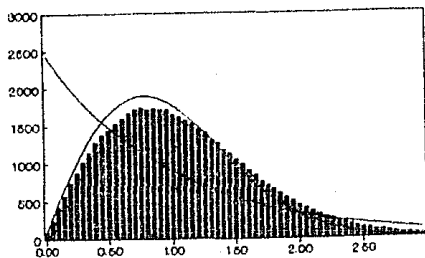
1 por 0.1



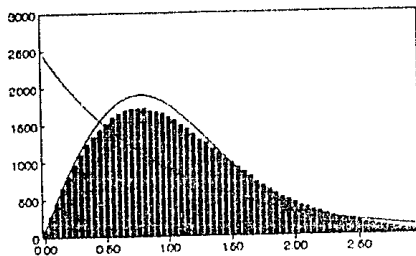
1 por 0.01



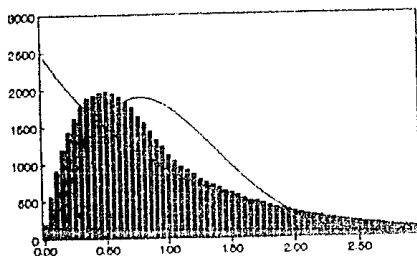
1 por 0.001



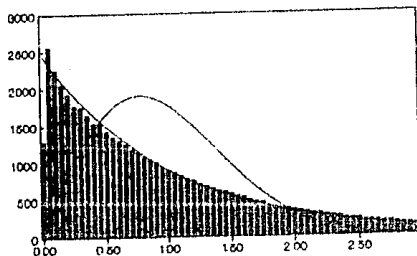
1 por 0.0001



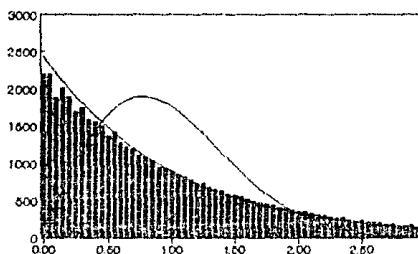
1 por 0.000 01



1 por 0.000 001



1 por 0.000 000 1



En todas las figuras, se grafican la función exponencial con $\lambda = 1$ y la distribución de Wigner (que corresponden a la función de Brody cuando $d = 1$ y cuando $d = 2$) sobre el histograma obtenido numéricamente de las distancias mínimas en un rectángulo de las dimensiones especificadas. Como podrá claramente observarse, cuando la altura del rectángulo es mayor que 0,1, no cambia la distribución obtenida para el cuadrado de lado uno. Sin embargo, al ir disminuyendo la altura, la cola de la distribución va aumentando. Además, el punto donde se alcanza el máximo se recorre hacia la izquierda. Como por construcción la media es uno (ya que se dividió entre el promedio para normalizar), esto significa que se da un aumento gradual de la varianza.

En la *Tabla 8* se muestra cual es la distancia mínima promedio según la altura. El mínimo se alcanza cuando estamos en el caso de la línea recta.

Tabla 8 : Valores promedio de la distancia mínima

Altura	Distancia Media
0.01	0.000 225 7
0.001	0.000 072 8
0.000 1	0.000 024 4
0.000 01	0.000 011 2
0.000 001	0.000 010 3
0.000 000 1	0.000 010 0

Si comparamos las distancias mínimas promedio y las gráficas, podemos observar que conforme esta distancia promedio es menor que la altura del rectángulo, la curva sigue manteniendo una forma parecida a la distribución de Wigner. En el rectángulo de 1 por 0.000 01, la distancia promedio es mayor que la altura, y es en este caso cuando la parte asintótica de la curva se asemeja mucho a la distribución exponencial. El único punto en donde hay un poco de problema es en el intervalo alrededor del cero. La probabilidad de que la distancia entre un punto y su vecino más cercano sea cero sigue siendo cero (i.e., $\lim_{x \rightarrow 0} P(X < x) = 0$). Esto deja de ser cierto cuando tomamos un rectángulo de altura 0.000 0001. Sin embargo, en este último caso, dada la precisión numérica de la computadora, y comparando los resultados que se obtienen cuando tomamos la línea recta, la perturbación causada por la altura es despreciable y nos encontramos en el caso de la distribución exponencial.

Para ningún caso es posible el hacer el ajuste de la distribución de Brody para algún valor de d . Es decir, el cambio de dos a una dimensión es discontinuo en d .

Sigue entonces quedando una pregunta abierta. ¿Qué representa entonces la distribución de Brody cuando se toma una d no entera?

VII. CONCLUSIONES Y RESULTADOS

1. La distribución de Brody es la función de probabilidad que representa la distancia entre un punto y su vecino más cercano si éstos se distribuyen uniformemente y el espacio es no acotado y sin frontera. La función depende únicamente de un parámetro: la dimensionalidad del espacio (d).
2. Si el espacio es acotado, la distribución de Brody sigue siendo efectiva para representar la distribución de vecinos más cercanos en espacios euclidianos de d dimensiones. Sin embargo, el parámetro en cuestión es ligeramente menor que la dimensionalidad del espacio en el que se está trabajando. Es claro que el efecto de borde disminuye (de manera exponencial) con respecto al número de puntos por cuadrado, es decir, el número de *puntos peligrosos* es inversamente proporcional a la densidad de puntos.
3. Los resultados del generador de números al azar de la programación del IMSL son significativamente menos buenos que los resultados obtenidos utilizando el generador de Brody; conforme aumenta el número de casuales utilizado, los resultados del IMSL empeoran, siendo que los resultados obtenidos con el generador de Brody se mantienen relativamente constantes. Es decir, desde el punto de vista del problema tratado en este trabajo, el generador IMSL no produce el resultado esperado, siendo que el generador de Brody sí; esto es, el generador IMSL no cumple con la definición de un buen generador de números al azar, como lo expusimos en un principio.
4. Una buena manera de aminorar las fluctuaciones estadísticas cuando se tiene un problema en que no existe el problema de muestreo es, siguiendo el espíritu de las matrices estocásticas, promediar los histogramas; esto es posible gracias a que sólo se depende de un generador de casuales y de tiempo de procesador de computadora. Los resultados así obtenidos son radicalmente mejores que los resultados obtenidos de otras maneras. Los valores de la prueba χ^2 de bondad

de ajuste son particularmente buenos cuando se hace esta corrección (llegando al extremo de obtener valores menores que 2 para una χ^2 con 31 grados de libertad).

5. Si se utiliza como método de ajuste de un parámetro el *Simplex*, entonces es necesario tener cuidado con la prueba de bondad de ajuste que se aplique. En la mayoría de los casos, la estimación por *Simplex* sobrestima el valor obtenido por el método de mínima χ^2 , por lo que si su valor mínimo está muy cerca del valor crítico, el valor obtenido con el *Simplex* en cuestión será rechazado ante la prueba χ^2 de bondad de ajuste.
6. Pueden hacerse pruebas de hipótesis para una infinidad de valores de d encontrando el intervalo para d donde el valor de la χ^2 es menor que el valor crítico. Esto es posible gracias a la continuidad de la función de Brody al variar la dimensionalidad y a la continuidad de la función χ^2 .
7. Al hacer el cambio de dos dimensiones a una dimensión haciendo tender a cero el lado de un rectángulo y manteniendo el otro constante (e igual a uno), se observa que mientras la distancia mínima promedio sea menor que el lado variable, se obtiene la distribución de Wigner para alturas mayores que 0.1, y distribuciones muy parecidas a la de Wigner con una mayor varianza para alturas menores. Cuando la distancia mínima promedio está en la vecindad de la altura del rectángulo, entonces el máximo de la función se recorre hacia la derecha, pero la probabilidad de que la distancia sea cero sigue siendo cero (i.e. se mantiene el fenómeno de repulsión).⁽⁴⁾ Existe un punto en que los cálculos dejan de ser afectados por la altura del rectángulo y nos encontramos en presencia de una línea recta y la distribución exponencial.
8. Sigue quedando abierto el significado del parámetro d cuando no es un entero. ¿Representará la distribución de distancias entre vecinos más cercanos en un espacio de dimensión no entera?

VIII. BIBLIOGRAFÍA

1. R. Abilock, *American Mathematical Monthly*, 1967, vol.74, p.720.
2. R. Tao & F. Y. Wu, *The vicious neighbour problem*, *Journal of Physics A*, 1987, L299-L306.
3. William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, *Numerical Recipes in C. The Art of Scientific Computing*, Cambridge University Press, 1988, p.204.
4. T. A. Brody, *A statistical measure for the repulsion of energy levels*, *Lettere al Nuovo Cimento*, 1973, p.482.
5. C. E. Porter, editor, *Statistical Theories of Spectra: Fluctuations*, Academic Press, 1965.
6. T. A. Brody, *A random-number generator*, *Computer Physics Communications*, 1984, vol. 34, p.39-46.
7. IMSL, International Mathematics and Statistics Library, Sixth Floor, NBC Building, 7500 Bellaire Boulevard, Houston, Texas, 77036.
8. J. A. Nelder & R. Mead, *Computer Journal*, vol. 7, 1965, p. 308.

IX. APENDICE

TABLA A : 10,000 puntos
 Generador de Brody
 Histogramas con 32 clases

$d(\text{Simple})$	Error	$d(\chi^2)$	χ^2
2.051 887	8,273	2.014 393	30.41
1.971 658	13,906	1.979 993	43.25
2.008 726	9,210	1.980 262	30.40
1.988 962	10,194	1.979 744	41.69
2.006 205	8,899	1.977 453	50.27*
2.030 753	9,851	2.001 158	36.34
2.009 100	12,143	1.973 011	47.28*
2.041 458	14,192	2.021 177	30.80
2.006 771	15,193	1.986 796	48.67*
1.978 296	11,429	1.976 192	32.71
1.976 962	21,222	1.958 157	61.56*
2.014 143	12,935	1.997 530	38.94
1.992 244	7,044	1.991 738	27.22
1.992 266	10,198	1.984 448	28.53
1.993 212	16,341	1.990 374	55.69*
1.941 116	15,498	1.940 134	42.35
1.997 865	10,137	1.978 592	42.26
2.005 179	20,619	1.987 187	68.76*
1.999 589	21,797	1.995 386	40.77
2.019 399	13,386	1.989 017	35.49
1.992 085	13,384	1.973 606	40.27
2.022 705	20,241	2.011 647	54.86*
1.966 924	16,075	1.965 114	42.11
1.995 277	17,516	1.966 387	51.71*
1.973 856	11,905	1.957 497	32.84
1.996 740	15,889	1.967 091	52.98*
2.032 428	10,749	2.011 187	26.99
1.997 499	18,796	1.987 251	51.97*
2.006 594	8,880	1.994 151	20.63
1.981 851	11,274	1.982 560	31.76

TABLA A : 10,000 puntos

Generador IMSL

Histogramas con 32 clases

$d(\text{Simpl}x)$	<i>Error</i>	$d(\chi^2)$	χ^2
1.959 280	19,054	1.975 545	42.74
1.992 266	17,777	1.967 763	54.57*
1.985 904	8,127	1.980 409	29.13
1.976 672	15,970	1.975 968	41.71
1.957 195	18,621	1.959 315	47.04*
1.924 485	26,665	1.942 585	68.75*
2.003 364	8,211	1.982 566	28.05
1.991 288	12,769	1.967 251	33.22
1.916 215	14,884	1.953 644	41.93
1.933 974	18,002	1.948 051	50.19*
1.997 310	12,296	1.981 907	31.97
1.974 845	6,833	1.985 344	26.60
1.971 888	17,257	1.965 567	53.88*
1.977 863	22,896	1.965 881	47.22*
2.012 146	19,070	1.962 906	54.63*
2.032 662	21,056	1.989 017	62.26*
1.987 368	10,016	1.974 118	26.26
1.981 626	13,842	1.968 831	53.91*
1.974 697	13,857	1.973 005	46.44*
2.023 000	23,608	1.984 435	65.67*
2.015 141	10,949	1.983 500	37.06
2.028 320	9,420	2.011 033	25.50
2.014 837	13,026	1.967 309	58.99*
2.014 655	15,045	2.003 136	45.31*
1.986 616	15,147	1.960 698	47.00*
2.005 481	4,950	1.986 169	37.77
1.964 254	5,817	1.947 532	28.75
2.022 083	10,684	2.018 425	32.14
1.964 496	25,393	1.990 047	62.88*
1.995 073	13,323	1.987 660	38.01

TABLA B : 25,000 puntos

Generador de Brody

Histogramas con 32 clases

$d(\text{Smpl}x)$	<i>Error</i>	$d(\chi^2)$	χ^2
2.001 352	53,390	1.996 198	52.23*
2.001 076	35,098	1.999 194	37.08
2.031 372	50,150	2.005 286	53.73*
2.029 397	67,028	1.997 383	70.69*
2.026 214	65,446	2.002 899	66.55*
2.006 390	26,069	1.987 475	43.06
1.997 356	28,854	1.992 608	39.03
1.990 115	40,791	1.979 897	43.23
1.963 259	23,635	1.971 296	29.56
1.992 027	44,727	1.982 605	63.44*
1.999 722	27,715	1.985 030	35.84
2.036 768	35,205	2.019 929	40.38
2.001 963	25,629	1.995 095	45.07*
1.964 184	24,050	1.969 772	34.04
2.006 534	25,081	1.992 679	40.62
1.991 004	36,752	1.983 368	41.08
1.991 324	35,006	1.991 341	54.47*
1.997 125	47,822	1.991 289	53.66*
2.010 089	30,926	2.003 001	43.45
2.004 933	35,743	2.000 134	49.66*
1.998 569	33,378	1.995 738	37.91
2.043 728	45,232	2.026 988	49.55*
1.988 285	6,890	1.986 745	20.65
1.995 744	31,075	1.994 176	29.99
2.016 778	44,480	2.011 756	44.98*
1.966 235	63,653	1.971 328	45.58*
2.000 716	28,153	1.989 683	46.43*
2.007 071	44,821	2.007 488	43.64
1.994 530	50,702	1.983 750	61.89*
2.007 877	42,293	1.994 643	53.58*

TABLA B : 25,000 puntos

Generador IMSL

Histogramas con 32 clases

$d(\text{Simplex})$	Error	$d(\chi^2)$	χ^2
2.014 553	62,093	2.012 601	46.24*
1.983 565	18,320	1.972 012	22.46
1.995 272	37,249	1.988 403	42.05
1.976 230	46,847	1.973 907	52.91*
1.988 443	44,493	1.976 397	50.01*
1.976 515	37,391	1.965 440	52.11*
2.041 401	21,245	2.025 484	25.79
1.995 391	25,824	1.993 037	42.91
1.947 762	23,869	1.957 715	36.96
2.016 937	23,402	2.008 025	32.98
1.980 900	34,341	1.975 449	41.29
2.013 114	23,794	1.990 208	51.19*
1.973 094	39,736	1.971 065	40.26
2.002 527	42,603	1.997 012	45.98*
1.989 256	26,054	1.980 704	30.86
2.010 843	55,079	1.996 589	47.79*
2.001 210	62,594	1.975 270	68.99*
1.987 730	47,958	1.988 665	44.42*
2.013 574	48,922	1.979 411	56.08*
1.990 829	48,432	1.994 592	47.75*
2.008 551	57,867	2.003 515	48.58*
1.973 231	39,588	1.973 318	65.63*
1.981 525	36,638	1.987 200	56.05*
2.002 584	26,675	2.002 701	36.49
1.991 479	38,362	1.991 962	42.24
1.983 801	58,356	1.983 027	45.55*
1.991 444	33,813	1.986 150	50.51*
1.997 034	25,830	1.982 431	34.06
1.999 215	32,508	1.994 272	30.90
1.971 017	31,214	1.987 558	39.69

TABLA C : 50,000 puntos

Generador de Brody

Histogramas con 32 clases

$d(\text{Simple})$	Error	$d(\chi^2)$	χ^2
2.015 542	35,923	2.006 381	34.41
2.026 897	59,622	2.020 908	30.42
2.007 346	61,421	2.000 698	39.52
2.017 146	47,266	2.008 985	37.16
2.021 865	52,972	2.014 477	38.35
1.994 568	73,054	1.991 655	43.95*
1.990 509	38,004	1.991 303	37.13
1.995 696	72,438	1.984 800	45.92†
1.982 740	39,059	1.988 914	26.47
2.014 656	108,710	2.002 483	48.97*
2.019 913	45,635	2.006 995	39.70
2.013 989	45,364	2.007 878	24.01
1.995 207	86,834	1.989 856	39.85
2.003 524	40,834	1.995 885	30.42
1.991 045	35,923	1.990 528	26.63
2.001 992	53,187	1.996 685	31.67
1.997 351	87,111	1.991 302	52.05*
2.015 472	122,462	2.005 497	57.02*
2.011 276	67,581	2.005 913	51.05*
1.990 199	65,882	1.991 821	35.10
1.997 116	80,159	1.999 143	37.65
2.014 850	64,134	2.007 737	35.48
1.983 742	96,652	1.986 310	50.68*
1.995 816	85,131	1.990 106	53.68*
1.999 774	91,826	1.998 541	47.30*
1.990 446	39,042	1.989 247	23.69
1.992 025	75,526	1.986 464	55.98*
1.990 458	102,138	1.989 229	57.29*
1.981 150	92,738	1.984 921	53.44*
2.021 331	49,262	2.003 609	39.55

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

TABLA C : 50,000 puntos

Generador IMSL

Histogramas con 32 clases

$d(\text{Simpl}x)$	Error	$d(\chi^2)$	χ^2
1.992 051	47,436	1.984 217	40.69
1.987 890	51,460	1.991 290	32.04
2.001 472	35,116	2.001 382	30.40
1.965 248	66,964	1.980 601	46.07*
1.992 477	96,153	1.999 085	50.50*
1.978 836	88,465	1.978 950	46.40*
2.021 905	85,336	2.011 577	41.26
2.001 535	74,390	1.999 156	40.37
1.969 930	76,172	1.972 640	38.71
1.982 053	91,004	1.984 473	61.45*
2.018 830	64,345	2.015 210	38.43
2.001 708	115,456	1.987 027	61.60*
1.982 055	44,935	1.982 393	30.03
1.980 533	50,150	1.984 633	50.82*
1.992 400	58,056	1.991 962	39.28
2.006 100	59,446	1.996 499	37.13
1.998 661	162,099	1.989 676	73.97*
1.996 372	55,463	1.984 115	37.09
1.988 479	55,975	1.978 515	46.26*
1.981 276	66,860	1.989 760	37.22
1.995 295	185,193	1.994 548	97.44*
1.975 861	68,921	1.986 157	47.86*
1.977 561	116,695	1.983 974	47.70*
2.009 309	92,206	2.010 918	48.26*
1.987 123	64,427	1.988 595	41.31
1.996 006	150,365	1.993 651	64.39*
2.001 547	46,735	1.997 799	29.93
1.988 280	77,420	1.992 672	52.15*
1.980 388	70,728	1.994 912	39.54
1.999 704	50,245	1.990 566	39.03

TABLA D : 100,000 puntos

Generador de Brody

Histogramas con 32 clases

$d(\text{Simplx})$	Error	$d(\chi^2)$	χ^2
2.009 272	96,818	2.005 849	29.97
2.014 079	109,879	2.008 697	43.35
2.006 457	111,913	2.000 294	28.25
2.005 082	91,628	2.000 615	24.28
2.006 210	75,662	2.001 018	23.29
1.998 656	157,481	1.993 818	43.68
1.997 128	111,517	1.995 559	41.35
1.996 995	102,869	2.001 331	29.38
1.997 166	79,946	2.000 873	25.86
2.014 143	111,835	2.009 446	29.95
2.005 742	188,940	2.006 937	47.54*
2.003 708	138,889	2.001 843	36.65
2.006 340	138,780	2.004 268	37.49
2.026 971	112,177	2.013 081	46.44*
2.000 260	235,259	1.997 248	49.59*
1.993 284	154,270	1.996 813	50.25*
2.001 443	75,843	1.993 869	21.71
2.011 238	201,789	2.004 665	54.28*
1.990 746	86,542	1.989 100	27.23
1.992 686	146,234	1.994 535	38.38
1.992 044	165,044	1.993 075	51.14*
2.007 007	100,021	2.003 194	36.43
1.995 337	125,773	1.996 269	37.23
2.011 147	74,859	2.007 967	22.47
2.005 169	178,656	2.003 865	41.75
1.977 709	126,828	1.983 193	34.86
1.996 973	97,420	1.994 208	35.19
1.995 412	131,714	1.992 768	45.44*
1.996 446	82,874	1.993 338	30.58
2.005 406	81,112	2.000 730	30.07

TABLA D : 100,000 puntos

Generador IMSL

Histogramas con 32 clases

$d(\text{Smpl}x)$	<i>Error</i>	$d(\chi^2)$	χ^2
2.006 093	543,952	2.008 069	100.22*
1.990 126	520,058	1.993 664	100.44*
1.998 479	149,051	1.991 955	39.95
2.000 000	294,724	1.997 434	70.64*
1.995 207	354,289	1.996 053	80.59*
1.991 152	192,167	1.988 077	47.89*
2.003 471	191,884	2.001 408	47.47*
2 002 361	171,553	1.994 868	58.78*
1.986 258	192,117	1.990 042	52.15*
1.979 454	318,618	1.981 932	56.98*
2.009 673	482,347	2.003 718	104.67*
1.993 835	334,087	1.994 496	68.60*
1.995 790	279,234	1.992 365	59.11*
2.003 628	215,153	2.002 029	56.30*
1.992 879	187,063	1.997 005	36.64
1.993 803	220,250	1.995 239	42.82
1.997 483	238,693	1.994 823	48.13*
1.989 269	168,671	1.990 420	49.04*
2.008 282	293,207	2.010 393	72.57*
1.981 665	232,747	1.986 047	57.19*
1.993 423	315,188	1.994 535	66.50*
2.014 492	173,541	1.999 994	63.81*
1.984 949	179,871	1 993 312	58.55 ¹
2.009 543	119,611	2.006 278	46.20*
2.000 815	137,701	1.994 727	47.73*
1.996 296	237,663	1.993 498	47.25*
2.007 411	241,931	2.000 947	58.17*
1.991 248	623,223	1.986 995	140.13*
1.997 691	433,295	1.995 610	86.93*
1.995 690	309,246	1.989 459	59.89*

TABLA E : 50 puntos
Histogramas promedio con 6 clases

$d(\text{Smpl}x)$	<i>Error</i>	$d(\chi^2)$	χ^2_{\min}	<i>lim.inf.</i>	<i>lim.sup.</i>
1.956 637	1	1.869 577	0.18	1.187 820	2.535 444
2.023 119	2	1.944 850	0.49	1.212 328	2.573 754
1.951 342	2	1.893 637	0.22	1.181 119	2.534 943
2.054 290	2	1.966 577	0.54	1.245 857	2.580 479
1.951 342	2	1.893 637	0.22	1.181 119	2.534 943
1.951 342	2	1.893 637	0.22	1.181 119	2.534 943
1.962 201	1	1.868 353	0.22	1.177 666	2.545 693
2.007 997	3	1.918 925	0.37	1.217 809	2.542 629
1.899 506	3	1.868 353	0.20	1.148 834	2.524 462
1.956 637	1	1.866 577	0.18	1.187 820	2.535 444
1.989 011	3	1.919 831	0.35	1.205 712	2.551 936
1.956 911	2	1.896 755	0.26	1.170 635	2.545 063
1.951 342	2	1.893 637	0.22	1.181 119	2.534 943
1.956 637	1	1.896 137	0.18	1.187 820	2.535 444
1.956 637	1	1.896 137	0.18	1.187 820	2.535 444

TABLA F : 100 puntos
Histogramas promedio con 6 clases

$d(\text{Smpl}x)$	<i>Error</i>	$d(\chi^2)$	χ^2_{\min}	<i>lim.inf.</i>	<i>lim.sup.</i>
1.930 074	9	1.928 931	0.23	1.381 641	2.438 616
2.002 174	10	1.920 453	0.35	1.435 316	2.473 088
1.951 342	9	1.893 637	0.45	1.377 751	2.370 516
2.017 745	11	1.930 716	0.43	1.452 827	2.477 397
1.947 622	6	1.941 656	0.15	1.399 166	2.445 940
2.011 465	5	1.919 036	0.17	1.431 161	2.483 936
1.989 356	6	1.918 473	0.58	1.403 988	2.388 151
1.975 224	9	1.916 231	0.59	1.400 346	2.386 709
1.954 519	9	1.917 996	0.45	1.377 751	2.370 516
1.974 369	7	1.917 344	0.59	1.403 459	2.386 822
1.985 491	7	1.869 036	1.04	1.379 155	2.315 910
1.922 790	11	1.909 750	0.77	1.343 827	2.291 184
1.978 308	7	1.918 473	0.52	1.394 786	2.393 952
1.962 269	8	1.937 423	0.30	1.405 342	2.460 917
1.970 253	10	1.867 139	1.10	1.384 459	2.309 212

TABLA G : 500 puntos
Histogramas promedio con 6 clases

$d(\text{Smplr})$	<i>Error</i>	$d(\chi^2)$	χ^2_{crit}	<i>lim.inf.</i>	<i>lim.sup.</i>
1.977 169	201	1.913 831	2.69	1.705 396	2.114 664
1.971 343	230	1.933 331	2.55	1.721 296	2.137 165
1.980 252	172	1.926 899	2.21	1.709 663	2.135 734
1.974 469	174	1.922 088	2.05	1.703 052	2.132 723
1.978 791	193	1.916 412	2.67	1.706 777	2.118 246
1.968 421	186	1.919 749	2.04	1.700 113	2.130 784
1.977 728	191	1.914 329	2.53	1.703 294	2.117 363
2.000 550	153	1.938 195	2.63	1.726 960	2.141 029
1.976 379	207	1.938 433	2.68	1.726 998	2.141 467
1.981 004	208	1.916 669	2.60	1.706 434	2.119 102
1.990 164	191	1.931 243	2.40	1.716 807	2.137 477
1.978 227	172	1.914 483	2.45	1.702 648	2.118 317
1.984 996	176	1.919 373	2.57	1.708 338	2.122 206
1.980 092	209	1.923 788	2.38	1.709 353	2.130 223
1.984 231	184	1.919 006	2.69	1.709 771	2.126 039

TABLA H : 1000 puntos
Histogramas promedio con 12 clases

$d(\text{Smplr})$	<i>Error</i>	$d(\chi^2)$	χ^2_{crit}	<i>lim.inf.</i>	<i>lim.sup.</i>
1.977 599	32	1.972 331	0.42	1.749 494	2.181 166
1.979 333	23	1.977 481	0.19	1.752 444	2.188 916
1.985 468	17	1.979 851	0.18	1.754 814	2.190 686
1.986 110	17	1.974 439	0.24	1.751 602	2.183 074
1.991 734	14	1.979 554	0.28	1.756 517	2.187 988
1.990 250	17	1.978 913	0.29	1.756 076	2.187 348
1.984 391	14	1.978 739	0.15	1.754 101	2.189 574
1.975 792	15	1.978 225	0.27	1.752 787	2.189 860
1.970 898	13	1.960 403	0.36	1.738 767	2.168 438
1.967 452	11	1.960 915	0.29	1.738 878	2.169 149
1.984 112	13	1.970 962	0.46	1.750 125	2.177 796
1.981 228	22	1.961 104	0.67	1.742 668	2.165 538
1.981 852	11	1.977 915	0.16	1.752 678	2.189 350
1.990 587	16	1.985 220	0.29	1.760 383	2.195 855
1.977 904	23	1.978 782	0.35	1.753 944	2.189 817

TABLA I : 1500 puntos
 Histogramas promedio con 12 clases

$d(\text{Smpl}x)$	<i>Error</i>	$d(\chi^2)$	χ_{\min}^2	<i>lim.inf.</i>	<i>lim.sup.</i>
1.984 913	33	1.972 811	0.46	1.792 181	2.143 839
1.977 977	31	1.975 400	0.21	1.792 370	2.149 229
1.981 197	28	1.988 002	0.18	1.723 758	2.257 247
1.987 117	33	1.973 949	0.51	1.793 719	2.144 777
1.990 356	31	1.981 871	0.51	1.800 441	2.153 900
1.985 132	43	1.973 679	0.53	1.793 249	2.144 708
1.974 611	13	1.973 817	0.12	1.790 187	2.148 246
1.985 936	13	1.979 819	0.30	1.797 789	2.152 248
1.989 114	20	1.982 459	0.40	1.800 629	2.154 688
1.973 153	18	1.970 841	0.25	1.788 410	2.144 069
1.972 547	32	1.963 372	0.24	1.781 942	2.136 000
1.982 200	27	1.981 861	0.31	1.798 631	2.155 690
1.980 557	29	1.974 357	0.37	1.792 326	2.146 985
1.985 316	24	1.976 216	0.29	1.794 186	2.148 844
1.989 995	23	1.984 292	0.28	1.801 461	2.157 721