

REVISTA ENERGETICA ENERGY MAGAZINE



ORGANIZACION LATINOAMERICANA DE ENERGIA

LATIN AMERICAN ENERGY ORGANIZATION

**METODOLOGIAS DE EVALUACION DE PERDIDAS
ELECTRICAS NO TECNICAS**

**METHODOLOGIES FOR ASSESSING NON-TECHNICAL
POWER LOSSES**

José Luis Calabrese

**COMPARACION ECONOMICA ENTRE CENTRALES
GEOTERMoeLECTRICAS Y PLANTAS TERMoeLECTRICAS**

**ECONOMIC COMPARISON BETWEEN GEOTHERMAL
POWER PLANTS AND THERMoeLECTRIC PLANTS**

Gerardo Hiriart Le Bert

**DOCUMENTO REGIONAL DE EXPERIENCIAS NACIONALES
EN PEQUEÑAS CENTRALES HIDROELECTRICAS**

**REGIONAL DOCUMENT ON NATIONAL EXPERIENCES
WITH SMALL HYDROPOWER STATIONS**

OLADE

AÑO 12 No. 2 AGOSTO 1988

YEAR 12 No. 2 AUGUST 1988

METODOLOGIAS DE EVALUACION DE PERDIDAS ELECTRICAS NO TECNICAS 1/

Jose Luis Calabrese *

1. INTRODUCCION

En un sistema eléctrico integrado, las pérdidas no técnicas representan un rubro muy importante que merece ser evaluado. Sin embargo cabe señalar que la falta de antecedentes sobre el tema bajo análisis le convierte en un trabajo de investigación, en el que se plantearán hipótesis a ser confirmadas o rechazadas luego de efectuar mediciones en el terreno práctico.

Mediante este estudio se busca determinar:

- El número de clientes infractores y el de no infractores en algunas ramas industriales, comerciales y de servicios.
- Las cantidades de energía subfacturadas a causa de las infracciones.
- Un sistema de facturación equitativo que permita recuperar las pérdidas por infracciones.

Para cumplir estos propósitos se emplearán técnicas de muestreo y la teoría estadística como medio para obtener los estimadores más apropiados para dar respuesta a los problemas planteados.

1/ Este documento puede ser considerado como la base del artículo que será presentado en el Simposio Latinoamericano sobre Control de Pérdidas Eléctricas, que se efectuará en Bogotá - Colombia, del 10 al 13 de octubre de 1988.

* Ingeniero Consultor.

2. DEFINICIONES GENERALES

El ámbito de ejecución de este estudio está conformado por los clientes industriales, comerciales y de servicios de la Empresa de Energía Eléctrica de Bogotá (EEEB), de acuerdo a la siguiente correspondencia con la codificación internacional CIU:

Industrial	Gran División 3
Comercial	Gran División 6
Servicios Comunales	Gran División 9
Servicios Financieros	Gran División 8

Se trata de un universo muy numeroso y heterogéneo con subdivisiones internas a 2 y 4 dígitos; cuanto más finas son estas subdivisiones, las mismas representan grupos de comportamiento más homogéneo respecto al consumo de electricidad.

Es sabido que en el universo en cuestión existe un número considerable, pero no bien conocido, de infractores, los que, mediante alteración de los contadores, logran reducir el consumo real en una magnitud también desconocida.

En consecuencia, este estudio tiene como objetivo central investigar si es posible, mediante la aplicación del método estadístico, aportar elementos de juicio cuantitativos para plantear y resolver los problemas de subfacturación que se originan por las infracciones.

La EEEB posee un Archivo Electrónico Maestro (AEM), en el cual para cada cliente se consignan las siguientes características:

- Código CIU del cliente a 4 dígitos.
- Carga contratada, W_c , que es el número de kW que el cliente solicita a la empresa.
- Último consumo en kWh, Cf.
- Promedio de los últimos consumos.
- Factor de potencia, ϕ_c .

Esta última característica se define como:

$$\phi_c = \frac{Cf \times 100}{Wc \times 1\,400} \quad (1)$$

en la cual, 1 440 es el número de horas que hay en un bimestre, que corresponde al período de facturación.

Se propone realizar encuestas sobre muestras tomadas en el AEM, con el objeto de medir las potencias de placa de todos los equipos encontrados en el establecimiento. A esta nueva potencia se le denomina potencia aforada, W_a , que define un nuevo factor:

$$\phi_a = \frac{Cf \times 100}{W_a \times 1\,440} \quad (2)$$

Los factores definidos por la ecuaciones (1) y (2) son los factores de utilización de carga contratada y aforada, respectivamente.

Mediante el barrido de las muestras se puede determinar qué clientes son infractores y cuáles no, analizando los contadores. A los primeros se les denominará L y a los segundos H.

Para los no infractores, el consumo real C_r es igual al consumo facturado:

$$C_r = C_f = 14,40 \phi_a W_a ; \quad (H) \quad (3)$$

Mientras que para los infractores:

$$C_r > C_f = 14,40 \phi_a W_a ; \quad (L) \quad (4)$$

Se formula la hipótesis, bastante razonable, de que no hay correlación entre la potencia aforada y el factor calculado con la ecuación (2). Esta hipótesis, que puede ser verificada por muestreo, significa que altas (bajas) cargas instaladas no implican necesariamente un mayor (menor) uso de las mismas. En efecto, la carga corresponde a una expectativa de nivel económico o tamaño del negocio, en tanto que el uso de ese equipamiento depende del movimiento económico real que el establecimiento tuvo al facturársele su último consumo.

Esta hipótesis se expresa mediante la condición:

$$\text{COV}(\phi_a, W_a) = 0 \quad (5)$$

Los estimadores de proporción, $p(H)$ y $p(L)$, que representan respectivamente el número de no infractores e infractores que hay en el AEM, tienen una distribución binomial, con una varianza igual a $p(1-p)/n-1$. Donde n es el respectivo tamaño muestral, y el número de individuos en el universo N , es mucho mayor que n ; en caso contrario hay que corregir por el correspondiente término de finitud $(N - n)/N$.

El universo de referencia contenido en el AEM tiene los siguientes tamaños:

CIUU	SECTOR	TAMAÑO
31	Alimentos, Bebidas y Tabaco	2 427
32	Textil, Confecciones, Calzado y Cuero	1 874
33	Maderas y Muebles	2 459
34	Papel, Imprenta	933
35	Químicos, Caucho, Plásticos	840
36	Piedras, Vidrio, Cerámica	387
37	Hierro, Acero y no Ferrosos	286
38	Maquinaria y Equipos	2 977
39	No Clasificados	162
3	TOTAL INDUSTRIAS	12 345
6310	Restaurantes	5 893
6320	Hoteles	644
8	Servicios Financieros	9 929
9	Servicios Comunales	12 573
Resto 6	Comercio General	21 536
	TOTAL COMERCIO Y SERVICIOS	50 575
	GRAN TOTAL	62 920

Cada uno de los grupos listados se considera una subpoblación del AEM, sin perjuicio de lo cual se podrán tomar divisiones más finas si la evolución del estudio así lo aconseja.

3. ENCUESTA PILOTO: RESULTADOS Y CONCLUSIONES

Para enfocar el estudio se efectuaron dos encuestas piloto en el Grupo 6320 - Hoteles y en el Sector 36 - Piedras, Vidrio y Cerámica. La encuesta tenía por objeto determinar la potencia de placa de cada establecimiento para los siguientes grupos:

MOTORES
BOMBILLOS
HORNOS
OTROS

Dicha potencia de placa se asimila a la potencia aforada. En los casos en que las placas no son legibles o simplemente no existen, se infieren de la marca y otras características del equipo. Cuando esto tampoco es posible, se efectúan mediciones con pinzas voltiamperométricas para determinar la potencia instantánea. Se busca así un procedimiento rápido que permita cubrir muchos establecimientos en corto tiempo.

Se realiza un muestreo estratificado sobre subpoblaciones en el AEM, ordenadas según consumos decrecientes; los estratos se forman arbitrariamente por rangos de consumo y dentro de cada estrato se toman muestras simples al azar. Esto permite que todos los rangos de consumo estén debidamente representados.

Los estimadores de proporción obtenidos son:

	HOTELES	PIEDRAS, ETC.
$\hat{p}(H)$	66,67 %	85,71 %
$\hat{p}(L)$	33,33 %	14,28 %

Como se ve, la proporción de infractores es alta y para el caso específico de Hoteles es sorprendentemente alta. El fenómeno de la infracción adquiere pues dimensiones de magnitud y parece depender fuertemente del tipo de actividad económica, que está muy bien captada por la codificación CIU. Con esto se comprueba que el método de muestreo propuesto es el adecuado para estimar la magnitud de los infractores.

Al examinar ahora el factor de utilización de la carga aforada y contratada, las medias y desviaciones están en el cuadro No. 1. En el caso de Hoteles (que por tener una proporción de infractores muy grande se considera una muestra muy representativa), se observa que en la clase H, prácticamente el factor

Cuadro No. 1

MEDIAS Y DESVIACIONES DE LOS FACTORES DE
UTILIZACION MUESTRALES

FACTOR DE UTILIZACION DE LA MUESTRA PILOTO PARA HOTELES

GRUPO	CARGA AFORADA	CARGA CONTRATADA
H	18 473 (16 278)	18 221 (25 258)
L	15 810 (14 327)	20 386 (19 764)

FACTOR DE UTILIZACION DE LA MUESTRA PILOTO PARA EL
SECTOR PIEDRAS, VIDRIO Y CERAMICA.

GRUPO	CARGA AFORADA	CARGA CONTRATADA
H	15 571 (14 806)	12 946 (11 567)
L	9 295 (7 775)	12 946 (10 972)

aforado y el contratado no difieren. En cambio en la clase L se cumple que:

$$\phi_a < \phi_c ; (L) \quad (6)$$

Esta desigualdad en la clase L se cumple en la muestra del Sector Piedras, debiendo anotarse que esta muestra es menos representativa por la menor cantidad de infractores. En cuanto a la clase H en el mismo sector, el factor aforado es ligeramente mayor que el contratado.

Al comparar las dos clases se comprueba que:

$$\phi_a (L) < \phi_a (H) \quad (7)$$

Nótese que esta desigualdad no necesariamente se cumple con los factores contratados.

Otra observación importante es la enorme dispersión existente alrededor de los valores medios, tanto para los factores aforados como para los contratados. Esta dispersión se mantiene también al observar la relación entre ϕ_a y ϕ_c para cada cliente: en efecto, en un caso particular puede suceder que un ϕ_c alto (bajo) dé un ϕ_a bajo (alto) independientemente de que el cliente sea H o L.

¿Cómo se puede interpretar estos resultados?

Respecto de los factores medios, e independientemente del gran tamaño de las desviaciones, la expresión (7) indica que el aforo reduce el factor del grupo L respecto del grupo H; o sea que los infractores poseen, en promedio, más carga que la que contratan, y ésta solo se conoce al aforar. En el grupo H, en cambio, la relación entre ambos factores no se puede predecir, aunque es de esperarse que el aforo no introduzca diferencias apreciables, y es lo que parecen indicar los resultados muestrales. Esto no es sino la verificación de un hecho experimental ya conocido por la EEEB, y que es el que dio la orientación inicial a este estudio.

De manera que el tipo de muestreo y tamaño muestral realizados en la encuesta piloto, además de permitir una buena estimación de la proporción de infractores, posibilita una buena estimación de los factores medios. Esto da el camino para determinar los volúmenes de energía utilizada ilícitamente. El razonamiento sería el siguiente:

El consumo real esperado del grupo H en el AEM se obtiene tomando esperanzas en la ecuación (3), lo cual, en ausencia de correlaciones, da:

$$E [Cr (H)] = 14,40 E [\phi_a (H)] E [Wa (H)] \quad (8)$$

Para el grupo L se puede suponer que si no estuviera constituido por infractores, el valor esperado de su factor aforado coincidiría con el de los del grupo H, dentro de una aproximación estadística. En consecuencia, su consumo real esperado sería:

$$E [Cr (L)] = 14,40 E [\phi_a (H)] E [Wa (L)] \quad (9)$$

Como también la potencia aforada esperada sería la misma en ambos grupos, se concluye que los consumos esperados serían los mismos en los dos grupos. El consumo esperado facturado por la empresa a los L sería aparente y se calcularía como:

$$E [Cf (L)] = 14,40 E [\phi_a (L)] E [Wa (L)] \quad (10)$$

Al efectuar la diferencia entre las ecuaciones (9) y (10), se obtiene la magnitud esperada del uso ilícito en el universo:

$$E [\Delta C (L)] = 14,40 E [Wa (L)] E [\phi_a (H) - \phi_a (L)] \quad (11)$$

Ahora, al reemplazar las esperanzas en el universo por las medias muestrales, se obtiene la magnitud promedio de la infracción estimada con un cierto error, que se lo conoce como error estándar:

$$\hat{E} [\Delta C (L)] = 14,40 \bar{W}_a (L) [\bar{\phi}_a (H) - \bar{\phi}_a (L)] \quad (12)$$

En consecuencia el total de kilovatios utilizados ilícitamente sería:

$$\Delta \hat{C} = 14,40 \hat{p}(L) N \bar{W}_a (L) [\bar{\phi}_a (H) - \bar{\phi}_a (L)] \quad (13)$$

donde:

N es el número total de establecimientos en la subpoblación.

La confiabilidad de este estimador puede analizarse sin dificultad siguiendo el método de muestreo aplicado, si bien las muestras piloto son demasiado pequeñas para dar resultados confiables.

La ecuación (12) es válida para valores esperados y la (13) para totales estimados pero no para casos individuales. Solamente en el caso en que las varianzas fueran suficientemente pequeñas, se podrían aplicar esas ecuaciones a casos individuales, y esta no es la situación que presentan las muestras. El efecto perturbador de la situación económica de ambas clases, H y L, es demasiado grande como para ser ignorada. La varianza que capta este efecto debe ser también incorporada al modelo de facturación.

4. PROPUESTA DE UN MODELO ESTADISTICO PARA LA FACTURACION DE INFRACCIONES

Lo discutido en el punto anterior muestra que sobre el factor de utilización aforado de un establecimiento particular convergen una multitud de factores, indistintamente al hecho de que el cliente sea o no infractor. Una industria o comercio puede tener una determinada potencia instalada, que no necesariamente utiliza todo el tiempo en forma constante. Como los valores promedio no reflejan las situaciones particulares de cada uno de los usuarios, si se quiere contar con un método de facturación válido para todos los casos, habrá que trabajar con toda la función de distribución de los factores de utilización. A esta función se la denomina $f(\phi)$.

Después de hacer una serie de análisis sobre las muestras piloto, se llega a la conclusión de que es posible representar razonablemente bien $f(\phi)$, por medio de algunas funciones técnicas distintas a la distribución normal. Se examinaron una serie de estas funciones y se concluyó que la FUNCION GAMMA y la FUNCION BETA poseen interesantes propiedades parecidas a las que presenta el factor de utilización. Ambas permiten, por ejemplo, captar sesgos a la izquierda o a la derecha, y ambas muestras piloto exhiben pronunciados sesgos a la izquierda. ¿Qué significa un sesgo a la izquierda en el factor de utilización aforado? Este sesgo refleja una situación económica no muy próspera donde hay muchos establecimientos con grandes cargas instaladas pero con poco uso. Hay, sin embargo, algunos establecimientos muy prósperos en medio de una mala situación general, lo cual se refleja en grandes colas a la derecha.

La función Gamma tiene por expresión:

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \quad (14)$$

para:

$$\begin{aligned} \lambda &> 0 \\ r &> 0 \end{aligned}$$

y,

$$\mu = \frac{r}{\lambda} ; \quad \sigma^2 = \frac{r}{\lambda^2}$$

donde μ y σ son la media y la varianza de x , siendo además:

$$\Gamma(r) = \int_0^{\infty} t^{r-1} e^{-t} dt \quad (15)$$

Para probar la función Gamma para el grupo H de la muestra de Hoteles, del cuadro No. 1 se obtienen los valores de μ y σ^2 :

$$\mu = 0,18473 \quad ; \quad \sigma^2 = 0,02650$$

con lo cual resulta:

$$\lambda = 6,97165 \quad r = 1,28787 \quad \Gamma(r) = 0,897$$

Se ajusta entonces la función:

$$f(x) = 13,59296 x^{0,28787} e^{-6,97165} \quad (16)$$

La función Beta tiene por expresión:

$$f(x) = \frac{1}{\beta(a,b)} x^{a-1} (1-x)^{b-1} ; \quad \begin{matrix} 0 < x < 1 \\ a > 0 \\ b > 0 \end{matrix} \quad (17)$$

$$\mu = \frac{a}{a+b} ; \quad \sigma^2 = \frac{ab}{(a+b+1)(a+b)^2} ;$$

$$\beta(a,b) = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)}$$

Y reemplazando los valores de μ y σ^2 , en las ecuaciones de los parámetros, se tiene:

$$a = 0,86513 \quad ; \quad b = 3,81807 \quad ; \quad \frac{1}{\beta(a,b)} = 2,86195$$

con lo cual se llega a la función:

$$f(x) = 2,86195 x^{-0,13487} (1-x)^{2,81807} \quad (18)$$

En el cuadro No. 2 se comparan los valores de las funciones teóricas (15) y (17) con los resultados experimentales. Para ello es más cómodo trabajar con las funciones integrales de probabilidad, obtenidas de la Encuesta Piloto de Hoteles, como son: la Función de Distribución Acumulada Experimental del Factor de Utilización Aforado, las Funciones Teóricas Gamma y Beta, y las diferencias observadas entre la Función Experimental y la Función Gamma.

Según la prueba de Kolmogorov-Smirnov para la bondad de ajuste de la función Gamma, la máxima diferencia observada (0,109) es menor que el valor crítico (0,27) para un nivel de significación de 0,01.

$$F(x) = P(X \leq x) = \int_0^x f(x) dx \quad (19)$$

La ecuación (19) indica que la probabilidad de que x sea menor que un determinado valor es $F(x)$. Como ejemplo, la figura No. 1 indica que la probabilidad de que el factor aforado de los Hoteles en el grupo de no infractores sea menor que el 40%, es de 0,875 para la función Gamma, 0,817 para la función Beta, y 0,909 para la distribución de frecuencias experimentales.

Ambas funciones teóricas parecen adaptarse bien a los resultados experimentales. Sin embargo, con un examen más cuidadoso, se observa que, para el juego particular de parámetros investigados, la función Beta no representa bien los valores pequeños, puesto que $f(x)$ tiende a ∞ cuando x tiende a 0.

Cuadro No.2

FUNCIONES INTEGRALES DE PROBABILIDAD
ENCUESTA PILOTO HOTELES

FACTOR DE POTENCIA	FUNCION	GAMMA	FUNCION	BETA	EXPERI-MENTAL	
$\phi = x$	f(x)	F(x)	f(x)	F(x)	F(ϕ)	D(x)
0,000	0,000	0,000		0,062	0,000	
0,025	3,949	0,049	4,383	0,118		
0,050	4,049	0,149	3,710	0,219		
0,100	3,489	0,337	2,901	0,384	0,394	0,057
0,200	2,121	0,618	1,896	0,624	0,727	0,109
0,300	1,187	0,783	1,232	0,780	0,818	0,035
0,400	0,642	0,875	0,768	0,800	0,909	0,034
0,500	0,341	0,924	0,446	0,941	0,939	0,015
0,600	0,179	0,950	0,232	0,975	0,970	0,020
0,700	0,093	0,964	0,101	0,992	0,970	0,006
0,800	0,048	0,971	0,032	0,998	1,000	0,029
0,900	0,025	0,974	0,004	1,000	1,000	0,026
1,000	0,013	0,976	0,000		1,000	

Para probar la bondad de ajuste de la función Gamma a los datos experimentales, se utiliza la prueba de Kolmogorov-Smirnov. Para ello se calculan los desvíos:

$$D(x) = | F(\phi) - F(x) | \quad (20)$$

y se formula la prueba de hipótesis:

$$H^0 : F(x) = F(\phi) \quad (21)$$

El máximo desvío observado es $D(0,20) = 0,109$, el cual resulta menor que el valor crítico $D_{cr} = 0,27$ leído en las tablas de bondad de ajuste para un nivel de significación del 1%. En consecuencia, la hipótesis de que Gamma representa bien la curva experimental no puede ser rechazada.

En base a lo expuesto, y dentro de las limitaciones de la encuesta piloto, se acepta que la ecuación (15) es la curva de densidad de probabilidad poblacional de los hoteles no infractores.

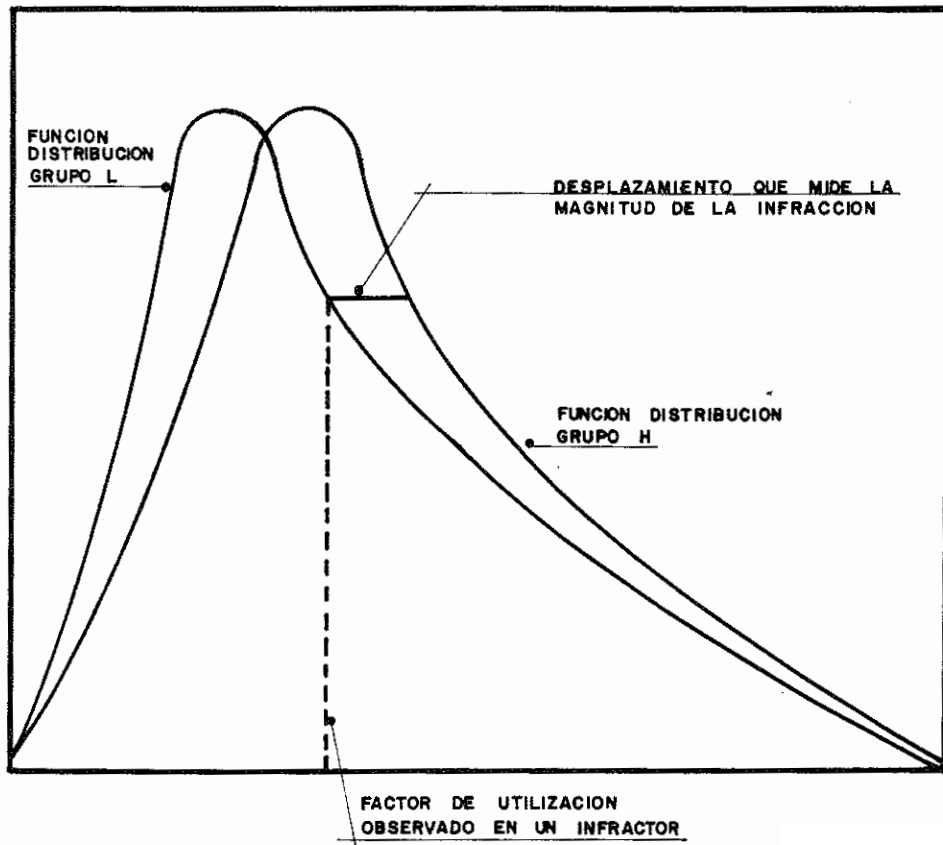
Se asume entonces que la correspondiente curva de densidad de la clase L es también una función Gamma desplazada hacia la izquierda, cuya expresión es:

$$\begin{aligned} \mu &= 0,15810 \quad ; \quad \sigma^2 = 0,02053 \\ \lambda &= 7,70232 \quad ; \quad r = 1,21774 \quad ; \quad \Gamma(r) = 0,914 \\ g(x) &= 13,14399 \times 0,21774 e^{-7,70232} \quad (22) \end{aligned}$$

Los resultados mostrados por las ecuaciones (15) y (21) permiten proponer un modelo matemático para la facturación de infractores, que será ajustado con las encuestas definitivas. La idea es que si todos los infractores de una subpoblación del AEM fueran detectados, habría que facturarles un consumo tal que permita corregir el desplazamiento hacia la izquierda de su función de distribución. La cantidad a facturar estaría entonces medida por dicho desplazamiento, tal como se ilustra en la figura No. 1. El procedimiento es equitativo ya que castigaría a cada cual según "el lugar que ocupa en la función de densidad", la cual es un reflejo de la buena o mala situación económica del cliente. Al cabo de dicho procedimiento, la EEEB recuperaría las

Figura No. 1

ILUSTRACION GRAFICA DEL PROCEDIMIENTO DE FACTURACION
PROPUESTO PARA RECUPERAR LAS PERDIDAS POR INFRACCION



pérdidas de acuerdo con las ecuaciones (12) y (13), puesto que las medias de los factores de utilización aforados que aparecen en ellas son las mismas que se emplearon para calcular los parámetros de la función Gamma, y sus respectivas varianzas son también las varianzas de esa distribución.

Para cada factor de utilización observado en la clase L se trata de encontrar un incremento δ tal, que sumado a ese factor transforme la función de distribución de los infractores en la de los no infractores. El procedimiento matemático para hallar δ en función de x es bastante complicado ya que implica trabajar con las funciones integrales de probabilidad. La ecuación de equilibrio que permite estimar el incremento δ es la siguiente:

$$\int_0^x \Gamma(r_L) \times r_L^{-1} e^{-\lambda_L x} dx = \int_0^{x+\delta} \Gamma(r_H) \times r_H^{-1} e^{-\lambda_H x} dx \quad (23)$$

Reemplazando en la ecuación (23) los parámetros de (15) en el primer miembro y los de (22) en el segundo, se obtiene la relación buscada que debe resolverse por aproximaciones.

5. PROGRAMA DE ENCUESTAS DEFINITIVAS

Lo discutido anteriormente exige modificar el plan de encuestas originalmente propuesto, ya que el mismo había sido diseñado bajo el supuesto de que sería suficiente encontrar estimadores de intervalo para cumplir con los propósitos del estudio. La necesidad, demostrada a través de la encuesta piloto, de trabajar con funciones de distribución completas y aplicar tests de estadísticas no paramétricas, implica incrementar los tamaños muestrales respecto a los inicialmente previstos.

Los nuevos tamaños muestrales vienen determinados por una doble condición:

- a. El coeficiente de variación del estimador de proporción de clase tiene por expresión:

$$CV = \sqrt{\frac{N - n}{N} \frac{1 - \hat{p}(L)}{(n - 1) \hat{p}(L)}} \quad (24)$$

Fijando este coeficiente, que se llama también error relativo de muestreo, se tendrían ciertos tamaños muestrales. Como el muestreo es sin reposición, estrictamente se debería aplicar la distribución geométrica para el cálculo de los intervalos confidenciales; la aproximación binomial o aún normal es frecuentemente utilizada. En cualquier caso la ecuación (24) expresa que el tamaño muestral para un error prefijado aumenta cuanto más pequeño es el estimador de proporción, y éste solo se conoce después de haber tomado las muestras.

- b. El parámetro de bondad de ajuste de Kolmogorov-Smirnov, para tamaños mayores que 35 y niveles de significación del 1%, tiene por expresión:

$$N(D) = \frac{1,63}{\sqrt{n}} \quad (25)$$

donde:

D es el desvío definido por la ecuación (20).

Como se puede apreciar, sería extremadamente difícil determinar a priori los tamaños muestrales requeridos para verificar ambas condiciones con la sola información de la encuesta piloto. Es más conveniente aplicar un MUESTREO EXPERIMENTAL RECURRENTE, que consiste en escoger una muestra inicial por medio de un sorteo abierto en cada estrato y cuyo tamaño muestral no sea excesivamente grande. Las muestras son luego analizadas, y se examina cómo se ajustan a las condiciones (24) y (25). Si ambas o alguna de esas condiciones no verifica los requerimientos de confiabilidad, el sorteo prosigue y el tamaño se incrementa. Con este método los tamaños muestrales son abiertos y solo se conocen cuando las muestras se juzgan satisfactorias.

Debe notarse que todo el estudio descansa en el hecho de que se verifique la inecuación (7), lo cual, vale la pena recordar, significa que el aforo permite medir el desplazamiento a la izquierda de la función de distribución de los infractores, y en base a ello montar el mecanismo matemático de la facturación a través de dicho desplazamiento. Ello quiere decir que, los efectos perturbadores, por grandes que puedan ser, afectan por igual a los grupo H y L.

Este hecho fue observado en ambas encuestas piloto y por eso se acepta como un supuesto de trabajo. Sin embargo, si alguna de las subpoblaciones no respondiera a la expresión (7), el modelo propuesto para ella ya no funcionaría y deberían emplearse otras herramientas de análisis, ya que entonces los efectos perturbado-

res (que reflejan la buena o mala situación económica) estarían correlacionados con las variables de clase H y L. El análisis se complicaría y posiblemente excedería los límites de este estudio, puesto que en vez de trabajar con funciones de densidad univariadas como las discutidas en el punto 3., se deberán considerar funciones dobles. Por ejemplo, la distribución conjunta del factor de utilización y la situación económica del negocio.

6. RESULTADOS

6.1 Análisis de las Muestras

El plan de muestreo propuesto se llevó a cabo en los siguientes grupos:

TXT	Textil, Calzado y Cuero
QMC	Químicos
PVC	Piedras, Vidrio y Cerámica
HTL	Hoteles
RST	Restaurantes
LVN	Lavanderías

De las entrevistas planificadas, un número importante no pudieron realizarse; entre las razones por las cuales una encuesta no podía concretarse a pesar de haber hecho la visita, caben mencionar: i) direcciones inexistentes; ii) dirección existente pero que no correspondía a la actividad especificada por el código CIU, al respecto se constató que hay ciertos errores de asignación de códigos en el AEM; iii) usuarios que se negaban a que se efectuara la encuesta de aforos; iv) usuarios que permitían la realización de la encuesta de aforos, pero no autorizaban que las cuadrillas de la EEEB practicaran los barridos sobre los contadores; y finalmente, v) como siempre sucede en trabajos de este tipo, un cierto número de encuestas deben rechazarse por inconsistencia en los datos.

Muchas de las unidades entrevistadas debieron visitarse dos y hasta tres veces para resolver problemas de subcaptación en los aforos.

El método de muestreo fue estratificado y se dividió el AEM en cuatro estratos:

- N1 Donde están los mayores consumidores que explican más del 60% del consumo.
- N2 Tramo intermedio con cobertura sobre el 20 ó 30% del consumo.

N3 Pequeños consumidores que explican un orden inferior al 10%.

-N Es una suerte de "cola" bastante numerosa de consumidores muy pequeños pero cuyos consumos son irrelevantes; en general no se tomaron muestras en este grupo.

La situación poblacional de los estratos se detalla en el cuadro No. 3.

En una primera inspección de las muestras, se nota que la carga contratada puede, en algunos casos, diferir significativamente de la aforada sin que haya un patrón definido para el signo de la diferencia; en otros casos, en cambio existe una notable coincidencia entre ambas magnitudes. Se concluye entonces que la carga contratada no refleja la realidad en todos los casos y ello se debe, entre otros motivos, a que no se verifica con la potencia instalada ni se actualiza con el tiempo.

Los factores de utilización de las muestras se comportan tal como se había previsto en la fase piloto, siendo $\bar{\Phi}_L < \bar{\Phi}_H$ y sus respectivas desviaciones, del mismo orden de magnitud que las medias. Nuevamente se constata que la gran dispersión es una característica de los factores de utilización y que los valores medios no son suficientes para caracterizar a todos los usuarios de un grupo CIU.

Las medias del último consumo facturado y del consumo promedio, por una parte para la clase no infractora y por otra la infractora, coinciden muy estrechamente, y las diferencias entre los factores $\bar{\Phi}_L$ y $\bar{\Phi}_H$ se mantienen, sea que los mismos hayan sido calculados con el último consumo o con el consumo promedio. Para los efectos de este estudio se trabajará con los factores referidos al último consumo facturado.

Las medias de las cargas contratadas y aforadas presentan diferencias debido a que las primeras, como ya se ha dicho, no siempre reflejan, en la realidad, la potencia instalada de los usuarios.

La situación de la carga contratada y el último consumo de los usuarios en el universo, está resumida en el cuadro No. 4.

Un test de gran importancia, antes de pasar al cálculo de los estimadores, es averiguar cuál es la influencia de la estratificación sobre el hecho de ser no infractor o infractor. En otros términos: cuál es la correlación entre los estimadores de clase H y L, y la posición que cada usuario ocupa en el AEM ordenado por consumos decrecientes. Se llega a la conclusión, no

sin sorpresa, que dicha correlación es nula y que el hecho de que un usuario consuma 10 000 ó 100 kWh al bimestre, no tiene nada que ver con que sea o no infractor. La hipótesis de no correlación había sido ya formulada a título simplificador en la fase piloto, y es ampliamente verificada en todas las submuestras.

Esto significa que la estratificación adoptada, si bien es útil porque garantiza una buena representatividad de todos los niveles de consumo, no hace ningún aporte a la bondad de la estimación, de manera que la inferencia estadística puede hacerse con independencia del criterio de estratificación. En consecuencia, las ecuaciones (8) y (9), así como las que de ellas se deducen, resultan valederas y pueden ser utilizadas para la estimación.

6.2 Estimación de la Magnitud de las Infracciones

Se determinan para cada rama de la proporción \hat{p} y el total de la clase L , que dan respectivamente el porcentaje y el número de infractores. Si en una muestra de tamaño n se observan l infractores, la distribución de probabilidades de l sigue una ley binomial.

$$P(l) = \frac{n!}{l! (n-l)!} p^l (1-p)^{n-l} \quad (26)$$

La varianzas de la proporción y el total de clase se estiman como:

$$\hat{V}(\hat{p}) = \frac{N-n}{N(n-1)} \hat{p}(1-\hat{p}) ; \quad \hat{V}(L) = N^2 \hat{V}(\hat{p}) \quad (27)$$

Para el cálculo de los límites de confianza, y para tamaños muestrales relativamente grandes, como los de esta encuesta, se puede utilizar la aproximación normal de la distribución binomial agregando un término de corrección por continuidad, con lo cual el intervalo confidencial se vuelve simétrico. Si z es el valor de la densidad de probabilidad normal para el nivel de confiabilidad escogido, dicho intervalo resulta ser igual a:

$$IC = \pm z \sqrt{\hat{V}(\hat{p})} + \frac{1}{2n} \quad (28)$$

Cuadro No. 3

ESTRATOS EN EL UNIVERSO

SECTOR	N1	N2	N3	-N	N
TXT	106	636	714	417	1 873
QMC	106	424	173	137	840
PVC	159	106	122	0	387
HTL	212	212	220	0	644
RST	4 374	0	0	1 519	5 893
LVN	544	0	0	131	675

Cuadro No. 4

CARGA Y CONSUMO REGISTRADO DE LOS GRUPOS
ENCUESTADOS EN EL UNIVERSO

Septiembre de 1985

SECTOR	Total Carga Contrada kW	Total Ultimo Consumo kWh/bimes.	Número Usuarios	Ultimo Consumo Promedio kWh/bimes.
TXT	51 502,10	6 652 445	1 873	3 532,00
QMC	45 396,20	9 269 119	840	11. 035,00
PVC	24 966,50	3 094 697	387	7 996,00
HTL	8 736,50	1 313 790	644	2 040,00
RST	44 215,40	7 779 769	5 893	1 320,00
LVN		1 059 967	675	1 370,00

En el cuadro No. 5 se observa que la proporción de infractores varía mucho con la actividad económica expresada por la rama del código CIU, desde un mínimo del 10% en el Sector Piedras, Vidrio y Cerámica hasta un 37% en Hoteles. Asimismo los intervalos confidenciales están de acuerdo con las fracciones de muestreo empleadas y se consideran razonables.

Según los resultados, hay en el grupo de usuarios investigados alrededor de un 25% de infractores, lo que confirma la presunción de la encuesta piloto, en el sentido de que la infracción adquiriría niveles sumamente altos. De esta manera, de los 10 312 usuarios investigados 2 522 son infractores, y este total de clase se da con un intervalo de ± 77 , para un nivel de significación del 90%, resultando entonces una medida muy confiable. Como es obvio, es menos confiable decir en qué rama de actividad económica se encuentran esos infractores, pero cabe señalar que si se procede a una búsqueda sistemática de ellos, trabajando conjuntamente en todas las ramas, este procedimiento se beneficia de la alta probabilidad que se tiene de que el número total sea del orden de los 2 500. El alto grado de confianza con que se puede afirmar que si un infractor no está en una rama debe necesariamente estar en otra, respalda la búsqueda conjunta, ya que al fin de cuentas la empresa busca infractores que le permitan recuperar sus pérdidas, casi con independencia de en qué grupo estén. Los estimadores por grupo tienen gran utilidad para trazar líneas de acción, hecho lo cual hay que trabajar simultáneamente en todos ellos.

La infracción promedio por ramas es extremadamente variable (cuadro No. 6) y conduce a conclusiones curiosas. El grupo donde menos infractores hay es en el Sector Piedras, Vidrio y Cerámica, pero es al mismo tiempo de los que más usan en forma ilícita, ya que un 45% de lo que consumen no está registrado por los contadores. El grupo donde hay más infractores es Hoteles, pero la magnitud de su infracción es un discreto 17% de su consumo real. Los Sectores Químicos y Restaurantes tienen muchos infractores con una apreciable magnitud de uso ilícito, pero es en el Sector Lavanderías el caso crítico, presentando numerosos infractores con un nivel de utilización ilegal del 60%, que es el más alto de la muestra.

En el cuadro No. 7 se indica la resolución final del problema de estimar infracciones, indicando los consumos registrados en los listados del AEM y los totales usados ilícitamente estimados a partir de la encuesta de aforos. Al sumar las infracciones a los registros se obtiene el consumo real, y se ve que la infracción es un 5,54% del consumo real y que es mayor que lo que se factura a la totalidad de los Hoteles o a la totalidad de las Lavanderías.

Cuadro No. 5

ESTADÍSTICAS DE LA PROPORCIÓN Y EL TOTAL DE CLASE PARA INFRACTORES

SECTOR	Proporción Infractores	Error Estandar	Intervalo Confidencial para 90%	Total de Clase Infractor	Intervalo Confidencial para 90%
TXT	17,04	3,93	± 6,57	319	± 120
QMC	24,19	5,27	± 8,91	203	± 72
PVC	10,42	4,17	± 7,23	40	± 26
HTL	37,50	5,78	± 9,92	241	± 61
RST	25,48	2,97	± 4,95	1 501	± 287
LVN	32,26	8,32	± 13,67	218	± 92
TOTAL	24,47	1,87	± 3,07	2 522	± 77

Cuadro No. 6

CARGA AFORADA Y CONSUMO REGISTRADO DE INFRACTORES Y NO INFRACTORES EN EL UNIVERSO

SECTOR	Consumo Registrado Promedio		Potencia Aforada Promedio en kW	Infracción Promedio en kWh por bimestre
	NO INFRACTOR en kWh por bimestre	INFRACTOR		
TXT	3 600,47	3 315,59	13,72	284,88 (8%)
QMC	11 569,73	9 353,18	38,73	2212,82 (20%)
PVC	8 354,90	4 908,74	37,30	3443,16 (43%)
HTL	2 177,48	1 811,56	8,86	365,92 (17%)
RST	1 414,90	1 044,28	4,05	370,62 (26%)
LVN	2 933,76	1 153,27	10,05	1780,49 (60%)

Cuadro No. 7

CONSUMO REGISTRADO, INFRACCION Y CONSUMO REAL

SECTOR	Consumo Registrado kWh/bimestre	Infracción kWh/bimestre	Consumo Real kWh/bimestre	Porcentaje de la Infracción
TXT	6 652 445	90 922	6 743 367	1,35
QMC	9 269 119	449 636	9 718 755	4,63
PVC	3 094 697	138 847	3 233 544	4,29
HTL	1 313 790	88 370	1 402 160	6,30
RST	7 779 769	556 499	8 336 268	6,68
LVN	1 059 967	387 711	1 447 678	26,78
TODOS	29 169 787	1 711 985	30 881 772	5,54

Comparando la situación entre las ramas de actividad, se observa que los mayores robos se localizan en el Sector Lavanderías y los menores en el Sector Textil, Calzado y Cueros. Los otros grupos tienen niveles de uso ilegal entre el 4 y 7%.

De acuerdo con la teoría implícita en la expansión de las muestras, los más altos niveles de infracción se producen cuando se cumplen dos condiciones: i) una elevada proporción de infractores y, ii) una gran diferencia entre los factores de utilización de H y L. Cuando se cumple solo con la primera, habrá "muchos infractores que hacen uso ilegal de poco" y si se cumple solo la segunda serán los "pocos que hacen uso ilegal de mucho". Como se ve, de todo hay en el AEM, y el caso extremo de "muchos que hacen uso ilegal de mucho" es el de las Lavanderías.

Los infractores de los sectores de Lavanderías y Piedras, Vidrio y Cerámica tienen factores de utilización especialmente bajos y es posible que esto se deba a que son sectores que hacen un uso térmico intensivo de la electricidad.

6.3 Modelo para la Recuperación de Pérdidas

El modelo discutido en el punto 3 se aplica aquí para representar el factor de utilización de no infractores e infractores. Para ello se emplea la función Gamma de la ecuación (14); para mayor simplicidad de notación se expresan como $f(x)$ y $F(x)$ las funciones de probabilidad de no infractores y con $g(x)$ y $G(x)$ las correspondientes a los infractores.

El modelo que da el mejor ajuste corresponde a una función Gamma truncada y que además cumple la condición de que la función de densidad de probabilidad de los infractores se origina al desplazar la de los no infractores hacia la izquierda, conservando la misma dispersión. Como al aplicar dicho procedimiento es posible que la función de densidad en el origen se haga infinita, el modelo se denomina Gamma Infinito.

En la figura No. 2 se presentan dos ejemplos de bondad de ajuste para los sectores de Restaurantes y Textiles, entre el modelo Gamma Infinito y los datos experimentales. La aplicación del test de Kolmogorov-Smirnov expresado por las ecuaciones (20), (21) y (25), conduce a aceptar la hipótesis de que el modelo representa bien los datos experimentales.

En el cuadro No. 8 se presentan los parámetros de calibración del modelo para cada una de las ramas. Se comprueba que el modelo adoptado funciona bien en todas las ramas investigadas. Sin embargo, los parámetros varían mucho de una rama a otra y no

es conveniente hacer extrapolaciones para grupos en que no se han realizado encuestas, puesto que la función de distribución Gamma truncada debe calibrarse sobre un conjunto de datos experimentales de agregados relativamente homogéneos, y no cabe duda que esos parámetros dependen estrechamente de la tecnología del consumo de energía. En el cuadro No. 9 se indican los valores de las funciones de ajuste y en ellas pueden apreciarse las diferencias por ramas. Se nota, por ejemplo, que para factores menores o iguales al 10% la función de distribución de no infractores cubre el 38% en TXT pero solamente el 24% en RST.

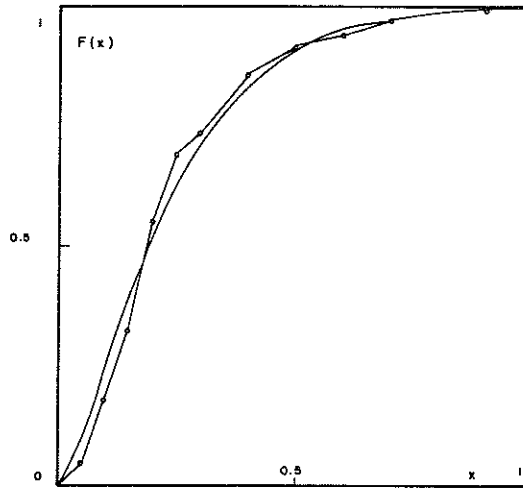
En la figura No. 3 se dan las gráficas de $F(x)$ y $G(x)$ para los 6 grupos investigados. El desplazamiento hacia la izquierda de $G(x)$ con respecto a $F(x)$ mide la infracción, que en términos del factor de utilización se designa por δ , que es una función de x cuya esperanza es:

$$E(\delta) = \int_0^1 \delta(x)g(x)dx = \bar{\phi}_H - \bar{\phi}_L \quad (29)$$

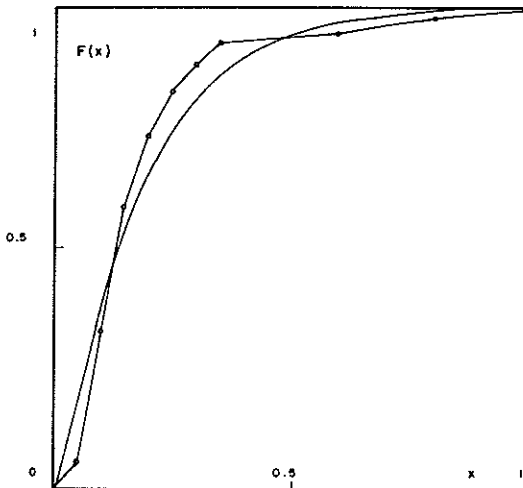
La ecuación (29) expresa que el desplazamiento es tal que en promedio debe producir las diferencias promediadas entre los factores de utilización de no infractores e infractores. Para encontrar el valor de δ en cada punto, se debe resolver la ecuación (23); todo el proceso de cálculo que comienza con la calibración de las funciones de densidad de probabilidad y termina con el cálculo de los desplazamientos, implica grandes dificultades prácticas, aun en computador, debido a que se necesita hacer iteraciones recurrentes sobre funciones que, o bien tienen valores infinitos, o representan curvaturas muy pronunciadas con abruptos cambios de pendiente. Las respectivas soluciones solo pueden ser aproximadas y es conveniente primero efectuar una calibración gruesa para conocer sus propiedades. Se obtienen los resultados de la figura No. 4, que representa una familia de curvas de δ como función de x : las ordenadas dan la cantidad que hay que aumentar al factor de utilización de un infractor perteneciente a una cierta rama para que, al hacer lo propio con todos los infractores que hay en esa rama, se obtenga un desplazamiento de la función de densidad hacia la derecha, como consecuencia de lo cual se recuperan las pérdidas.

Figura No. 2

AJUSTE DE LA FUNCION GAMMA TRUNCADA A LOS DATOS
EXPERIMENTALES DE LA MUESTRA DE RESTAURANTES



AJUSTE DE LA FUNCION GAMMA TRUNCADA A LOS DATOS
EXPERIMENTALES DE LA MUESTRA DE TEXTILES



Cuadro No. 8

MODELO GAMMA INFINITO

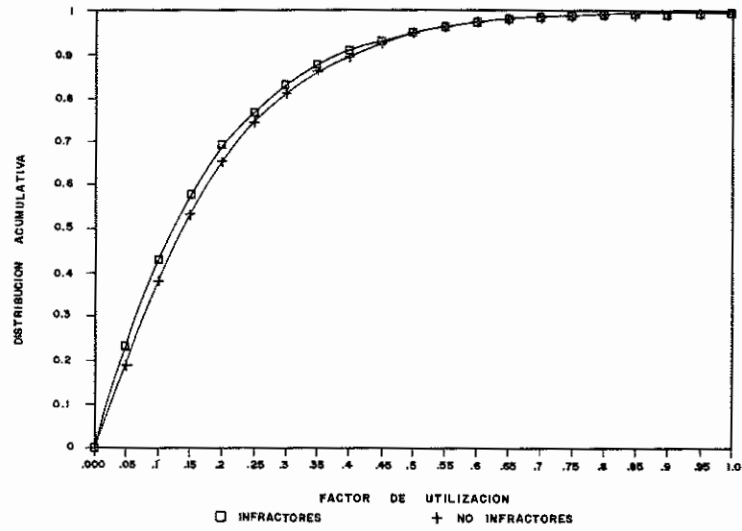
(Parámetros)

	TXT	QMC	PVC	HTL	RST	LVN
MU:L	16,782	16,776	9,139	14,199	17,906	7,969
MU:H	18,224	20,745	15,555	17,067	24,261	12,216
SD:L	11,355	16,506	5,537	13,449	17,564	3,862
SD:H	16,766	19,617	13,415	15,735	19,116	5,955
r:L	1,0019	0,7313	0,4641	0,8143	0,8774	1,7908
λ :L	5,9701	4,3594	5,0783	5,7349	4,9001	22,4719
$\Gamma(r)$:L	1,0000	1,2507	1,9082	1,1471	1,0885	0,9288
r:H	1,1815	1,1183	1,3445	1,1765	1,6107	4,2082
λ :H	6,4831	5,3907	8,6435	6,8932	6,6392	34,4481
$\Gamma(r)$:H	0,9237	0,9436	0,8922	0,9237	0,8947	7,8365
$\lambda^r/\Gamma(r)$:L	5,9905	2,3468	1,1141	3,6146	3,7048	283,5259
r-1:L	0,0019	-0,2687	-0,5359	-0,1857	-0,1226	0,7908
- λ :L	-5,9701	-4,3594	-5,0783	-5,7349	-4,9001	-22,4719
Trunc	0,9723	0,9725	0,9886	0,9881	0,9899	1,0039
$\lambda^r/\Gamma(r)$:H	9,8530	6,9731	20,3657	10,4914	23,5797	375 448
r-1:H	0,1815	0,1183	0,3445	0,1765	0,6107	3,2082
- λ :H	-6,4831	-5,3907	-8,6435	-6,8932	-6,6392	-34,4481
Trunc	0,9843	0,9808	0,9897	0,9858	0,9897	0,9724

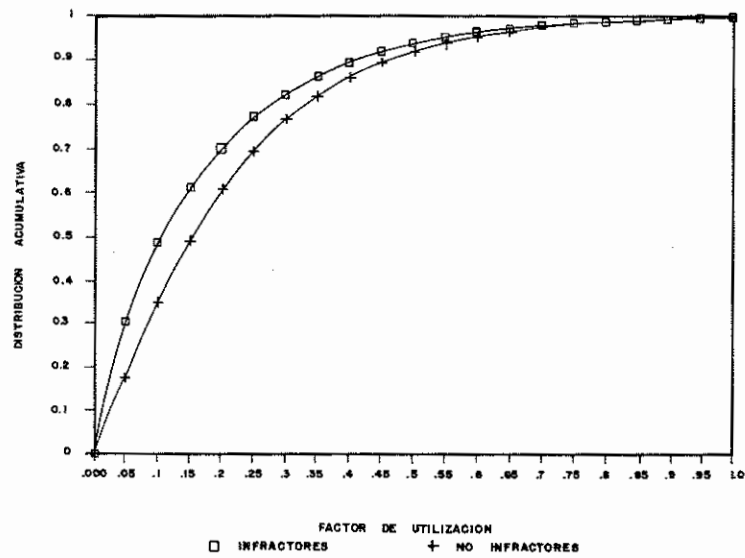
MU Media Aforada
SD Desviación Aforada

Figura No. 3

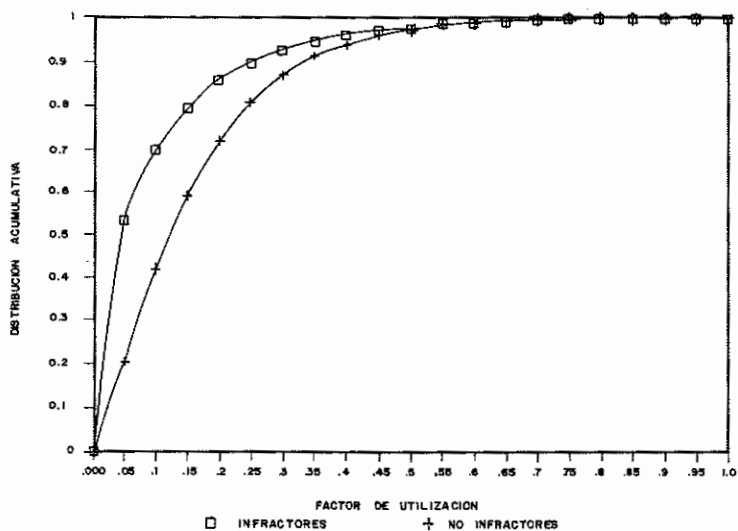
TEXTIL, CALZADO Y CUERO



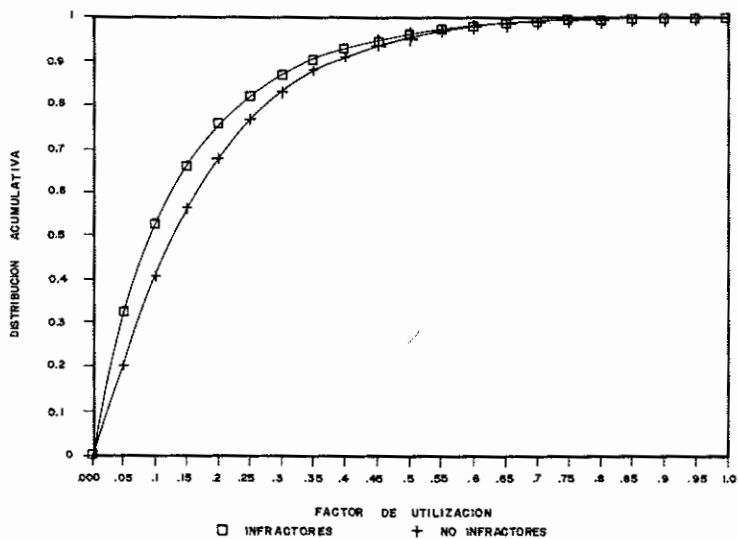
QUIMICOS



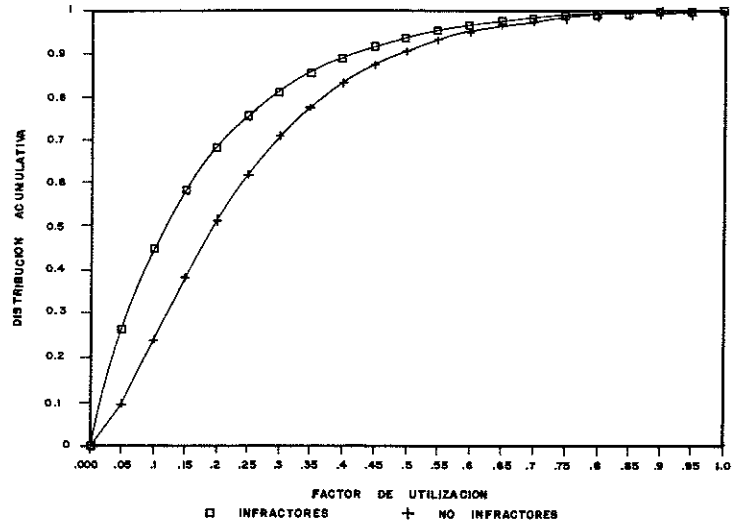
PIEDRAS, VIDRIO Y CERAMICA



HOTELES



RESTAURANTES



LAVANDERIAS

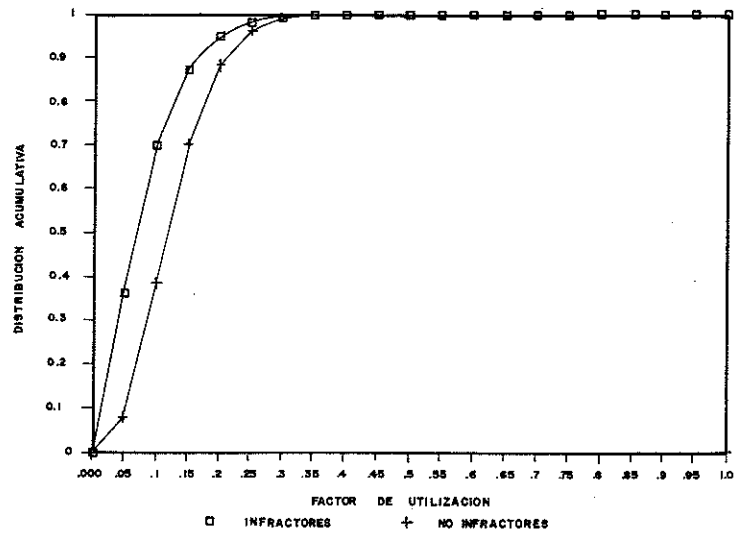
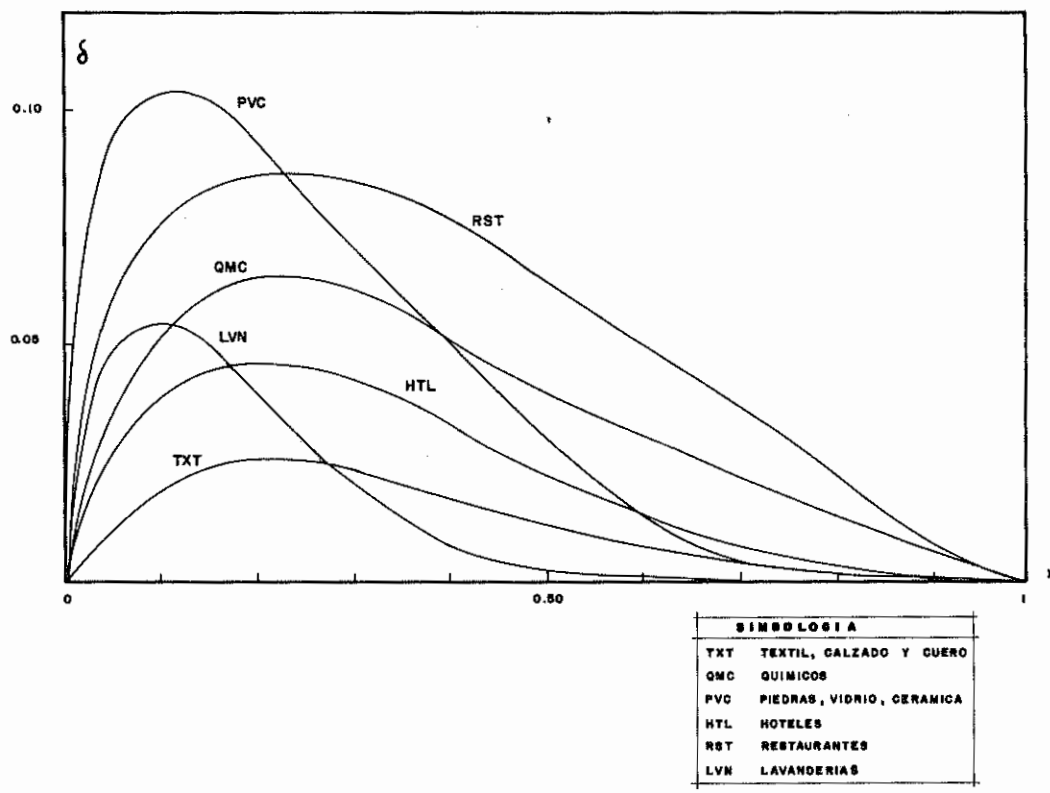


Figura No. 4

DESPLAZAMIENTO DEL FACTOR DE ACTUALIZACION DE LOS INFRACTORES
PARA RECUPERAR LAS PERDIDAS POR INFRACCION



Cuadro No. 9

MODELO GAMMA INFINITO PARA LOS DIFERENTES SECTORES

x	TEXTIL, CALZADO Y CUERO				QUIMICOS			
	g(x)	G(x)	f(x)	F(x)	g(x)	G(x)	f(x)	F(x)
0,00	6,0271	0,0000	1,2387	0,0000	53,2028	0,0000	1,8209	0,0000
0,01	5,7535	0,0288	4,0673	0,0203	7,9616	0,0796	3,9069	0,0195
0,02	5,4273	0,0847	4,3229	0,0623	6,3269	0,1511	4,0182	0,0592
0,03	5,1167	0,1374	4,3609	0,1057	5,4318	0,2099	3,9924	0,0992
0,04	4,8228	0,1871	4,3062	0,1490	4,8134	0,2611	3,9159	0,1388
0,05	4,5453	0,2339	4,2027	0,1916	4,3399	0,3068	3,8096	0,1774
0,10	3,3767	0,4320	3,4465	0,3828	2,8969	0,4878	3,1582	0,3516
0,15	2,5072	0,5791	2,6826	0,5360	2,0891	0,6124	2,5305	0,4938
0,20	1,8612	0,6883	2,0439	0,6542	1,5550	0,7035	1,9996	0,6071
0,25	1,3814	0,7694	1,5391	0,7438	1,1777	0,7718	1,5680	0,6963
0,30	1,0253	0,8295	1,1504	0,8110	0,9018	0,8238	1,2236	0,7660
0,35	0,7609	0,8742	0,8555	0,8612	0,6957	0,8638	0,9517	0,8204
0,40	0,5647	0,9073	0,6338	0,8984	0,5398	0,8946	0,7384	0,8627
0,45	0,4190	0,9319	0,4682	0,9260	0,4205	0,9187	0,5719	0,8954
0,50	0,3110	0,9502	0,3451	0,9463	0,3287	0,9374	0,4422	0,9208
0,55	0,2308	0,9637	0,2539	0,9613	0,2577	0,9520	0,3416	0,9404
0,60	0,1712	0,9738	0,1866	0,9723	0,2024	0,9636	0,2636	0,9555
0,65	0,1271	0,9812	0,1369	0,9804	0,1593	0,9726	0,2032	0,9672
0,70	0,0943	0,9867	0,1003	0,9863	0,1256	0,9797	0,1566	0,9762
0,75	0,0700	0,9908	0,0735	0,9906	0,0991	0,9853	0,1206	0,9831
0,80	0,0519	0,9939	0,0537	0,9938	0,0783	0,9898	0,0928	0,9884
0,85	0,0385	0,9962	0,0393	0,9962	0,0620	0,9933	0,0714	0,9926
0,90	0,0286	0,9978	0,0287	0,9979	0,0491	0,9961	0,0549	0,9957
0,95	0,0212	0,9991	0,0210	0,9991	0,0389	0,9993	0,0422	0,9981
1,00	0,0157	1,0000	0,0153	1,0000	0,0309	1,0000	0,0324	1,0000

Cuadro No. 9

(Continuación)

x	PIEDRAS, VIDRIOS Y CERAMICA				HOTELES			
	g(x)	G(x)	f(x)	F(x)	g(x)	G(x)	f(x)	F(x)
0,00	538,6685	0,0000	0,3899	0,0000	31,0282	0,0000	1,3953	0,0000
0,01	12,6358	0,2527	3,8626	0,0193	8,1238	0,0812	4,4072	0,0220
0,02	8,2838	0,3573	4,4983	0,0611	6,7445	0,1556	4,6489	0,0673
0,03	6,3359	0,4304	4,7443	0,1073	5,9067	0,2188	4,6611	0,1139
0,04	5,1618	0,4879	4,8047	0,1551	5,2873	0,2748	4,5772	0,1601
0,05	4,3532	0,5355	4,7590	0,2029	4,7900	0,3252	4,4439	0,2052
0,10	2,3292	0,7025	3,9222	0,4199	3,1616	0,5240	3,5580	0,4052
0,15	1,4540	0,7971	2,9275	0,5912	2,2012	0,6581	2,7077	0,5619
0,20	0,9668	0,8576	2,0982	0,7168	1,5665	0,7522	2,0182	0,6800
0,25	0,6655	0,8984	1,4706	0,8060	1,1283	0,8196	1,4872	0,7676
0,30	0,4682	0,9268	1,0166	0,8682	0,8188	0,8683	1,0881	0,8320
0,35	0,3344	0,9469	0,6958	0,9110	0,5973	0,9037	0,7921	0,8790
0,40	0,2415	0,9613	0,4729	0,9402	0,4374	0,9296	0,5746	0,9231
0,45	0,1759	0,9717	0,3197	0,9601	0,3213	0,9485	0,4156	0,9379
0,50	0,1290	0,9793	0,2152	0,9734	0,2365	0,9625	0,3000	0,9558
0,55	0,0951	0,9849	0,1443	0,9824	0,1744	0,9727	0,2161	0,9687
0,60	0,0704	0,9890	0,0965	0,9884	0,1288	0,9803	0,1555	0,9780
0,65	0,0523	0,9921	0,0644	0,9925	0,0953	0,9859	0,1117	0,9847
0,70	0,0390	0,9944	0,0429	0,9951	0,0706	0,9901	0,0802	0,9895
0,75	0,0292	0,9961	0,0285	0,9969	0,0523	0,9932	0,0575	0,9929
0,80	0,0218	0,9974	0,0189	0,9981	0,0388	0,9954	0,0412	0,9954
0,85	0,0165	0,9983	0,0125	0,9989	0,0288	0,9971	0,0295	0,9972
0,90	0,0123	0,9991	0,0083	0,9994	0,0214	0,9984	0,0211	0,9984
0,95	0,0093	0,9996	0,0055	0,9998	0,0159	0,9993	0,0151	0,9994
1,00	0,0070	1,0000	0,0036	1,0000	0,0118	1,0000	0,0108	1,0000

Cuadro No. 9
(Continuación)

x	RESTAURANTES				LAVANDERIAS			
	g(x)	G(x)	f(x)	F(x)	g(x)	G(x)	f(x)	F(x)
0,00	15,3487	0,0000	0,0211	0,0000	0,0314	0,0000	0,0000	0,0000
0,01	6,2668	0,0627	1,3389	0,0067	5,9121	0,0296	0,1049	0,0005
0,02	5,4810	0,1214	1,9132	0,0230	8,1696	0,1000	0,6869	0,0045
0,03	4,9658	0,1736	2,2933	0,0440	8,9920	0,1858	1,7873	0,0169
0,04	4,5645	0,2213	2,5582	0,0682	9,0170	0,2758	3,1872	0,0417
0,05	4,2290	0,2653	2,7434	0,0948	8,5922	0,3639	4,6208	0,0808
0,10	3,0404	0,4470	3,0059	0,2385	4,8326	0,6995	7,6286	0,3870
0,15	2,2643	0,5796	2,7628	0,3827	2,1650	0,8744	5,0043	0,7028
0,20	1,7109	0,6790	2,3631	0,5108	0,8837	0,9506	2,2498	0,8842
0,25	1,3030	0,7543	1,9431	0,6185	0,3427	0,9813	0,8223	0,9610
0,30	0,9973	0,8118	1,5584	0,7060	0,1287	0,9931	0,2636	0,9881
0,35	0,7660	0,8559	1,2286	0,7757	0,0473	0,9975	0,0772	0,9966
0,40	0,5898	0,8898	0,9564	0,8303	0,0171	0,9991	0,0212	0,9991
0,45	0,4550	0,9159	0,7374	0,8727	0,0061	0,9997	0,0055	0,9998
0,50	0,3516	0,9361	0,5643	0,9052	0,0022	0,9999	0,0014	0,9999
0,55	0,2720	0,9517	0,4292	0,9301	0,0008	1,0000	0,0003	1,0000
0,60	0,2106	0,9638	0,3247	0,9489	0,0003	1,0000	0,0001	1,0000
0,65	0,1632	0,9731	0,2447	0,9632	0,0001	1,0000	0,0000	1,0000
0,70	0,1266	0,9804	0,1837	0,9739	0,0000	1,0000	0,0000	1,0000
0,75	0,0983	0,9860	0,1375	0,9819	0,0000	1,0000	0,0000	1,0000
0,80	0,0763	0,9903	0,1026	0,9879	0,0000	1,0000	0,0000	1,0000
0,85	0,0593	0,9937	0,0764	0,9924	0,0000	1,0000	0,0000	1,0000
0,90	0,0461	0,9964	0,0568	0,9957	0,0000	1,0000	0,0000	1,0000
0,95	0,0358	0,9984	0,0421	0,9982	0,0000	1,0000	0,0000	1,0000
1,00	0,0279	1,0000	0,0312	1,0000	0,0000	1,0000	0,0000	1,0000

6.4 Ejemplo de Cálculo

Para un restaurante infractor cuyo factor de utilización aforado fuera $x = 0,10$ y cuyo consumo registrado sea $Cr = 850$ kWh/bimes. Para ese valor de x , se lee en la correspondiente curva $\delta = 0,075$. El factor de utilización corregido será igual a: $x + \delta = 0,175$. Si se aplica ahora la ecuación (2), siendo Cr' el consumo real, se tiene:

$$\frac{Cr}{Cr'} = \frac{x}{x + \delta} \quad (30)$$

con lo cual el consumo real resulta ser $Cr = 1\,487$ kWh/bimes. De modo que un restaurante de estas características está usando ilícitamente 637 kWh/bimestre, que es un 42,8% de lo que consume. Esto es así porque, dentro de la aproximación estadística del test de bondad de ajuste de Kolmogorov-Smirnov, los restaurantes no infractores se comportan de acuerdo con la función Gamma, con parámetros específicos del grupo de restaurantes, porque todo infractor sale de esa familia y se desplaza hacia la izquierda en una magnitud δ que por definición corresponde a la magnitud de su infracción.

Si se recorre toda la curva del Sector de Restaurantes en la figura No. 4 y se aplica a la ecuación (30) para todos los infractores detectados en esa rama, la ecuación (29) asegura que se recobrarán las pérdidas a lo largo del proceso de recuperación, puesto que debe cumplirse que:

$$Cr' - Cr = 14,4 N \hat{p} \bar{W}_a E(\delta) \quad (31)$$

Para obviar la aproximación gruesa del procedimiento gráfico, se elaboró un programa de computador que permite llegar a un ajuste más fino, de manera que, al introducir un valor de x y seleccionar la rama apropiada, se obtiene el desplazamiento por medio de un proceso iterativo.

Jose Luis Calabrese *

1. INTRODUCTION

In an interconnected power system, the non-technical losses constitute a significant element meriting assessment. However, it is worthwhile to note that the lack of previous studies on the subject under analysis makes it a topic for research, where hypotheses will be formulated for later confirmation or rejection based on in-the-field investigations.

The aim of this study is to determine:

- The number of non-paying energy users and of paying energy users in various branches of industry, commerce and services.
- The amounts of underbilled energy due to infractions.
- An equitable billing system making it possible to recover the losses due to infractions.

To do this, sampling techniques and statistical theory will be used to obtain the most suitable indicators to respond to the issues set forth above.

1/ This document may be considered as the basis for an article to be presented to the Latin American Symposium on Power Loss Reduction, which will be held in Bogota, Colombia, from October 10 to 13, 1988.

* Engineer Consultant.

2. GENERAL DEFINITIONS

This study covers the industrial, commercial and service sector customers of the Bogota Electric Power Company (EEEB), in keeping with the following items under the ISIC international code:

Industrial	Major Heading 3
Commercial	Major Heading 6
Financial Services	Major Heading 8
General Services	Major Heading 9

The universe is numerous and quite heterogeneous, with two- and four-digit subdivisions. The more refined the subdivisions, the more homogeneous the behavior of the groups with respect to electricity consumption.

In the universe in question there is a considerable, but unknown, number of non-paying energy users who, by altering their meters, manage to reduce real consumption by a magnitude also unknown.

Consequently, the main purpose of this study is to investigate whether, by applying statistical methods, it is possible to contribute quantitative criteria for tackling and solving the problems of underbilling due to infractions.

The EEEB has a Master Electronic File (AEM), in which each customer is identified by the following characteristics:

- Four-digit ISIC code number.
- Granted load, W_c , which is the number of kW which the customer contracts for with the power company.
- Last billed consumption, in kWh, C_f .
- Average recent consumption.
- Load factor, ϕ_c .

This last characteristic is defined as:

$$\phi_c = \frac{C_f \times 100}{W_c \times 1\,400} \quad (1)$$

where 1 440 is the number of hours in a two-month period, which is the billing period.

The proposal is to conduct surveys on the samples drawn from the AEM, in order to measure the nominal power of all of the equipment found in an establishment. This new power will be termed "measured" or "real" power, W_a , which defines a new factor:

$$\phi_a = \frac{C_f \times 100}{W_a \times 1\,440} \quad (2)$$

The factors defined by equations (1) and (2) are the utilization factors for granted and measured load, respectively.

Through a review of the sampling, it is possible to determine which clients are non-paying energy users and which are not, by analyzing the meters. The former will be denoted by "L" and the latter by "H".

For the paying energy users, real consumption C_r is equal to billed consumption:

$$C_r = C_f = 14.40 \phi_a W_a ; \quad (H) \quad (3)$$

Whereas for the non-paying energy users:

$$C_r \neq C_f = 14.40 \phi_a W_a ; \quad (L) \quad (4)$$

A fairly reasonable hypothesis is that there is no correlation between the measured power and the factor calculated from equation (2). This hypothesis, which may be verified by sampling, means that high (low) installed loads do not necessarily imply a greater (lesser) use of the loads. In fact, the load responds to expectations for economic level or size of the business, while the use of the equipment depends on the actual economic movement that the establishment had when its last consumption was billed.

This hypothesis is expressed by:

$$\text{COV}(\phi_a, W_a) = 0 \quad (5)$$

The proportion estimators $p(H)$ and $p(L)$, which respectively represent the number of paying energy users and non-paying energy users in the AEM, have a binomial distribution, with a variance equal to $p(1-p)/n-1$, where n is the respective sample size and the number of individuals in the universe N is much larger than n . Otherwise, it will be necessary to correct for the corresponding finite term $(N-n/N)$.

The reference universe contained in the AEM has the following sizes:

ISIC	SECTOR	SIZE
31	Food, Beverages and Tobacco	2 427
32	Textiles, Apparel, Footgear and Leather	1 874
33	Wood and Furniture	2 459
34	Paper and Pulp	933
35	Chemicals, Rubber, Plastics	840
36	Stone, Glass, Ceramics	387
37	Iron, Steel and Non-Ferrous Metals	286
38	Machinery and Equipment	2 977
39	Unclassified	162
3	TOTAL INDUSTRIES	12 345
6310	Restaurants	5 893
6320	Hotels	644
8	Financial Services	9 929
9	Community Services	12 573
Rest of 6	General Commerce	21 536
	TOTAL COMMERCE AND SERVICES	50 575
	GRAND TOTAL	62 920

Each one of the groups listed is considered to be a sub-population of the AEM, without detriment to which more refined divisions may be taken if advised.

3. PILOT SURVEY: FINDINGS AND CONCLUSIONS

To focus the study, two pilot surveys were taken in Group 6320 (Hotels) and in Sector 36 (Stone, Glass and Ceramics). The aim of the survey was to determine the nominal power of each establishment for the following groups:

MOTORS
 LIGHT BULBS
 FURNACES
 OTHERS

This nominal power is assimilated in the measured power. In the cases in which the power ratings are not legible or simply do not exist, they can be inferred from the brand name and other equipment features. When this is not possible either, measurements can be taken using voltiamperometric pins to determine the instantaneous power. In this way, a quick procedure is employed to make it possible to cover as many establishments as possible in the shortest time.

A stratified sample is done on the AEM subpopulations, in order of decreasing consumption; the strata are arbitrarily formed by ranges of consumption, and within each stratum simple random sampling is done. This makes it possible for all of the consumption ranges to be duly represented.

The proportion estimators obtained are:

	HOTELS	STONE, ETC.
$\hat{p}(H)$	66.67%	85.71%
$\hat{p}(L)$	33.33%	14.28%

As can be seen, the proportion of non-paying energy users is high; for the specific case of hotels, surprisingly so. The phenomenon of infractions takes on sizeable dimensions and seems to depend heavily on the type of economic activity, which is represented well by the ISIC code. This proves that the proposed sampling method is the most suitable for estimating the number of non-paying energy users.

The averages and deviations found in Table No. 1 will be useful in examining the utilization factor for the measured and granted loads. In the case of hotels (which is considered to be very representative due to the large number of non-paying energy users), it can be seen that the measured load and the granted load practically do not differ in class H. However, in class L:

$$\phi_a < \phi_c ; (L) \quad (6)$$

This inequality in class L is seen in the Stone Sector, where the sample is less representative due to the lower number of non-paying energy users. As for class H in the same sector, the measured factor is slightly higher than the contractual load.

Table No. 1

AVERAGES AND DEVIATIONS FOR THE SAMPLES'
UTILIZATION FACTORS

UTILIZATION FACTOR FOR THE PILOT SAMPLE OF HOTELS

GROUP	MEASURED LOAD	GRANTED LOAD
H	18 473 (16 278)	18 221 (25 258)
L	15 810 (14 327)	20 386 (19 764)

UTILIZATION FACTOR FOR THE PILOT SAMPLE OF THE
STONE, GLASS AND CERAMICS SECTOR

GROUP	MEASURED LOAD	GRANTED LOAD
H	15 571 (14 806)	12 946 (11 567)
L	9 295 (7 775)	12 946 (10 972)

On comparing the two classes, it can be seen that:

$$\phi_a (L) < \phi_a (H) \quad (7)$$

Note that this inequality does not necessarily hold true for the granted factors.

Another important observation is the enormous scattering around the mean values, both for the measured factors as well as for the granted ones. This scattering is also maintained in the ratio between ϕ_a and ϕ_c for each customer: in fact, in a given case it may occur that a high (low) ϕ_c would yield a high (low) ϕ_a , no matter whether the customer is H or L.

How can these findings be interpreted?

With respect to the average factors and independently of the large deviation, equation (7) indicates that the assessment reduces the factor of group L with respect to group H, i.e. on average the non-paying energy users have a larger load than the one they have contracted for, but this can only be learned through an assessment. In the case of group H, the ratio between the two factors cannot be predicted, although it is to be expected that the assessment will not introduce appreciable differences, and this is what the sample results seem to indicate. This is merely a verification of experimental findings which are already known by the EEEB, and which gave this study its initial orientation.

Thus, the type of sampling and sample size in the pilot survey, in addition to enabling a sound estimate of the proportion of non-paying energy users, permits a sound estimate of the mean factors and determination of the volumes of energy used illegally. The reasoning would be as follows:

In the AEM, the actual consumption expected for class H is obtained by taking expectations in equation (3). In the absence of correlations, this is:

$$E [Cr (H)] = 14.40 E [\phi_a (H)] E [Wa (H)] \quad (8)$$

For class L, it may be supposed that, if not made up of non-paying energy users, the expected value for its measured factor would coincide with the one for those in class H, within a statistical approximation.

$$E [Cr (L)] = 14.40 E [\phi_a (H)] E [W_a (L)] \quad (9)$$

Since the expected measured load would also be the same for both classes, it may be concluded that the expected consumption would be the same for the two. The expected consumption billed by the firm to members of class L would be apparent and would be calculated as follows:

$$E [Cf (L)] = 14.40 E [\phi_a (L)] E [W_a (L)] \quad (10)$$

In finding the difference between equations (9) and (10), the expected magnitude of fraudulent use expected in the universe is obtained:

$$E [\Delta C (L)] = 14.40 E [W_a (L)] E [\phi_a (H) - \phi_a (L)] \quad (11)$$

By replacing the expectations in the universe by means from the sample, the average magnitude of infraction estimated with a standard error is obtained:

$$\hat{E} [\Delta C (L)] = 14.40 \bar{W}_a (L) [\bar{\phi}_a (H) - \bar{\phi}_a (L)] \quad (12)$$

Consequently, the total number of kilowatts used illegally would be:

$$\Delta \hat{C} = 14.40 \hat{p}(L) N \bar{W}_a (L) [\bar{\phi}_a (H) - \bar{\phi}_a (L)] \quad (13)$$

where:

N is the total number of establishments in the subpopulation.

The reliability of this estimator can be analyzed without difficulty following the applied sampling method, even though the pilot surveys are too small to yield reliable results.

Equation (12) is valid for expected values and (13) for estimated totals but not for individual cases. Only if the variances were sufficiently small could those equations be applied to individual cases, and this is the situation presented by the samples. The disturbing effect of the economic situation of both classes, H and L, is too large to be overlooked. The variance which captures this effect should also be incorporated into the billing model.

4. PROPOSAL FOR A STATISTICAL MODEL FOR BILLING INFRACTIONS

The preceding section has shown that numerous factors are involved in the measured utilization factor for any given establishment, whether the customer is a violator or not. An industry or business may have a given installed power which it does not necessarily use constantly. Since the average values do not reflect the particular situations of each user, if what is sought is a billing method valid for all cases, it is necessary to work with the distribution function of the utilization factors, known as $f(\phi)$.

Following a series of analyses on the pilot surveys, one arrives at the conclusion that it is possible to reasonably represent $f(\phi)$ by means of technical functions other than normal distribution. A series of these functions was reviewed, and it was concluded that the GAMMA FUNCTION and the BETA FUNCTION offer interesting properties, similar to those offered by the utilization factor; for example, both make it possible to detect biases to the left or right. Both pilot surveys show notable biases to the left, but what does a bias to the left in the measured utilization factor mean? It reflects an economic situation which is not very prosperous, where there are many establishments with large installed loads but little use. However, there are some thriving establishments in the midst of an overall poor situation, which is reflected in tails to the right.

The gamma function is expressed by:

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \quad (14)$$

for:

$$\lambda > 0 \\ r > 0$$

and,

$$\mu = \frac{r}{\lambda} ; \quad \sigma^2 = \frac{r}{\lambda^2}$$

where μ and σ are the mean and the variance of x , and:

$$\Gamma(r) = \int_0^{\infty} t^{r-1} e^{-t} dt \quad (15)$$

To test the gamma function, for class H in the sample of Hotels, the values for μ and σ^2 can be obtained from Table No. 1.

$$\mu = 0.18473 \quad ; \quad \sigma^2 = 0.02650$$

So that:

$$\lambda = 6.97165 \quad r = 1.28787 \quad \Gamma(r) = 0.897$$

The function is then adjusted:

$$f(x) = 13.59296 x^{0.28787} e^{-6.97165} \quad (16)$$

The beta function is expressed by:

$$f(x) = \frac{1}{\beta(a,b)} x^{a-1} (1-x)^{b-1} ; \quad \begin{matrix} 0 < x < 1 \\ a > 0 \\ b > 0 \end{matrix} \quad (17)$$

$$\mu = \frac{a}{a+b} \quad ; \quad \sigma^2 = \frac{ab}{(a+b+1)(a+b)^2} \quad ;$$

$$\beta(a,b) = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)}$$

and replacing the values for μ and σ^2 in the parametric equations:

$$a = 0.86513 \quad ; \quad b = 3.81807 \quad ; \quad \frac{1}{\beta(a,b)} = 2.86195$$

So that:

$$f(x) = 2.86195 x^{-0.13487} (1-x)^{2.81807} \quad (18)$$

Table No. 2 compares the values for the theoretical functions (15) and (17) with the experimental findings. To do

Table No. 2

INTEGRAL PROBABILITY FUNCTIONS IN THE
PILOT SURVEY OF HOTELS

POWER FACTOR	GAMMA FUNCTION		BETA FUNCTION		EXPERI- MENTAL	
	$\phi = x$	f(x)	F(x)	f(x)	F(x)	F(ϕ)
0.000	0.000	0.000		0.062	0.000	
0.025	3.949	0.049	4.383	0.118		
0.050	4.049	0.149	3.710	0.219		
0.100	3.489	0.337	2.901	0.384	0.394	0.057
0.200	2.121	0.618	1.896	0.624	0.727	0.109
0.300	1.187	0.783	1.232	0.780	0.818	0.035
0.400	0.642	0.875	0.768	0.800	0.909	0.034
0.500	0.341	0.924	0.446	0.941	0.939	0.015
0.600	0.179	0.950	0.232	0.975	0.970	0.020
0.700	0.093	0.964	0.101	0.992	0.970	0.006
0.800	0.048	0.971	0.032	0.998	1.000	0.029
0.900	0.025	0.974	0.004	1.000	1.000	0.026
1.000	0.013	0.976	0.000		1.000	

this, it is more comfortable to work with the integral probability functions obtained from the pilot survey on hotels, e.g.: the experimental cumulative distribution function of the measured utilization factor, the theoretical functions of gamma and beta, and the differences observed between the experimental function and the gamma function.

According to the Kolmogorov-Smirnov test for a sound fit (consistency) in the gamma function, the maximum difference observed (0.109) is less than the critical value (0.27) for a significance level of 0.01.

$$F(x) = P(X \leq x) = \int_0^x f(x) dx \quad (19)$$

Equation (19) indicates that the probability that x is less than a given value is $F(x)$. Figure No. 1 indicates, for example, that the probability that the measured factor of hotels in the non-violator group is lower than 40%, is 0.875 for the gamma function, 0.817 for the beta function, and 0.909 for the distribution of experimental frequencies.

These theoretical functions may fit in well with the experimental findings. However, with a more careful review, it can be seen that, for the particular play of the parameters under analysis, the beta function does not represent the small values well, since $f(x)$ tends to ∞ when x tends to 0.

To test the soundness of the gamma function's fit to the experimental findings, the Kolmogorov-Smirnov test is used. The deviations are calculated:

$$D(x) = | F(\phi) - F(x) | \quad (20)$$

and the hypothesis test is formulated:

$$H^0 : F(x) = F(\phi) \quad (21)$$

The maximum deviation observed is $D(0.20) = 0.109$, which is smaller than the critical value $D_{cr} = 0.27$ found in the fit tables for a significance level of 1%. Consequently, the hypothesis that gamma offers a good representation of the experimental curve well may not be discounted.

On the basis of the aforesaid, and within the limitations of the pilot survey, it is accepted that equation (15)

is the curve for population probability density for non-violator hotels.

It is then assumed that the corresponding density curve for class L is also a gamma function displaced to the left and expressed by:

$$\mu = 0.15810 \quad ; \quad \sigma^2 = 0.02053$$

$$\lambda = 7.70232 \quad ; \quad r = 1.21774 \quad ; \quad \Gamma(r) = 0.914$$

$$g(x) = 13.14399 x^{0.21774} e^{-7.70232x} \quad (22)$$

The results shown in equations (15) and (21) make it possible to propose a mathematical model for billing non-paying energy users, to be adjusted according to the final surveys. The idea is that if all the non-paying energy users in one subpopulation of the AEM were detected, it would be necessary to bill their consumption so as to correct for the displacement to the left in the distribution function. The amount to be billed would then be measured by the displacement, as illustrated in Figure No. 1. Such a procedure would be equitable, since it would penalize each one according to "the place it occupies in the density function", which is a reflection of the good or poor economic situation of the customer. At the end of the procedure, the EEEB would recover the losses according to equations (12) and (13), since the means for the measured utilization factors which appear therein are the same ones used to calculate the parameters of the gamma function; and the respective variances are also the variances of this distribution.

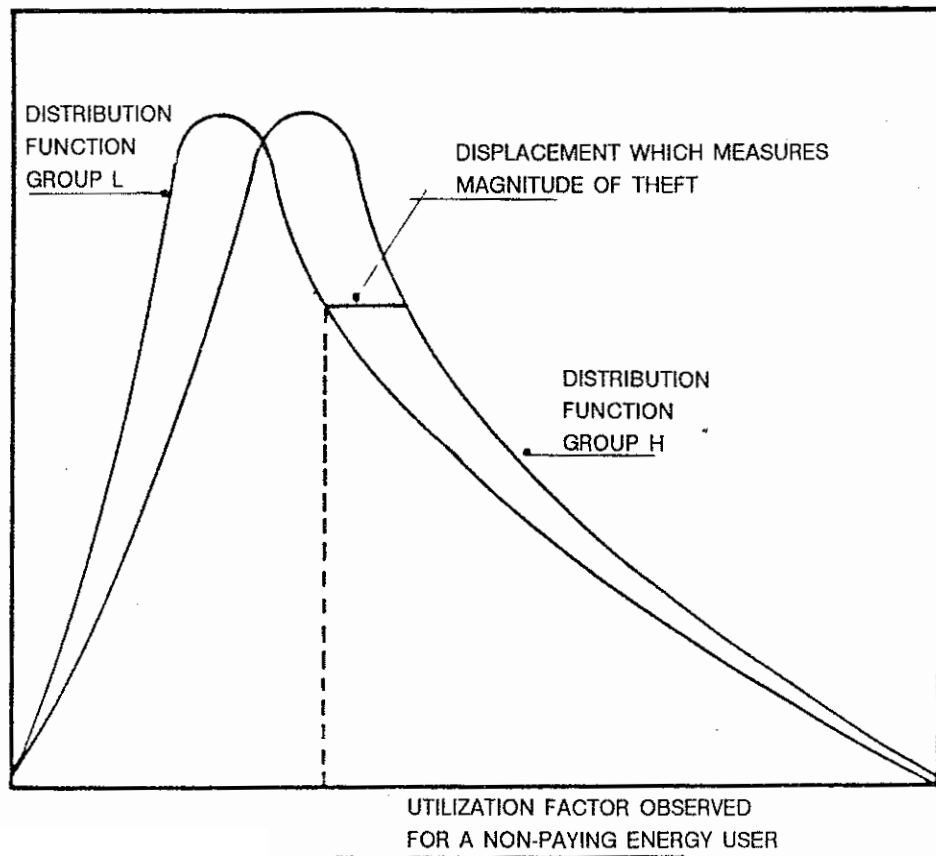
For each utilization factor observed in class L, an effort is made to find an increase δ such that, when added to this factor it transforms the distribution factor for the non-paying energy users into the one for the paying energy users. The mathematical procedure for finding δ as a function of x is quite complicated since it involves working with the integral probability functions. The equilibrium equation which makes it possible to estimate the increment δ is as follows:

$$\int_0^x \Gamma(r_L) x^{r_L-1} e^{-\lambda_L x} dx = \int_0^{x+\delta} \Gamma(r_H) x^{r_H-1} e^{-\lambda_H x} dx \quad (23)$$

Replacing the parameters of (15) in the first member and those of (22) in the second member of equation (23), the ratio can be obtained, but it must be solved through approximations.

Figure No. 1

PROPOSED BILLING PROCEDURE, GEARED TO RECOVERING
LOSSES DUE TO ENERGY THEFT



5. PROGRAM OF FINAL SURVEYS

On the basis of the foregoing, it is necessary to introduce modifications into the original survey proposal, since the plan was designed under the assumption that, for the purpose of the study, it would be sufficient to find interval estimators. The pilot survey demonstrated the need to work with complete distribution functions and to apply non-parametric statistical tests; this calls for increasing sample size over the size originally foreseen.

The new sample sizes will be determined by two conditions:

- a. The variation coefficient for the classes' proportional estimator, expressed by:

$$CV = \sqrt{\frac{N - n}{N} \frac{1 - \hat{p}(L)}{(n - 1) \hat{p}(L)}} \quad (24)$$

Once this coefficient (known also as relative sampling error) has been set, certain sample sizes will be obtained. Since the sample is without replacement, the geometric distribution must be applied to calculate the reliability intervals; a binomial approximation or even normal function could be used. In any case, equation (24) indicates that the smaller the proportion estimator, the larger the sample size for a pre-set error. However, this can only be known after sampling.

- b. The fitness parameter of Kolmogorov-Smirnov, for sizes larger than 35 and significance levels of 1%, is expressed by:

$$N(D) = \frac{1.63}{\sqrt{n}} \quad (25)$$

where:

D is the deviation defined by equation (20).

It would be extremely difficult to determine a priori the sample sizes required to verify the two conditions with only the information from the pilot survey. It is more convenient to apply a RECURRENT EXPERIMENTAL SAMPLING, which consists of choos-

ing an initial, not very large, sample by drawing lots in each stratum. The samples would then be analyzed and the adjustments to conditions (24) and (25) examined. If either or both of these conditions did not verify the reliability requirements, the drawing would continue and the size would increase. With this method, the sample sizes are open and can only be known when the samples are judged to be satisfactory.

The study rests on the fact that inequality (7) is verified, and it is worthwhile to recall that the assessment makes it possible to measure the displacement to the left of the violators' distribution function and, on that basis, to mount a mathematical mechanism for billing for that displacement. This means that, no matter how large, any disturbing effect would affect both groups H and L.

This fact was observed in both the pilot surveys and is therefore accepted as a working assumption. However, if one of the subpopulations should not respond to expression (7), the model proposed for it would no longer function; other analytical tools would have to be used, since the disturbing effects (which reflect a good or poor economic situation) would be correlated to the variables for class H or L. The analysis would become more complicated and would possibly exceed the scope of this study, since, instead of working with one-variable density functions such as those discussed under section 3, bivariate functions would have to be considered; for example, the joint distribution of the utilization factor and the economic situation of the business.

6. FINDINGS

6.1 Analysis of the Samples

The proposed sampling was done in the following groups:

TXT	Textiles, Leather and Footgear
QMC	Chemicals
PVC	Stone, Glass and Ceramics
HTL	Hotels
RST	Restaurants
LVN	Laundries

Many of the planned interviews could not be held, although the visits were actually made. Among the reasons for this, it is worthwhile to mention: i) non-existent addresses; ii) addresses not corresponding to the activity specified in the ISIC code, i.e. errors were detected in the AEM designations; iii) users'

refusal to permit the assessment; iv) users which permitted the assessment but did not authorize the EEEB teams to examine the meters; and finally v), as always occurs in efforts of this kind, rejection of a certain number of surveys due to inconsistent data.

Many of the interviewed units had to be visited two or three times in order to solve the problems of under-representativity in the assessment.

A stratified method was used, dividing the AEM into four strata:

- N1 The largest consumers, which account for over 60% of total consumption.
- N2 The intermediate consumers, which account for 20 or 30%.
- N3 Small consumers, which account for less than 10%.
- N A fairly numerous "tail" of very small consumers having a negligible consumption. This group was not sampled.

The population breakdown of the strata is detailed in Table No. 3.

A first inspection of the samples reveals that the granted load may in some cases differ significantly from the measured load, although there is no definite pattern for the sign of the difference. In other cases, however, there is a notable coincidence between the two values. It can thus be concluded that the granted load is not always a reliable figure; among other reasons, because the installed power is not brought up-to-date.

The utilization factors of the samples behave as foreseen in the pilot phase, $\bar{\Phi}_L$ being lower than $\bar{\Phi}_H$ and its respective deviations of the same magnitude as the means. Again, it can be seen that the great deal of scattering is a characteristic of the utilization factors and that the mean values are not sufficient to characterize all of the users in an ISIC grouping.

The means for the last billed consumption and the average consumption-- for the non-violator group and for the non-paying energy users-- coincide closely, and the differences between the factors $\bar{\Phi}_L$ and $\bar{\Phi}_H$ are maintained, whether they have been calculated using last or average consumption. For the purposes of this study, the factors referring to the last billed consumption will be used.

The means for the contractual and measured loads show differences due to the fact that the former, as mentioned previously, do not always really reflect the users' installed power.

The situation of the contractual load and the last consumption of the users in the universe is summarized in Table No. 4.

Prior to calculating estimators, it is very important to test the influence of stratification on the fact that a user is or is not a violator; in other words, the correlation between the estimators for classes H and L and the position each user occupies in the AEM, ordered according to decreasing consumption. One arrives at the conclusion that, somewhat surprisingly, the correlation is null and that whether a user consumes 10 000 or 100 kWh per bimester has nothing to do with whether he is or is not a violator. The hypothesis of non-correlation had been formulated for the sake of simplification in the pilot phase and has been widely verified in all of the subsamples.

This means that the stratification adopted, although useful because it guarantees good representativity at all levels of consumption, does not afford a sounder estimate, so the statistical inference may be made independently of the stratification criterion. Consequently, equations (8) and (9), as well as those deduced therefrom, prove to be valid and may be used for the estimate.

6.2 Estimation of the Magnitude of Non-Paying Energy Users

For each branch, the proportion \hat{p} and the total are determined for class L; these yield the percentage and the number of non-paying energy users, respectively. If in a sample size n there are l non-paying energy users, the probability distribution of l follows a binomial law.

$$P(l) = \frac{n!}{l! (n-l)!} p^l (1-p)^{n-l} \quad (26)$$

The variances in the proportion and the total are estimated as follows:

$$\hat{V}(\hat{p}) = \frac{N-n}{N(n-1)} \hat{p}(1-\hat{p}) \quad ; \quad \hat{V}(L) = N^2 \hat{V}(\hat{p}) \quad (27)$$

Table No. 3

STRATA IN THE UNIVERSE

SECTOR	N1	N2	N3	-N	N
TXT	106	636	714	417	1 873
QMC	106	424	173	137	840
PVC	159	106	122	0	387
HTL	212	212	220	0	644
RST	4 374	0	0	1 519	5 893
LVN	544	0	0	131	675

Table No. 4

GRANTED LOAD AND MEASURED CONSUMPTION IN THE
GROUPS SURVEYED IN THE UNIVERSE

September 1985

SECTOR	Total Granted Load kW	Total Last Consumption kWh/bimes.	Number of Users	Average Last Consumption kWh/bimes.
TXT	51 502.10	6 652 445	1 873	3 532.00
QMC	45 396.20	9 269 119	840	11 035.00
PVC	24 966.50	3 094 697	387	7 996.00
HTL	8 736.50	1 313 790	644	2 040.00
RST	44 215.40	7 779 769	5 893	1 320.00
LVN		1 059 967	675	1 370.00

To calculate the reliability limits, and for relatively large sample sizes such as in this survey, the normal approximation for the binomial distribution may be used by adding a term of correction for continuity, with which the reliability interval becomes symmetrical. If z is the value of the normal probability density for the chosen level of reliability, this interval comes from the following formula:

$$IC = \pm z \sqrt{V(\hat{p})} + \frac{1}{2n} \quad (28)$$

Table No. 5 shows that the proportion of non-paying energy users varies a great deal according to economic activity as expressed by the ISIC code: from a minimum of 10% in the Stone, Glass and Ceramics Sector, up to 37% in Hotels. The reliability levels are also in line with the sample fraction used and are considered reasonable.

According to the results, the non-paying energy users constitute around 25% of the groups of users studied. This confirms the assumption of the pilot survey in the sense that infractions would acquire extremely high levels. Of the 10 312 users studied, 2 522 are non-paying energy users, and this group total occurs with an interval of ± 77 , for a significance level of 90%, thus being a very reliable measure. Obviously, the branch of economic activity in which these non-paying energy users are found cannot be affirmed with the same reliability. However, it is worthwhile to note that if a systematic search is done in all the branches, this procedure is benefitted by the high degree of probability that the total number will be on the order of 2 500. The high degree of confidence with which it may be affirmed that if a violator is not in one branch he will necessarily be in another supports the joint search, since in the end the firm seeks non-paying energy users that will allow it to recover its losses, no matter what group they are in. The group estimators are very useful for formulating lines of action, after which work should be done simultaneously in all of them.

The average infraction by branches is extremely variable (Table No. 6) and leads to curious conclusions. The group having the fewest non-paying energy users is the Stone, Glass and Ceramics Sector, but, at the same time, it has the most illicit use: since 45% of the consumption is not registered by the meters. The group having the most non-paying energy users is Hotels, but the magnitude of their infraction is a discreet 17% of the actual consumption. The sectors Chemicals and Restaurants have many non-paying energy users with an appreciable magnitude of illicit use, but the Laundry Sector is the critical case; it has numerous non-paying energy users with an illegal level of utilization of 60%, which is the highest in the sample.

Table No. 5

STATISTICS ON THE PROPORTION AND TOTAL OF
NON-PAYING ENERGY USERS BY SECTOR

SECTOR	Proportion Non-Paying Users	Standard Error	Confidence Interval 90%	Total Non-Paying Group	Confidence Interval 90%
TXT	17.04	3.93	6.57	319	120
QMC	24.19	5.27	8.91	203	72
PVC	10.42	4.17	7.23	40	26
HTL	37.50	5.78	9.92	241	61
RST	25.48	2.97	4.95	1 501	287
LVN	32.26	8.32	13.67	218	92
TOTAL	24.47	1.87	3.07	2 522	77

Table No. 6

MEASURED LOAD AND CONSUMPTION FOR PAYING AND
NON-PAYING ENERGY USERS IN THE UNIVERSE

SECTOR	Measured Average PAYING Consumption in kWh per bimester	Measured Average NON-PAYING Consumption in kWh per bimester	Measured Average Load in kW	Average Amt. of Theft in kWh per bimester
TXT	3 600.47	3 315.59	13.72	284.88 (8%)
QMC	11 569.73	9 353.18	38.73	2212.82 (20%)
PVC	8 354.90	4 908.74	37.30	3443.16 (43%)
HTL	2 177.48	1 811.56	8.86	365.92 (17%)
RST	1 414.90	1 044.28	4.05	370.62 (26%)
LVN	2 933.76	1 153.27	10.05	1780.49 (60%)

Table No. 7 indicates the final solution to the problem of estimating infractions, by indicating the consumption recorded in the AEM lists and the totals used illegally as estimated on the basis of the assessment survey. By adding the infractions to the records, real consumption is obtained. It can be seen that infractions are 5.45% of real consumption, and this is greater than what is billed to all of the Hotels or all of the Laundries put together.

By comparing the situation among the branches of activity, it can be seen that the largest thefts are located in the Laundry Sector and the lowest in the Textiles, Leather and Footgear Sector. The other groups have levels of illegal use of between 4 and 7%.

According to the theory implicit in the sample expansion, the highest levels of infraction occur when two conditions are fulfilled: i) a high proportion of non-paying energy users and ii) a large difference between the utilization factors for H and L. When only the first holds true, there will be "many non-paying energy users who make illegal use of little". If only the second holds true, there will be "a few making illegal use of much". Since there is a little bit of everything in the AEM, the extreme case of "many making illegal use of much" is the case of the Laundries.

The non-paying energy users in the Laundry Sector and the Stone, Glass and Ceramics Sector have especially low utilization factors, and it is possible that this is due to the fact that they are sectors which make intensive use of electricity for thermal purposes.

6.3 Model for Recovering Losses

The model discussed under section 3 is applied here to represent the utilization factor for the paying energy users and the non-paying energy users. To do this, the gamma function of equation (14) is used; for the sake of simplicity, the probability functions for the paying energy users are expressed as $f(x)$ and $F(x)$ and as $g(x)$ and $G(x)$ for the non-paying energy users.

The model which offers the best fit corresponds to a truncated gamma function; it also fulfills the condition that the probability density function for the non-paying energy users originates with the displacement of the function for the paying energy users to the left, with the same scattering. When this procedure is applied, it is possible for the density at the origin to become infinite; the model is therefore termed "infinite gamma".

Table No. 7

MEASURED LOAD, THEFT AND ACTUAL CONSUMPTION

SECTOR	Measured Load kWh/bimester	Theft kWh/bimester	Actual Consumption kWh/bimester	Percentage of Theft
TXT	6 652 445	90 922	6 743 367	1.35
QMC	9 269 119	449 636	9 718 755	4.63
PVC	3 094 697	138 847	3 233 544	4.29
HTL	1 313 790	88 370	1 402 160	6.30
RST	7 779 769	556 499	8 336 268	6.68
LVN	1 059 967	387 711	1 447 678	26.78
ALL	29 169 787	1 711 985	30 881 772	5.54

Figure No. 2 presents two examples of sound fit between the infinite gamma model and the experimental data, for the Restaurants Sector and the Textiles Sector. The application of the Kolmogorov-Smirnov test expressed by equations (20), (21) and (25) leads to acceptance of the hypothesis that the model is a good representation of the experimental data.

Table No. 8 presents the parameters for gauging the model for each branch. It can be proven that the adopted model works well in all of the branches studied. However, the parameters vary a great deal from one branch to another, so it is advisable not to extrapolate to groups in which no surveys have been done, since the truncated gamma distribution function should be gauged according to a set of experimental data of relatively homogeneous aggregates. Undoubtedly, these parameters depend heavily on energy consumption technologies. Table No. 9 indicates the values for the fit functions and shows the differences by branches. It can be seen, for example, that for factors less than or equal to 10%, the distribution function for the non-paying energy users covers 38% in TXT but only 24% in RST.

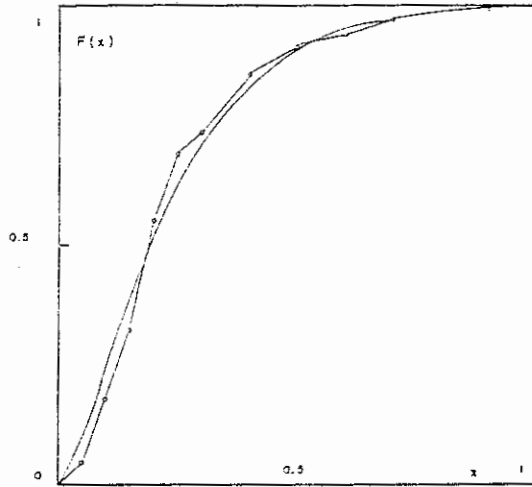
Figure No. 3 provides the graphs for $F(x)$ and $G(x)$ for the six groups under study. The displacement of $G(x)$ to the left with respect to $F(x)$ measures the infraction. In terms of the utilization factor, it is designated by δ , which is a function of x whose expectation is:

$$E(\delta) = \int_0^1 \delta(x)g(x)dx = \bar{\phi}_H - \bar{\phi}_L \quad (29)$$

Equation (29) expresses that the displacement is such that, on average, it should produce the averaged differences between the utilization factors for the paying energy users and the non-paying energy users. To find the value of δ for each point, equation (23) must be solved. The calculation procedure which begins with the gauging of the probability density functions and ends with the calculation of displacements entails major practical difficulties, even by computer, because it is necessary to do recurrent iterations on functions which either have infinite values or represent very sharp curves with abrupt changes in slope. The solution can only be approximate, and it is useful to first do a rough calibration to learn about properties. The results provided in Figure No. 4 represent a family of curves for as δ a function of x : the ordinates show the amount that the utilization factor for a violator belonging to a given branch must be increased by so that, when this is done for all of the non-paying energy users in that branch, there will be a density function shift to the right, as a consequence of which the losses will be recovered.

Figure No. 2

FIT OF THE TRUNCATED GAMMA FUNCTION TO THE
EXPERIMENTAL DATA FROM THE SAMPLE OF RESTAURANTS



FIT OF THE TRUNCATED GAMMA FUNCTION TO THE
EXPERIMENTAL DATA FROM THE SAMPLE OF TEXTILES

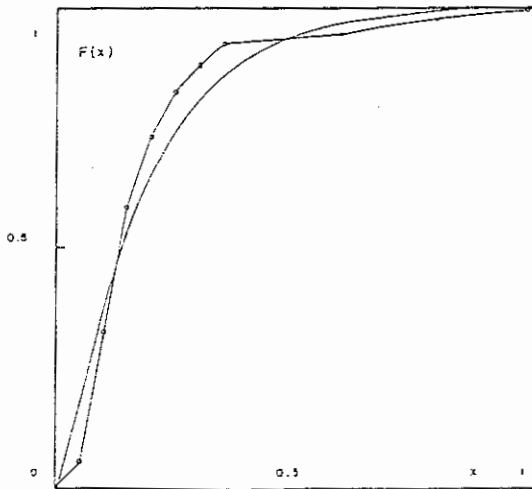


Table No. 8

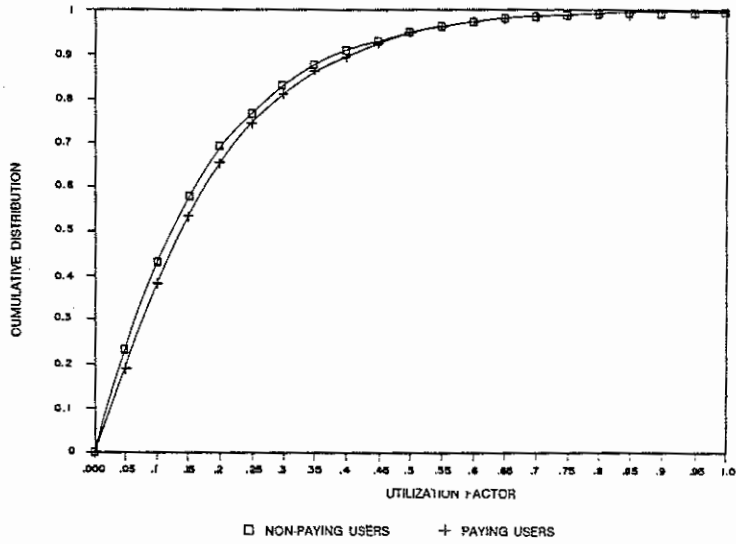
INFINITE GAMMA MODEL

(Parameters)

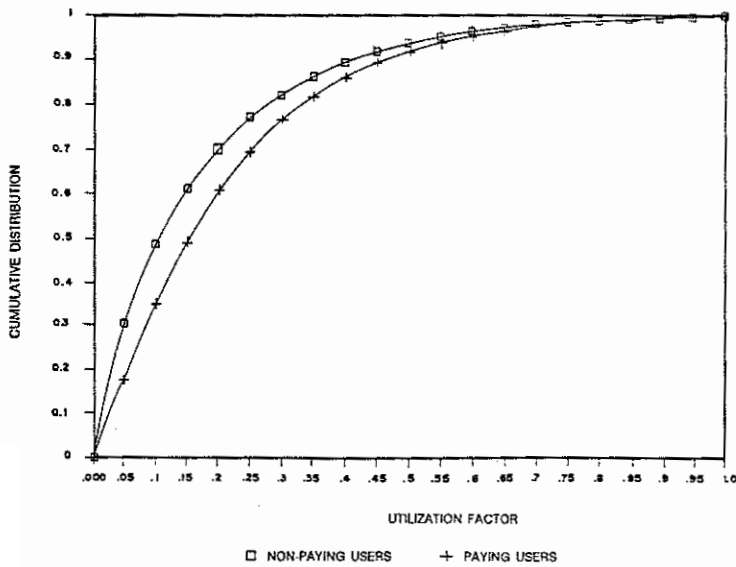
	TXT	QMC	PVC	HTL	RST	LVN
MU:L	16.782	16.776	9.139	14.199	17.906	7.969
MU:H	18.224	20.745	15.555	17.067	24.261	12.216
SD:L	11.355	16.506	5.537	13.449	17.564	3.862
SD:H	16.766	19.617	13.415	15.735	19.116	5.955
r:L	1.0019	0.7313	0.4641	0.8143	0.8774	1.7908
λ :L	5.9701	4.3594	5.0783	5.7349	4.9001	22.4719
$\Gamma(r)$:L	1.0000	1.2507	1.9082	1.1471	1.0885	0.9288
r:H	1.1815	1.1183	1.3445	1.1765	1.6107	4.2082
λ :H	6.4831	5.3907	8.6435	6.8932	6.6392	34.4481
$\Gamma(r)$:H	0.9237	0.9436	0.8922	0.9237	0.8947	7.8365
$\lambda^r/\Gamma(r)$:L	5.9905	2.3468	1.1141	3.6146	3.7048	283.5259
r-1:L	0.0019	-0.2687	-0.5359	-0.1857	-0.1226	0.7908
$-\lambda$:L	-5.9701	-4.3594	-5.0783	-5.7349	-4.9001	-22.4719
Trunc	0.9723	0.9725	0.9886	0.9881	0.9899	1.0039
$\lambda^r/\Gamma(r)$:H	9.8530	6.9731	20.3657	10.4914	23.5797	375.448
r-1:H	0.1815	0.1183	0.3445	0.1765	0.6107	3.2082
$-\lambda$:H	-6.4831	-5.3907	-8.6435	-6.8932	-6.6392	-34.4481
Trunc	0.9843	0.9808	0.9897	0.9858	0.9897	0.9724

MU Measured Average
SD Measured Deviation

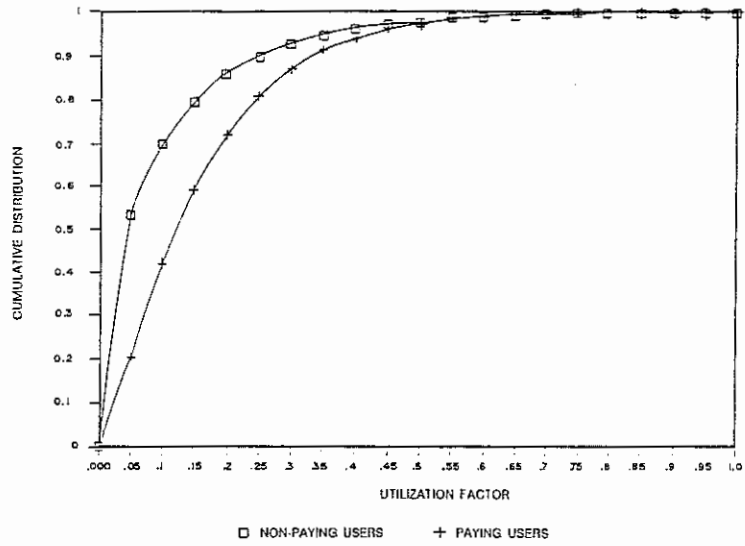
TEXTILES, LEATHER AND FOOTWEAR



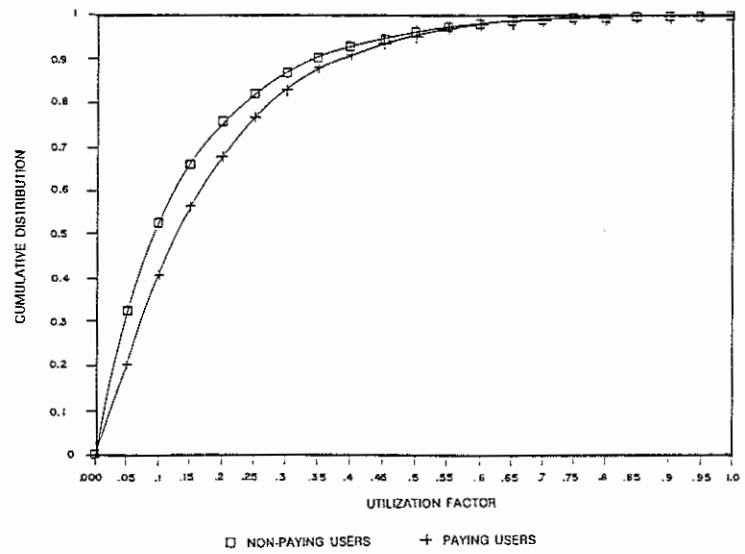
CHEMICALS



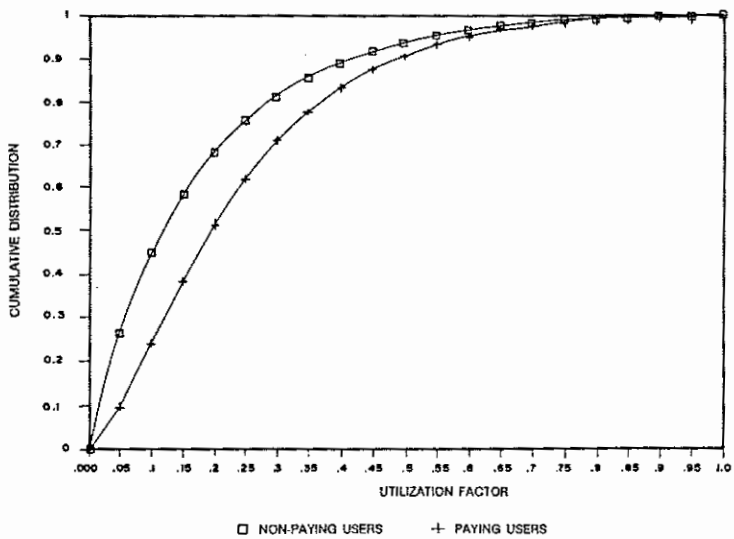
STONE GLASS AND CERAMICS



HOTELS



RESTAURANTES



LAUNDRIES

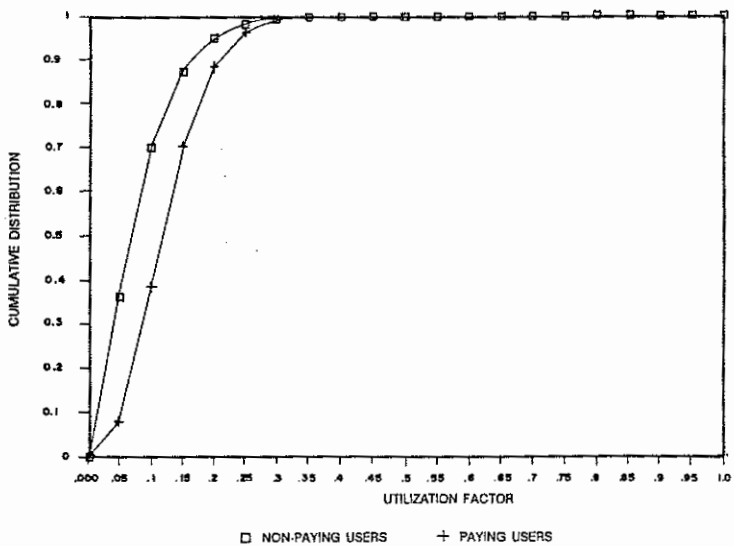


Figure No. 4

DISPLACEMENT OF THE UPDATING FACTOR FOR NON-PAYING ENERGY USERS
IN ORDER TO RECOVER LOSSES DUE TO THEFT

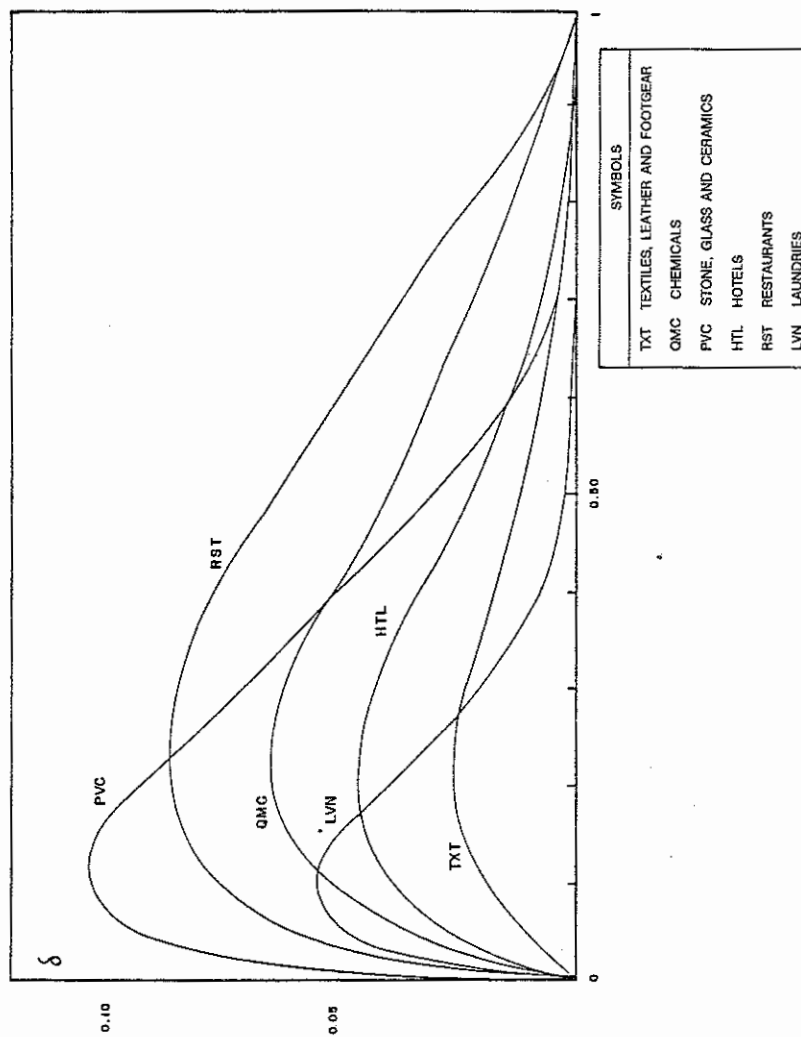


Table No. 9

INFINITE GAMMA MODEL FOR THE DIFFERENT SECTORS

TEXTILES, LEATHER AND FOOTGEAR					CHEMICALS			
x	g(x)	G(x)	f(x)	F(x)	g(x)	G(x)	f(x)	F(x)
0.00	6.0271	0.0000	1.2387	0.0000	53.2028	0.0000	1.8209	0.0000
0.01	5.7535	0.0288	4.0673	0.0203	7.9616	0.0796	3.9069	0.0195
0.02	5.4273	0.0847	4.3229	0.0623	6.3269	0.1511	4.0182	0.0592
0.03	5.1167	0.1374	4.3609	0.1057	5.4318	0.2099	3.9924	0.0992
0.04	4.8228	0.1871	4.3062	0.1490	4.8134	0.2611	3.9159	0.1388
0.05	4.5453	0.2339	4.2027	0.1916	4.3399	0.3068	3.8096	0.1774
0.10	3.3767	0.4320	3.4465	0.3828	2.8969	0.4878	3.1582	0.3516
0.15	2.5072	0.5791	2.6826	0.5360	2.0891	0.6124	2.5305	0.4938
0.20	1.8612	0.6883	2.0439	0.6542	1.5550	0.7035	1.9996	0.6071
0.25	1.3814	0.7694	1.5391	0.7438	1.1777	0.7718	1.5680	0.6963
0.30	1.0253	0.8295	1.1504	0.8110	0.9018	0.8238	1.2236	0.7660
0.35	0.7609	0.8742	0.8555	0.8612	0.6957	0.8638	0.9517	0.8204
0.40	0.5647	0.9073	0.6338	0.8984	0.5398	0.8946	0.7384	0.8627
0.45	0.4190	0.9319	0.4682	0.9260	0.4205	0.9187	0.5719	0.8954
0.50	0.3110	0.9502	0.3451	0.9463	0.3287	0.9374	0.4422	0.9208
0.55	0.2308	0.9637	0.2539	0.9613	0.2577	0.9520	0.3416	0.9404
0.60	0.1712	0.9738	0.1866	0.9723	0.2024	0.9636	0.2636	0.9555
0.65	0.1271	0.9812	0.1369	0.9804	0.1593	0.9726	0.2032	0.9672
0.70	0.0943	0.9867	0.1003	0.9863	0.1256	0.9797	0.1566	0.9762
0.75	0.0700	0.9908	0.0735	0.9906	0.0991	0.9853	0.1206	0.9831
0.80	0.0519	0.9939	0.0537	0.9938	0.0783	0.9898	0.0928	0.9884
0.85	0.0385	0.9962	0.0393	0.9962	0.0620	0.9933	0.0714	0.9926
0.90	0.0286	0.9978	0.0287	0.9979	0.0491	0.9961	0.0549	0.9957
0.95	0.0212	0.9991	0.0210	0.9991	0.0389	0.9938	0.0422	0.9981
1.00	0.0157	1.0000	0.0153	1.0000	0.0309	1.0000	0.0324	1.0000

Table No. 9

(Continued)

x	STONE, GLASS AND CERAMICS				HOTELS			
	g(x)	G(x)	f(x)	F(x)	g(x)	G(x)	f(x)	F(x)
0.00	538.6685	0.0000	0.3899	0.0000	31.0282	0.0000	1.3953	0.0000
0.01	12.6358	0.2527	3.8626	0.0193	8.1238	0.0812	4.4072	0.0220
0.02	8.2838	0.3573	4.4983	0.0611	6.7445	0.1556	4.6489	0.0673
0.03	6.3359	0.4304	4.7443	0.1073	5.9067	0.2188	4.6611	0.1139
0.04	5.1618	0.4879	4.8047	0.1551	5.2873	0.2748	4.5772	0.1601
0.05	4.3532	0.5355	4.7590	0.2029	4.7900	0.3252	4.4439	0.2052
0.10	2.3292	0.7025	3.9222	0.4199	3.1616	0.5240	3.5580	0.4052
0.15	1.4540	0.7971	2.9275	0.5912	2.2012	0.6581	2.7077	0.5619
0.20	0.9668	0.8576	2.0982	0.7168	1.5665	0.7522	2.0182	0.6800
0.25	0.6655	0.8984	1.4706	0.8060	1.1283	0.8196	1.4872	0.7676
0.30	0.4682	0.9268	1.0166	0.8682	0.8188	0.8683	1.0881	0.8320
0.35	0.3344	0.9469	0.6958	0.9110	0.5973	0.9037	0.7921	0.8790
0.40	0.2415	0.9613	0.4729	0.9402	0.4374	0.9296	0.5746	0.9231
0.45	0.1759	0.9717	0.3197	0.9601	0.3213	0.9485	0.4156	0.9379
0.50	0.1290	0.9793	0.2152	0.9734	0.2365	0.9625	0.3000	0.9558
0.55	0.0951	0.9849	0.1443	0.9824	0.1744	0.9727	0.2161	0.9687
0.60	0.0704	0.9890	0.0965	0.9884	0.1288	0.9803	0.1555	0.9780
0.65	0.0523	0.9921	0.0644	0.9925	0.0953	0.9859	0.1117	0.9847
0.70	0.0390	0.9944	0.0429	0.9951	0.0706	0.9901	0.0802	0.9895
0.75	0.0292	0.9961	0.0285	0.9969	0.0523	0.9932	0.0575	0.9929
0.80	0.0218	0.9974	0.0189	0.9981	0.0388	0.9954	0.0412	0.9954
0.85	0.0165	0.9983	0.0125	0.9989	0.0288	0.9971	0.0295	0.9972
0.90	0.0123	0.9991	0.0083	0.9994	0.0214	0.9984	0.0211	0.9984
0.95	0.0093	0.9996	0.0055	0.9998	0.0159	0.9993	0.0151	0.9994
1.00	0.0070	1.0000	0.0036	1.0000	0.0118	1.0000	0.0108	1.0000

Table No. 9

(Continued)

x	RESTAURANTS				LAUNDRIES			
	g(x)	G(x)	f(x)	F(x)	g(x)	G(x)	f(x)	F(x)
0.00	15.3487	0.0000	0.0211	0.0000	0.0314	0.0000	0.0000	0.0000
0.01	6.2668	0.0627	1.3389	0.0067	5.9121	0.0296	0.1049	0.0005
0.02	5.4810	0.1214	1.9132	0.0230	8.1696	0.1000	0.6869	0.0045
0.03	4.9658	0.1736	2.2933	0.0440	8.9920	0.1858	1.7873	0.0169
0.04	4.5645	0.2213	2.5582	0.0682	9.0170	0.2758	3.1872	0.0417
0.05	4.2290	0.2653	2.7434	0.0948	8.5922	0.3639	4.6208	0.0808
0.10	3.0404	0.4470	3.0059	0.2385	4.8326	0.6995	7.6286	0.3870
0.15	2.2643	0.5796	2.7628	0.3827	2.1650	0.8744	5.0043	0.7028
0.20	1.7109	0.6790	2.3631	0.5108	0.8837	0.9506	2.2498	0.8842
0.25	1.3030	0.7543	1.9431	0.6185	0.3427	0.9813	0.8223	0.9610
0.30	0.9973	0.8118	1.5584	0.7060	0.1287	0.9931	0.2636	0.9881
0.35	0.7660	0.8559	1.2286	0.7757	0.0473	0.9975	0.0772	0.9966
0.40	0.5898	0.8898	0.9564	0.8303	0.0171	0.9991	0.0212	0.9991
0.45	0.4550	0.9159	0.7374	0.8727	0.0061	0.9997	0.0055	0.9998
0.50	0.3516	0.9361	0.5643	0.9052	0.0022	0.9999	0.0014	0.9999
0.55	0.2720	0.9517	0.4292	0.9301	0.0008	1.0000	0.0003	1.0000
0.60	0.2106	0.9638	0.3247	0.9489	0.0003	1.0000	0.0001	1.0000
0.65	0.1632	0.9731	0.2447	0.9632	0.0001	1.0000	0.0000	1.0000
0.70	0.1266	0.9804	0.1837	0.9739	0.0000	1.0000	0.0000	1.0000
0.75	0.0983	0.9860	0.1375	0.9819	0.0000	1.0000	0.0000	1.0000
0.80	0.0763	0.9903	0.1026	0.9879	0.0000	1.0000	0.0000	1.0000
0.85	0.0593	0.9937	0.0764	0.9924	0.0000	1.0000	0.0000	1.0000
0.90	0.0461	0.9964	0.0568	0.9957	0.0000	1.0000	0.0000	1.0000
0.95	0.0358	0.9984	0.0421	0.9982	0.0000	1.0000	0.0000	1.0000
1.00	0.0279	1.0000	0.0312	1.0000	0.0000	1.0000	0.0000	1.0000

6.4 Example of Calculation

For a violating restaurant whose measured utilization factor is $x = 0.10$ and whose recorded consumption is $Cr = 850$ kWh/bimester. For this value of x , one reads $\delta = 0.075$ in the corresponding curve. The corrected utilization factor will be equal to $x + \delta = 0.175$. If equation (2) is now applied, with Cr' as the actual consumption:

$$\frac{Cr}{Cr'} = \frac{x}{x + \delta} \quad (30)$$

so that the actual consumption results in $Cr = 1\,487$ kWh/bimester. Thus, a restaurant having these features is illegally using 637 kWh/bimester, which is 42.8% of what it consumes. This is true because, within the statistical approximation of the sound-fit test of Kolmogorov-Smirnov, the non-violating restaurants behave according to the gamma function, with specific parameters for the group of restaurants, because all the violators come from this family and shift to the left in a magnitude δ , which by definition corresponds to the magnitude of their infraction.

If the entire curve for the Restaurant Sector is traced in Figure No. 4 and equation (30) is applied for all of the violators detected in this branch, equation (29) assures that the losses will be recovered during the recovery period, provided that the following condition is met:

$$Cr' - Cr = 14.4 N \delta \bar{W}_a E(\delta) \quad (31)$$

To omit the rough estimate of the graphing procedure, a computer program has been prepared to arrive at a more refined fit, so that, when a value for x is introduced and the appropriate branch is selected, the displacement is obtained by means of an iterative procedure.