Teorías de Pronóstico

(Incluye ejercicios)

De TETRAPRO

Editado por Sr. H. Leijon, UIT

UNION INTERNATIONALE DES TELECOMMUNICATIONS
INTERNATIONAL TELECOMMUNICATION UNION
UNION INTERNACIONAL DE TELECOMUNICACIONES

Teorías de Pronóstico

1. SUPUESTOS GENERALES PARA EL PRONOSTICO

1.1 Por qué pronosticar?

1.1.1 *Densidad telefónica y desarrollo económico*

Las facilidades de telecomunicaciones en un país dependen de su nivel económico. Para desarrollar la economía, son necesarias ciertas facilidades de telecomunicaciones básicas. Esto es válido para todos los países. Existe una correlación entre el número de teléfonos por habitante y el producto bruto interno per capita, tal como se muestra en la figura 1.1/1. Ocurren ciertas desviaciones debido a diferencias en la estructura económica de los diferentes países. Es fácil comprender que un país cuya economía está basada principalmente en la agricultura necesite menos teléfonos por habitante que un país altamente industrializado.

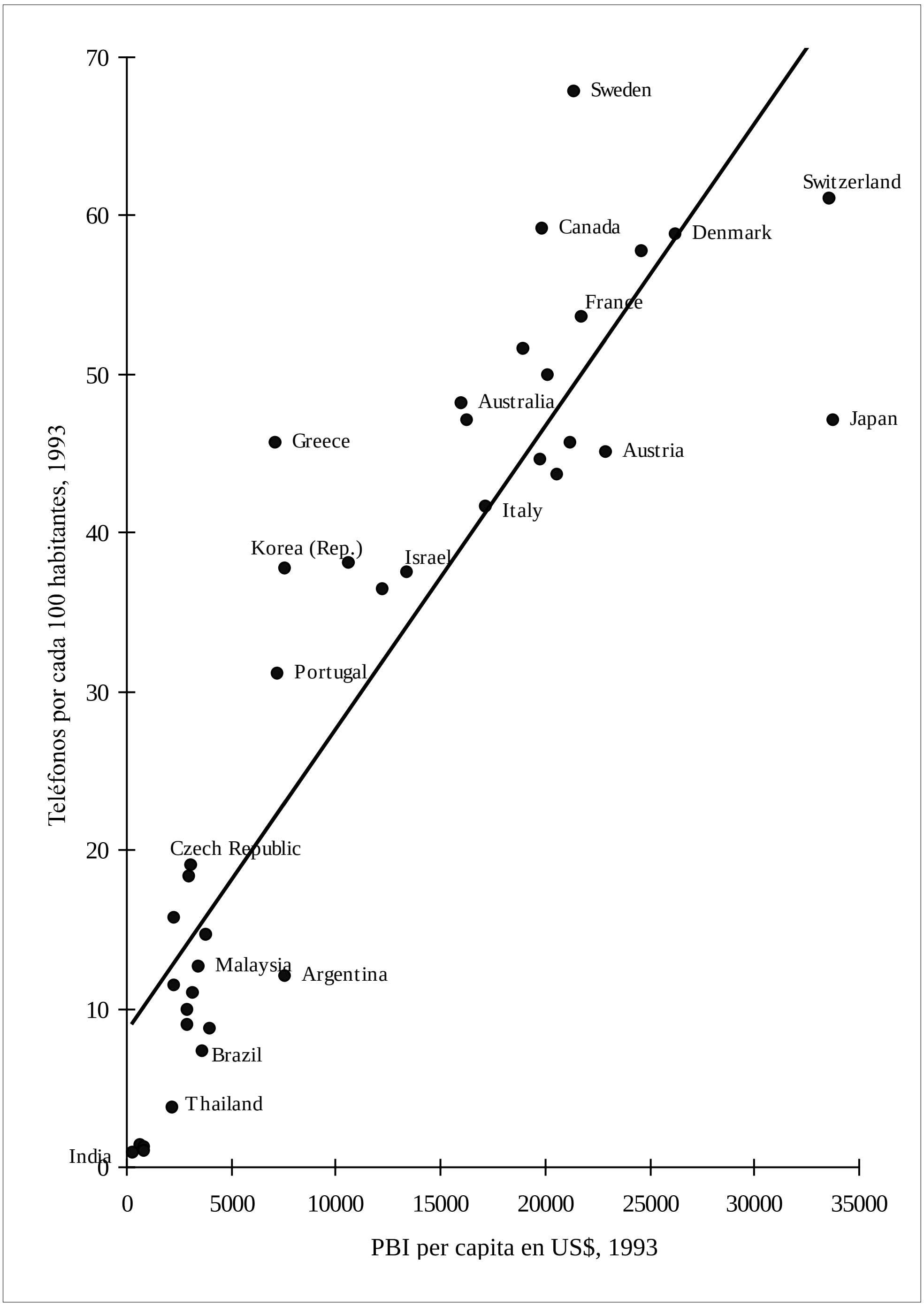


Figura 1.1/1 : Densidad telefónica y Producto Bruto Interno (PBI) per capita en algunos países, 1993

La demanda de teléfonos depende primariamente de la especialización ocupacional de la población. La necesidad de teléfonos es baja en una aldea autosuficiente, pero se incrementa en la medida que la gente deja de lado el sector agrícola por otras actividades (industria, comercio, servicios públicos, etc.). Las actividades gubernamentales y las privadas comerciales requieren de las telecomunicaciones para poder funcionar. La concentración de cada vez más gente en áreas urbanas puede también dar paso a una demanda mayor por teléfonos residenciales, pero esto solamente para aquellas personas que han alcanzado un nivel de vida suficientemente alto. Una cierta demanda para comunicación con las áreas rurales se derivará también de la urbanización. La demanda es, claro está, también dependiente de las preferencias y hábitos nacionales, así como también de las tarifas telefónicas. Para evitar la influencia tarifaría, toda investigación y pronóstico deben hacerse sobre el supuesto de que las tarifas telefónicas no son prohibitivas, sino suficientes para permitir un mantenimiento adecuado y una expansión del sonido de la red telefónica.

Para estimar la futura demanda de telecomunicaciones, se deben tomar en cuenta el desarrollo general esperado y, especialmente, el desarrollo económico. Son de utilidad datos publicados sobre la situación económica actual e información sobre planes de desarrollo económico.

1.1.2 *Propósito de los pronósticos*

La ampliación de los servicios telefónicos requiere la provisión de aparatos telefónicos, planta de línea de abonados, equipo de central, circuitos de empalme y troncales. Sin embargo, siempre habrá un cierto desfase de tiempo entre la identificación de una necesidad - o más bien una futura necesidad- y el momento en el cual puede ser satisfecha. El tiempo transcurrido entre la identificación y la satisfacción de la necesidad es considerable. Para evitar períodos de espera largos y alta congestión, es conveniente prever las necesidades con suficiente anticipación. Esto hace posible ampliar la planta en el momento oportuno porque las acciones necesarias pueden tomarse en el momento apropiado.

Un pronóstico, por tanto, producirá en primer lugar estimados precisos de la demanda futura de facilidades de telecomunicación.

1.1.3 *Pronóstico, planificación y programa de trabajo*

Los pronósticos proporcionan la base de los planes. Los planes son considerados por la administración que toma las decisiones. Conforme a las decisiones se formula un programa de trabajo. El programa de trabajo requiere planificación detallada de lo que se necesita hacer hasta que todo el equipamiento esté operativo. Hay planificación a largo, mediano y corto plazo, cada una de las cuales cuenta con sus propios requerimientos en cuanto a los detalles específicos en cada punto, de acuerdo a un cronograma de trabajo. Cada tipo de planificación debe usar, más o menos, pronósticos detallados de las cantidades requeridas.

Utilizaremos las siguientes definiciones:

Pronóstico	Un pronóstico es una predicción de la demanda futura, generalmente expresada en cantidades.
Plan	Un plan es una propuesta de acción y estimación del costo de llevar a cabo el plan. Puede incluir líneas alternativas de acción.
Decisión	La administración decide qué acción se tomará. Aprueba el plan que después se convierte en un programa de trabajo.
Programa de trabajo	Un programa de trabajo se basa en un plan aprobado y define las acciones a tomarse así como sus plazos.
Planificación	Un programa de trabajo generalmente requiere planificación detallada de todas las acciones a tomar para llevar a cabo el programa de trabajo. El plazo para cada acción es de suma importancia. Planes de trabajo más o menos detallados deben estar listos en un tiempo dado.

La relación entre los conceptos arriba mencionados se ilustra en la Figura 1.1/2.

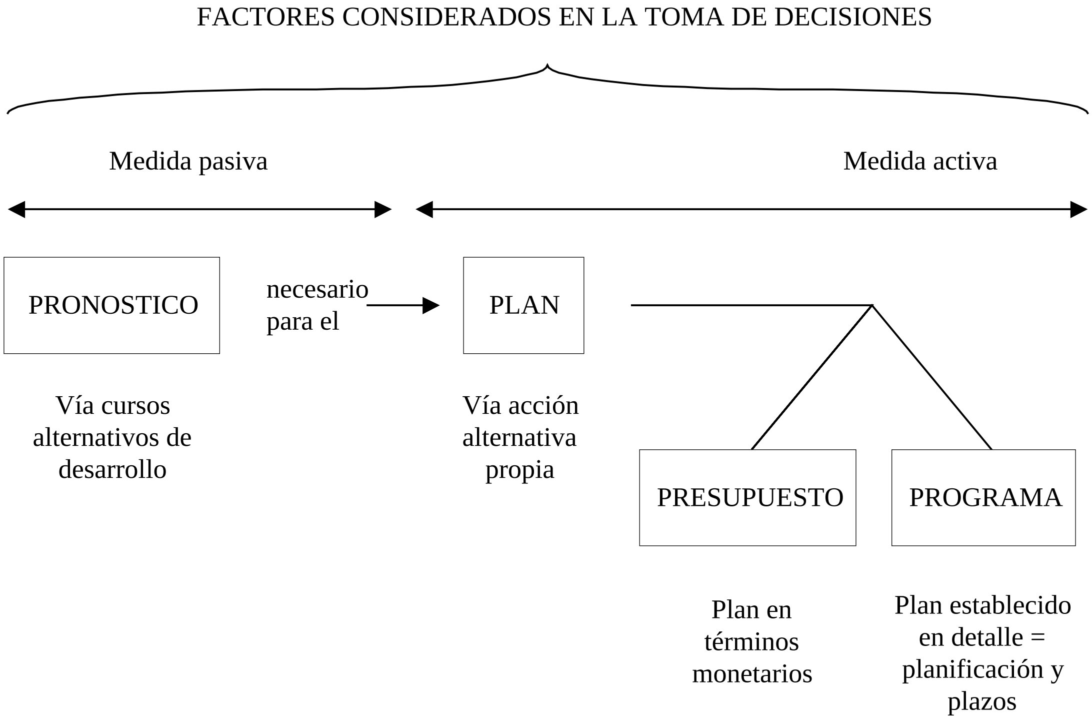


Figure 1.1/2 : Relación entre el pronóstico y el plan

Debido a que el programa de trabajo e inversiones se basan en pronósticos, es muy importante que estos últimos sean tan precisos como sea posible. Es también deseable que el planificador defina el grado de incertidumbre del pronóstico. Esto permite al planificador saber cuándo puede ser flexible.

La situación en muchos países en desarrollo puede ser tal que la demanda es mucho mayor que la oferta. En tal caso, prácticamente todas las ampliaciones se utilizarán de inmediato, sin disminuir esencialmente la brecha entre la oferta y la demanda. Prever la demanda en tales casos es muy difícil, debido a que la demanda existente no puede medirse o estimarse fácilmente. Es también tarea del planificador estimar esa demanda para que así la persona que toma las decisiones conozca el efecto de sus acciones sobre la situación de la oferta y la demanda. El tratar tales situaciones es también un asunto dependiente de la política que va a seguir la administración. Se le dará prioridad a los abonados de negocios, al desarrollo de áreas de centrales locales o a los servicios de larga distancia, etc.?

1.2 Cuándo se necesita un pronóstico?

Los períodos de pronóstico dependen de los períodos de planificación. Ellos, a su vez, dependen de los tiempos de entrega de los distintos tipos de equipo necesario. Damos algunos ejemplos a continuación:

Aparatos de abonados	1-2 años
Equipo de conmutación de centrales	3-4 años
Cables para abonados	6-10 años
Ductos en redes locales	10-15 años
Edificios	10-20 años
Terrenos para los edificios	hasta 50 años

La planificación así como los pronósticos se dividen por lo general en:

Planificación/pronóstico de largo plazo	más de 10-15 años
Planificación/pronóstico de mediano plazo	5-10 años
Planificación/pronóstico de corto plazo	menos de 5 años

De lo anterior se deriva que los varios tipos de pronósticos y planificación concernientes a diferentes tipos de equipo dependen de cuándo tendrá que usarse este equipo. Es evidente que un pronóstico de corto plazo debe ser más detallado, ya que sirve para la planificación a corto plazo en la cual debe especificarse cada detalle de la planta.

Ningún método de pronóstico será adecuado para todos los plazos de tiempo. Deben usarse diferentes métodos para diferentes períodos y esto hace surgir problemas de ajuste cuando los pronósticos para períodos de tiempo, ya sea traslapados o no, provienen de métodos diferentes, ya que pueden ocurrir discrepancias entre las curvas de crecimiento.

No existe procedimiento estándar para ajustar los pronósticos contradictorios. En la mayoría de los casos se emplea el juicio basado en la experiencia y se dibuja una curva que salva la inconsistencia entre los pronósticos. Esto se muestra en la Figura 1.3/1.

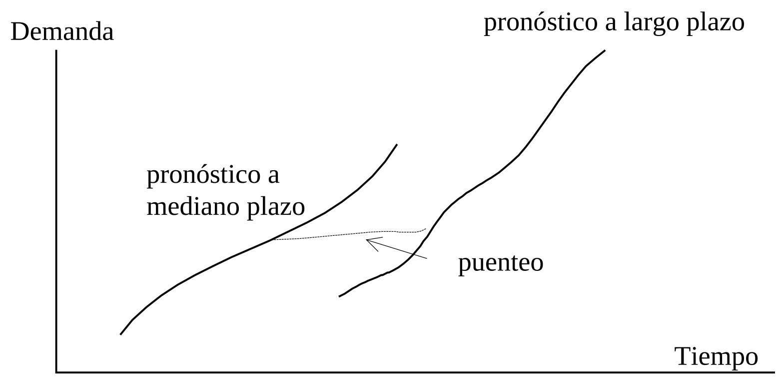


Figura 1.3/1 : Ajuste de pronósticos contradictorios

1.3 Qué se va a pronosticar?

1.3.1 *Tópicos de pronóstico*

La previsión concierne a cuatro áreas principales:

Ubicación	Terreno Edificios
Líneas	Líneas directas de central de abonado (pronóstico de la red de abonado) Red de empalme Red troncal
Equipo	Equipo de abonado Equipo de conmutación automático Equipo de conmutación manual
Numeración	Esquemas de numeración de centrales (estrategia a largo plazo)

Los pronósticos generalmente presentan el número futuro esperado de líneas directas de central (DEL's) y el tráfico futuro, siendo este último más o menos detallado. Consecuentemente hay:

- pronósticos de abonados;
- pronósticos de tráfico.

1.3.2 *Ambito del pronóstico*

Un pronóstico puede referirse al total de una entidad dada, tal como un país, cierta región de un país, un área metropolitana, etc. En este caso se llama un *pronóstico global* y contiene muy pocos detalles de la entidad. Usualmente es suficiente sólo para propósitos de planificación a largo plazo. Un *pronóstico detallado* proporciona información detallada sobre el desarrollo futuro esperado. Se hace separadamente para cada parte del país donde la división es tal que las unidades se vuelven manejables.

Solamente el pronóstico de abonado detallado especifica la ubicación exacta de cada abonado.

Un pronóstico de tráfico detallado especifica todo el tráfico futuro dentro del área en cuestión. Se basa generalmente en pronóstico de abonados y en mediciones.

1.3.3 *Información necesaria para la planificación*

Dependiendo de la necesidad de planificación, los resultados del pronóstico deben presentarse de diferente manera. Damos los ejemplos siguientes:

Distribución de abonados dentro de un área

- Ubicación exacta de los abonados individuales.
- Número de abonados en cada cuadrícula de un mapa de cuadrícula.
- Número de abonados por área de tráfico en el área en estudio.
- Número de abonados por categoría de abonado en cada área de tráfico.

Tráfico

- Tráfico de origen y de destino por abonado.
- Tráfico de origen y de destino por abonado en cada área de tráfico.
- Tráfico de origen y de destino por abonado por cada categoría de abonado.
- Flujos de tráfico entre áreas.

Podemos definir:

A = tráfico total para un grupo de abonados
N = número de abonados en el grupo
α = tráfico por abonado
y = intensidad de llamadas en el grupo
h = tiempo de ocupación promedio

Entonces, $A = y \cdot h$, pero también $A = N \cdot \alpha$

El tráfico A puede ahora pronosticarse de diferentes maneras:

1. Directamente, por ejemplo, por extrapolación de tendencias de valores de A registrados en los años anteriores.

2. Indirectamente, pronosticando primero N ó α, o alternativamente, y y h.

En lo concerniente al pronóstico del tráfico para un grupo de abonados, generalmente es preferible hacer un pronóstico para cada categoría de abonado antes de estimar el tráfico total. La razón para ello es que las proporciones entre las diferentes categorías pueden variar en el futuro.

1.4 Cómo se realiza el pronóstico?

1.4.1 *Requisitos básicos*

Hay dos requisitos principales para el pronóstico del sonido:

- Debe disponerse de una adecuada provisión de información exacta y relevante del pasado. Esto consistirá generalmente en registros de mediciones en equipo existente (conexiones, llamadas, tráfico, etc…) complementados con información general de base.

- Dicha información se usará sistemáticamente en el pronóstico.

- Debe contarse con una razonable conjetura acerca del desarrollo futuro. Este estimado del desarrollo futuro puede ser una extrapolación del desarrollo pasado, ajustado a veces para tomar en cuenta la información básica disponible. El planificador necesita información histórica precisa para mejorar su pronóstico.

Consecuentemente, la base del pronóstico es el estudio del pasado. Mientras mejor entendamos y describamos matemáticamente el desarrollo pasado, mejores serán nuestras probabilidades de realizar un pronóstico correcto.

También hay que destacar que debe proporcionarse el grado de incertidumbre del pronóstico a fin de que aquéllos que van a utilizar los datos la tomen en cuenta.

1.4.2 *Verificación del pronóstico*

Siempre existe una razón para ser críticos acerca de la viabilidad de un pronóstico. Deben formularse las siguientes preguntas:

- Es el pronóstico relevante (válido)?

- Cuán exacto es el pronóstico?

- Es el resultado creíble?

La relevancia de un pronóstico depende en algunos casos del uso correcto de los supuestos para el desarrollo futuro.

La exactitud de un pronóstico puede depender de la precisión estadística de la información histórica y del método de extrapolación usado. La credibilidad de un pronóstico con frecuencia es un asunto de juicio personal. Aquí la mente humana puede, sin embargo, tener sus limitaciones al imaginar el crecimiento futuro. Así, muchos pronósticos han sido disminuidos de tamaño debido a que el planificador no pudo visualizar tal crecimiento. El desarrollo de muchas redes ha sido entorpecido por la falta de imaginación del planificador.

Cuando se juzga la credibilidad de un pronóstico se debe recordar que:

- Los pronósticos concernientes a la población, la economía y el desarrollo industrial, etc., pueden ser engañosos.

- Las estadísticas de telecomunicaciones disponibles para períodos pasados pueden contener errores y pueden no siempre haberse recogido bajo las mismas condiciones.

- La información disponible sobre el desarrollo histórico puede cubrir sólo un período limitado.

- Las relaciones pasadas y presentes entre las variables en estudio pueden no ser necesariamente verdaderas en el futuro.

1.4.3 *Cómo comenzar?*

El proceso de pronóstico puede dividirse en las siguientes partes:

Definición del problema

Debe determinarse el propósito y los supuestos de los pronósticos.

Recolección de información básica

Se deben investigar varias fuentes para obtener datos básicos. Hay que estudiar el crecimiento poblacional y económico. También son esenciales los resultados de los pronósticos más recientes.

Selección del método de pronóstico

El método debe escogerse de acuerdo a la información disponible y a la exactitud requerida, etc.

Análisis y establecimiento de los pronósticos

El análisis consiste en la preparación metodología de la información básica y en la evaluación de los resultados obtenidos.

Documentación

El pronóstico debe presentarse en un formato de fácil entendimiento. El resultado debe contener pronósticos alternativos. Además del pronóstico más probable, debe existir también un pronóstico optimista y otro pesimista, para indicar a la persona que elaborará el plan dónde pueden encontrarse los limites superior e inferior.

2 METODOS DE PRONOSTICO

2.1 Introducción

Casi todos los métodos de pronóstico asumen que el futuro se asemejará de alguna manera al pasado. Esto se puede interpretar de varias maneras, por ejemplo:

1. Para métodos de pronóstico de tendencias de tiempo se asume que el desarrollo seguirá una curva que se ha sido ajustada a la información histórica existente.

2. Cuando se usan relaciones explícitas entre la demanda y diversos factores determinantes, éstas permanecerán iguales en el futuro.

3. Comparando varios pasos del desarrollo de las telecomunicaciones, se asume que el país (o área) menos desarrollada se desarrollará al nivel de los más desarrollados.

4. Cuando se aplica juicio personal (subjetivo) en el pronóstico, el futuro se asemejará al conocimiento previo personal y a la experiencia obtenida de anteriores desarrollos.

Aun cuando el futuro usualmente se asemeja al pasado en este sentido, nunca es una reproducción exacta. Por tanto, los pronósticos nunca deben aceptarse sin crítica. Esto quiere decir que los pronósticos producidos por ciertos métodos con frecuencia se modifican antes de ser aceptados para propósitos de planificación. Pese a que tales correcciones pueden parecer subjetivas, son necesarias siempre que haya una razón para creer que el futuro mostrará divergencias con el pasado que no fueron tomadas en cuenta por el método de pronóstico utilizado. Ejemplos de tales situaciones son los cambios tarifarios, las mejoras radicales del servicio, etc.

De ello se deduce que los supuestos utilizados en el pronóstico deben ser cuestionados, siempre que puedan ocurrir cambios en el ambiente en el cual opera el sistema telefónico.

Nunca se pone demasiado énfasis en la importancia de información precisa y detallada. La calidad de un pronóstico depende directamente de la calidad de los datos históricos empleados. Si no se dispone de información fiable, el planificador debe primero establecer procesos adecuados de recojo de datos antes de aplicar métodos de pronóstico. Muchos de los métodos discutidos en este capitulo serán difíciles, sino imposibles de utilizar sin los datos históricos adecuados.

2.2 Análisis estadístico de la demanda

Uno puede asumir que el desarrollo del número de líneas de central directas en un país sigue ciertos patrones. Por ejemplo, debiera ser dependiente del número de habitantes, del estándar de vida, del desarrollo tecnológico y económico, etc. Si algunas variables están lógicamente relacionadas con el desarrollo telefónico, esas variables pueden usarse para explicar el desarrollo. A continuación se describen algunas de dichas variables:

Tamaño de la Población - Al crecer la población, el número de consumidores aumenta y la demanda por líneas de abonados se incrementa.

Nivel de Vida.- Una medida del nivel de vida es el Producto Bruto Interno per capita. En las economías planificadas centralmente puede usarse el Producto Material Neto per capita. Ambos valores deben convertirse en un nivel de precios fijo. Cuando el nivel de vida aumenta, generalmente también aumenta la demanda de líneas de abonados.

Actividad de la Construcción.- Puede usarse el número de oficinas nuevas y apartamentos como una medida del cambio de estructura en la sociedad debido a la urbanización, el alto nivel de vida, el desarrollo poblacional, etc. La actividad de la construcción se puede usar para estimar la futura demanda de líneas de abonados.

Precio.- Se asume que un alto precio para el uso de un teléfono (es decir, cuota de inscripción, costo de conexión y servicio medido) reducen la demanda de líneas de abonados.

La elección de las variables explicativas depende de la disponibilidad de pronósticos fiables sobre ellas. Las variables que más comúnmente explican el número de líneas de abonados de un país son población y desarrollo económico.

La Figura 2.2/1 muestra la densidad telefónica y el Producto Bruto Interno per capita de Suecia de 1900 a 1965.

Sobre una escala logarítmica, la relación es aproximadamente lineal durante los períodos 1900-1915 y 1920-1965.

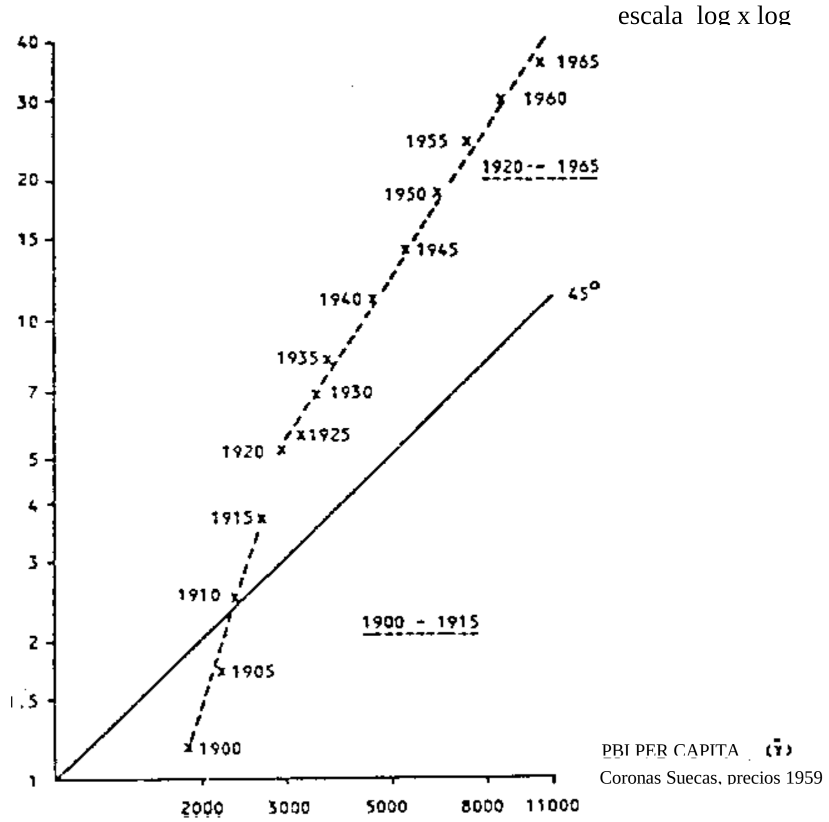


Figura 2.2/1 : Densidad telefónica y PBI per capita en Suecia, 1900-1965

El análisis de la demanda estadística por medio de <u>*modelos econométricos*</u> es una técnica diseñada para obtener relaciones cuantitativas explícitas entre el interés y otras variables que se cree tienen influencia sobre ella.

Un modelo econométrico de demanda es por tanto una representación matemática del comportamiento del cliente y depende de muchos más factores que aquéllos para los que se puede obtener información. Algunos factores, a pesar de tener una clara relación con las variables de pronóstico, son en sí mismos difíciles de pronosticar y otros factores pueden ser interdependientes, de modo que su influencia precisa no es clara.

Hay dos clases principales de modelos: los modelos de <u>*regresión de una sola ecuación*</u> y los modelos de <u>*regresión de ecuaciones múltiples*</u>. El primero comprende modelos en los cuales se asume que la variable dependiente es una función lineal de varias variables explicativas e independientes.

La segunda clase de modelo, el modelo de regresión de ecuaciones múltiples, nos da la posibilidad de tener en cuenta simultáneamente las interrelaciones entre un conjunto de variables. La debilidad de los modelos de ecuaciones múltiples es que las variables explicativas pueden ser interdependientes. La posible realimentación entre la variable dependiente y las variables explicativas es despreciable.

2.2.1 *Regresión de dos variables*

Supongamos que deseamos estudiar la relación entre dos variables, x y y. A fin de describir esta relación estadísticamente, necesitamos un conjunto de observaciones de pares de valores, uno para cada variable:

$$(x_1,y_1),(x_2,y_2),\ \mathrm{K}\ ,(x_n,y_n)$$

Si representamos gráficamente estas cantidades, podemos encontrar que ellas muestran un patrón consistente, por lo que es posible dibujar una curva que "se ajuste" a esos puntos. Se presentan entonces dos problemas:

a) decidir el tipo de curva a usar;

b) elegir la "mejor" curva de ese tipo.

El primer problema se resuelve por lo general simplemente mirando una representación gráfica de los pares de valores observados (x_i, y_i) y, de preferencia, eligiendo por observación un tipo de curva simple. Al respecto, son muy populares las líneas rectas.

El segundo problema: escoger los valores óptimos para los parámetros de la curva, puede resolverse solamente después de haber adoptado algún criterio que nos informe sobre la calidad de la curva. El criterio más usado es el de los mínimos cuadrados. Esto se debe a sus propiedades intuitivas y a sus propiedades matemáticas. Nosotros también lo usaremos. De acuerdo con el criterio de mínimos cuadrados una curva es buena si *S*, la suma de los cuadrados de las "distancias" desde los puntos dados a la curva, es pequeña. La mejor curva es aquélla en la cual el valor de *S* es el más pequeño.

El segundo problema puede entonces reformularse de la siguiente manera: *determinar los valores de parámetro que minimizan S.*

Supongamos que deseamos determinar una línea recta que se ajuste a las observaciones. Esto significa que asumimos una relación lineal entre *x*, la variable independiente, e *y*, la variable dependiente. El hecho de que las observaciones (x_i, y_i) no caigan sobre un línea se atribuye a la influencia de un término de error.

En resumen, se asume que:

$$y = \alpha + \beta\ x + \varepsilon \qquad (2.2.1)$$

Los parámetros α y β son desconocidos. El error ε es aleatorio pero tiene un valor esperado igual a cero.

Sean *a* y *b* las estimaciones de α, respecto de β B. El valor calculado para *y* a partir de un valor dado x_i usando estos estimados es:

$$\hat{y}_i = a + b \cdot x_i \qquad (2.2.2)$$

Sin embargo, observamos el par (x_i, y_i), de modo que $y_i - \hat{y}_i$ es la desviación de la línea de este punto particular. Así, tenemos que encontrar las estimaciones de *a* y *b*, que minimizan:

$$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

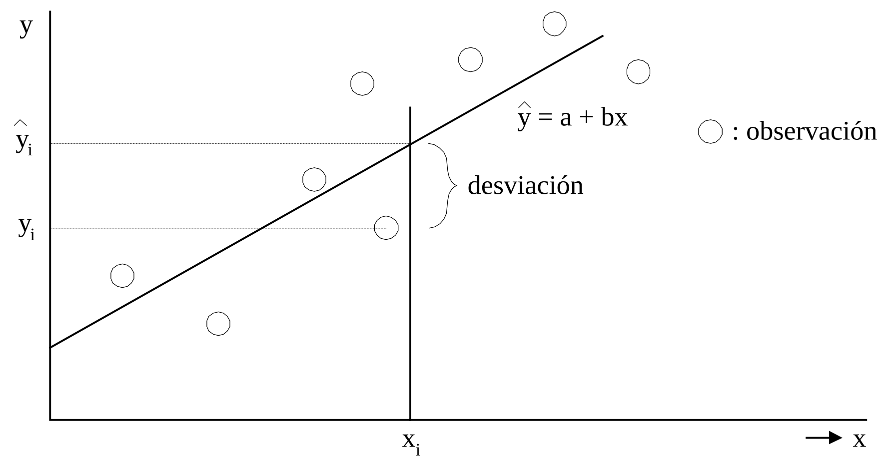


Figura 2.2/1 : Las observaciones y la línea

Nótese que solamente estamos interesados en las desviaciones verticales. Esto no es sorprendente ya que nuestro objetivo es predecir el valor *y*- cuando es dado un nuevo valor de *x*-. En general, tenemos que escoger entre tres posibilidades:

V: Distancias verticales
H: Distancias horizontales
P: Distancias perpendiculares a la línea

La siguiente figura da los resultados obtenidos, empezando por las mismas observaciones en estos tres casos.

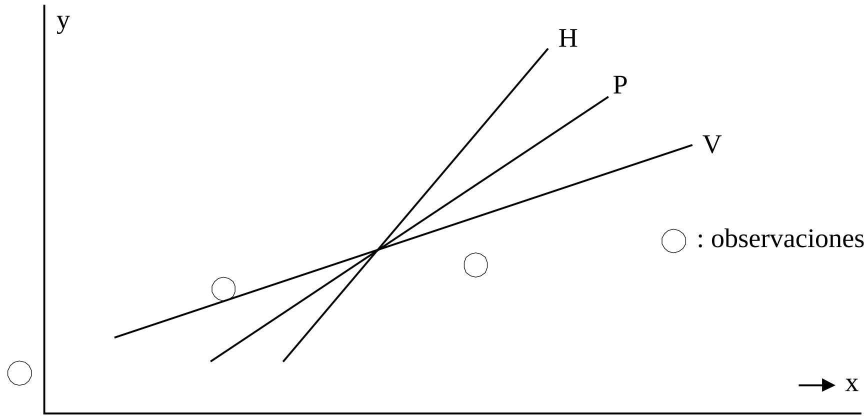


Figura 2.2/2 : Influencia de la "distancia" escogida

Las estimaciones de *a* y *b* se deciden resolviendo:

$$\frac{\partial S}{\partial a} = 0 \quad y \quad \frac{\partial S}{\partial b} = 0$$

donde

$$S = \sum_{i=1}^{n} (y_i - a - b \cdot x_i)^2$$

El resultado es:

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{2.2.3}$$

$$a = y - b \cdot x \tag{2.2.4}$$

Para juzgar la calidad de la línea así producida, podemos calcular el coeficiente de correlación *r*.

$$r^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = \frac{Explained\ variance}{Total\ variance} \tag{2.2.5}$$

r tiene el mismo signo que *b* y cae en [-1, +1]. $r = 0$ indica falta de correlación, en otras palabras: una línea recta no puede representar la relación entre *y* e *x*.

Si *r* se desvía de cero, hay dificultades para juzgar la bondad del ajuste entre las observaciones y la línea recta. Por lo tanto, tienen que aplicarse otros métodos.

Si estamos satisfechos con el valor obtenido de r , podemos usar

$$y = a + b \cdot x$$

para hallar los estimados de y, dado cualquier valor de x.

Hay, sin embargo, otros métodos para verificar que la líneas de regresión proporciona una descripción fiable de cómo y depende de x. El primer paso es estimar la varianza de b y verificar que está dentro de límites razonables. El segundo paso es verificar que no hay errores sistemáticos con el test de “Durbin-Watson”.

El primer test requiere calcular la cantidad

$$T = \frac{b\left(\sum (x - \bar{x})^2\right)^{1/2}}{s} \tag{2.2.6}$$

donde

$$s^2 = \frac{\sum y^2 - a\sum y - b\sum xy}{n-2} \tag{2.2.7}$$

Para una revisión precisa, es necesario usar tablas estadísticas para la distribución de t, pero T debe ser menor que -2 o mayor que +2 para que sea aceptable el análisis de regresión.

Para $-2 < T < 2$, el valor de b no se puede decir que sea significativamente diferente de cero y no podemos estar seguros que nuestra línea recta con $b \neq 0$ represente realmente un cambio de y con x.

Si deseamos diseñar intervalos de confianza para los parámetros α y β, o para nuestros valores de y predichos, debemos asumir lo siguiente:

Las observaciones (x_i, y_i) forman una muestra al azar de una población de dos variables distribuida normalmente, de modo tal que:

$$\begin{aligned} E\{Y \mid X = x\} &= \alpha + \beta \cdot x \\ V\{Y \mid = x\} &= \sigma^2 \end{aligned} \tag{2.2.8}$$

De ahora en adelante y sin pérdida de generalidad, también asumiremos que $\bar{x} = \bar{y} = 0$; esto significa que escogemos un nuevo origen, de manera que antes de empezar los cálculos tenemos que restar $\bar{x}$ a todos los valores x_i - y restar $\bar{y}$ a todos los valores y_i .

El estimado para β puede ahora escribirse así

$$b = \frac{\sum x_i \cdot y_i}{\sum x_i^2} \tag{2.2.3a}$$

y

$$a = 0 \tag{2.2.4a}$$

La varianza σ^2 es desconocida, pero podemos estimarla por

$$s_e^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y})^2 \tag{2.2.8}$$

s_e se conoce como el error estándar del estimado. s_e^2 es un estimador insertado de σ^2.

Además,

$$\frac{(n-2)s_e^2}{\sigma^2} = \chi^2(n-2) \qquad \text{(Chi cuadrado distribuido con } n\text{-}2 \text{ grados de libertad)}$$

$$a \approx N\left(\alpha, \frac{\sigma^2}{n}\right) \quad y \quad b \approx N\left(\beta, \frac{\sigma^2}{\sum x_i^2}\right), \quad así \quad \frac{a-\alpha}{\sigma}\sqrt{n} \quad y \quad \frac{b-\beta}{\sigma}\sqrt{\sum x_i^2}$$

se distribuyen normalmente, *N(0,1)*, (con media = 0 y desviación estándar = 1).

Puede mostrarse que *a, b* y *s* son independientes entre sí, de forma que:

$$\frac{a-\alpha}{\sigma}\cdot n \cdot \frac{\sigma^2}{s_e^2} \quad y \quad \frac{b-s}{\sigma}\cdot\sqrt{\sum x_i^2}\cdot\sqrt{\frac{\sigma^2}{s_e^2}}$$

se distribuyen de acuerdo a una distribución *t* con *n-2* grados de libertad. Pueden calcularse los intervalos de confianza y realizarse las pruebas de hipótesis. Para diseñar un intervalo de confianza para el valor de *y* cuando $x = x_0$, debemos ver la varianza de $y_o = bx_o$. Debido al nuevo origen, el cual es diferente para cada muestra, la varianza de y_o incluye la varianza del origen, V(a), y es igual a:

$$V(y_0) = V(a) + x_0^2 \cdot V(b) = \sigma^2\left(\frac{1}{n} + \frac{x_0^2}{\sum x_i^2}\right) \tag{2.2.9}$$

De nuevo, llegamos a una distribución *t(n-2)* eliminando la incógnita σ^2 porque:

$$\frac{y_0 - \alpha - \beta \cdot x_0}{\sigma \cdot \sqrt{\frac{1}{n} + \frac{x_0^2}{\sum x_i^2}}} \approx N(0,1)$$

y por tanto

$$\frac{y - \alpha - \beta \cdot x_0}{s_e \cdot \sqrt{\frac{1}{n} + \frac{x_0^2}{\sum x_i^2}}} \approx t(n-2)$$

En lo antes expuesto nos hemos limitado a una relación lineal entre *x* e *y*. Sin embargo, el mismo método puede aplicarse a ciertos casos donde parece más apropiada una relación no lineal.

Para esto es suficiente que la relación pueda obtenerse en una forma lineal, por medio de una transformación adecuada de los datos.

Por ejemplo, si asumimos que la relación puede expresarse por

$$y = \alpha \cdot e^{\beta \cdot x} \tag{2.2.10}$$

entonces $$ln\, y = ln\,\alpha + \beta \cdot x \tag{2.2.10a}$$

o $$z = \gamma + \beta \cdot x \tag{2.2.10b}$$

Determinamos los estimados de *c* y *b* para γ , respecto de β, en la forma usual y realizamos el pronóstico con:

$$y = e^{z} = e^{c+b\cdot x} = e^{c} \cdot e^{b\cdot x} \qquad (2.2.10c)$$

Si la transformación de datos a una relación lineal es imposible, entonces debe tratarse con regresión curvilínea.

La idea principal es ahora que la actual relación entre x y y pueda ser representada por:

$$y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot x^2 + \ldots + \beta_k \cdot x^k \qquad (2.2.11)$$

Nuestras observaciones (x_i, y_i) con el resultado de fluctuaciones aleatorias ε_I alrededor de esta línea.

Aplicando los criterios de mínimos cuadrados para la estimación de $\beta_0, \beta_1, \ldots, \beta_k$ significa otra vez que $\sum \varepsilon_i$ tiene que minimizarse.

Así, hemos tratado de encontrar el mínimo de:

$$\sum_{i=1}^{n} \left(y_i - (b_0 + b_1 \cdot x_1 + b_2 \cdot x_i^2 + \ldots + b_k \cdot x_i^k) \right)^2$$

donde los b_i estiman los β_i.

Derivando parcialmente respecto a $b_0, b_1, \ldots, b_k$, e igualando todas las derivadas a cero, se obtiene el sistema de ecuaciones lineales:

$$\sum y_i = nb_0 + b_1 \sum x_i + b_2 \sum x_i^2 + \ldots + b_k \sum x_i^k$$

$$\sum x_i y_i = b_0 \sum x_i + b_1 \sum x_i^2 + b_2 \sum x_i^3 + \ldots + b_k \sum x_i^{k+1}$$

$$\sum x_i^k y_i = b_0 \sum x_i^k + b_1 \sum x_i^{k+1} + \ldots + b_k \sum x_i^{2k}$$

La solución de este sistema de ecauciones lineales con $k+1$ incógnitas proporcionan las estimaciones requeridas. Es aconsejable limitar el número de parámetros b_k usados.

2.2.2 *Regresión múltiple*

Si no estamos satisfechos con el valor de r, lo cual significa que la varianza residuo (= varianza total - varianza explicada) es muy larga, podemos verlos residuos. La varianza residuales

$$(y_i - \hat{y}_i)^2$$

y puede tener un alto valor sólo si al menos algunos de los residuos $y - \hat{y}$ son grandes, lo cual significa que las distancias verticales entre las observaciones y la línea $y = a + bx$ son amplias.

No obstante, si la variable y estudiada parece depender de otras variables además de x, uno puede tratar de reducir los residuos introduciendo uno o más factores “explicativos”, x_1, x_2, etc. Los principales métodos a ser aplicados son el *Método Iterativo* y el *Método Directo,* los cuales se explicarán mediante dos ejemplos numéricos.

Método Iterativo

Ejemplo: Pronóstico del número de llamadas locales

El método mostrado abajo, se llama iterativo porque trabaja en pasos. Se lo prefiere cuando los factores explicativos a utilizarse no se han decidido por anticipado.

Debe notarse que aunque el método implica cierto grado de aproximación, los resultados numéricos serán suficientemente exactos para casos prácticos como los que se dan aquí.

Paso 1 Seleccionar una pareja de factores explicativos posibles, $x_1, \ldots, x_n$ para lo cual se dispone de buenos pronósticos (número de abonados, habitantes, etc.).

Paso 2 Calcular los coeficientes de correlación entre la variable dependiente y (en el ejemplo de abajo el número de llamadas locales) y cada uno de los factores explicativos $x_1, \ldots, x_n$.

Paso 3 Seleccionar el factor explicativo que tiene el coeficiente de correlación más alto, con y_1.

Paso 4 Si se supone que este factor es x_2, hacer el primer intento con la función de regresión:

$$y_1 = a_1 + b_1 \cdot x \tag{2.2.12}$$

y calcular los parámetros de a y b con los métodos indicados en la sección 2.2.1.

Paso 5 Calcular los residuos

$$r_1 = y_1 - (a_1 + b_1 \cdot x_2) \tag{2.2.13}$$

Paso 6 Calcular los coeficientes de correlación entre r_1 :s y cada uno de los factores explicativos sobrantes, en este caso $x_1, x_3, \ldots, x_m$

Paso 7 Seleccionar el factor explicativo restante que tiene el coeficiente de correlación más alto con r_1.

Paso 8 Si a este factor se le llama x1, hacer el segundo intento utilizando la función de regresión

$$r_1 = a_2 + b_2 \cdot x_1 \tag{2.2.14}$$

y calcule los parámetros a_2 y b_2, de acuerdo con los métodos de la sección 2.2.1.

Paso 9 La función de la prueba 1, junto con la función de la prueba 2, dan ahora

$$\begin{aligned} y_1 &= (a_1 + b_1 \cdot x_2) + r_1 = \\ &= (a_1 + b_1 \cdot x_2) + (a_2 + b_2 \cdot x_1) = \\ &= (a_1 + a_2) + b_2 \cdot x_1 + b_1 \cdot x_2 \end{aligned} \tag{2.2.15}$$

los cual es una mejora comparada con la función de regresión de la primera prueba.

Paso 10 Calcular los residuos

$$r_2 = y_1 - [(a_1 + a_2) + b_2 \cdot x_1 + b_1 \cdot x_2] \tag{2.2.16}$$

Paso 11 Calcular los coeficientes de correlación entre los r_2:s y cada uno de los factores explicativos restantes, en este caso $x_3, \ldots, x_n$

Paso 12 Llevar adelante la prueba 3 y seguir de la misma forma mientras los residuos calculados no sean lo suficientemente pequeños.

Un ajuste perfecto no significará que todos los residuos sean cero. Este estado solamente se alcanzará cuando el número de parámetros sea el mismo que el número de observaciones en el material estadístico.

Cálculos Numéricos

Año fiscal	Llamadas locales y_1	Llamadas De Larga Distancia		
		Servicio manual	Servicio automático	Total y_2
1958/59	2305	110	125	235
1959/60	2423	103	165	268
1960/61	2605	95	220	315
1961/62	2852	84	267	351
1962/63	3040	78	322	400
1963/64	3376	73	372	445
1964/65	3642	55	445	500
1965/66	3700	38	530	568
1966/67	3884	21	609	630
1967/68	4059	13	650	663

Tabla 2.2/1 : Llamadas Telefónicas (millones)

Paso 1: Elegir los factores explicativos posibles

Año 1o. enero	Población x_1	Abonados Telefónicos			Aparatos Telefónicos		
		Centrales manuales	Centrales automáticas	Total x_2	Centrales manuales	Centrales automáticas	Total x_3
1959	7 436	352	1 712	2 064	408	2 118	2 526
1960	7 471	311	1 839	2 150	359	2 278	2 637
1961	7 499	271	1 973	2 244	313	2 448	2 761
1962	7 542	207	2 143	2 350	236	2 668	2 904
1963	7 581	174	2 294	2 468	198	2 856	3 054
1964	7 627	144	2 454	2 598	163	3 060	3 223
1965	7 695	101	2 622	2 723	114	3 273	3 387
1966	7 773	75	2 785	2 860	86	3 487	3 573
1967	7 844	41	2 956	2 997	46	3 711	3 757
1968	7 894	27	3 095	3 122	32	3 903	3 935

Tabla 2.2/2 : Factores Explicativos (miles)

1ero de enero	Población x_1	Total Abonados x_2	Total Aparatos x_3
1969	7 964	3 214	4 111
1970	8 042	3 335	4 288
1971	8 120	3 453	4 458
1972	8 201	3 566	4 626
1973	8 281	3 672	4 793
1974	8 360	3 776	4 957
1975	8 430	3 879	5 120

Tabla 2.2/3 : Pronósticos Disponibles para los Factores Explicativos (miles)

Paso 2: Calcular los coeficientes de correlación Nótese que los valores básicos dados en las Tablas 2.2/1-2 se han redondeado en los cálculos siguientes.

El cálculo de coeficientes de correlación entre el y número de llamadas locales y los factores explicativos x_1, x_2 y x_3 dan los resultados que se muestran a continuación. Los datos se han tomado de las Tablas 2.2/1-2.

y_1 10^9	x_1 10^6	x_2 10^6	x_3 10^6
2.31	7.44	2.06	2.53
2.42	7.47	2.15	2.64
2.60	7.50	2.24	2.76
2.85	7.54	2.35	2.90
3.04	7.58	2.47	3.05
3.38	7.63	2.60	3.22
3.64	7.70	2.72	3.39
3.70	7.77	2.86	3.57
3.88	7.84	3.00	3.76
4.06	7.89	3.12	3.94
31.88	76.36	25.57	31.76

Tabla 2.2/4 : Factores Explicativos Dados

Se calculan las siguientes sumas:

$$\sum y_i = 31.88 \qquad \sum y_i^2 = 105.22$$

$$\sum x_1 = 76.36 \qquad \sum x_2 = 25.57 \qquad \sum x_3 = 31.76$$

$$\sum x_1^2 = 583.31 \qquad \sum x_2^2 = 66.58 \qquad \sum x_3^2 = 102.97$$

$$\sum x_1 y_1 = 244.32 \qquad \sum x_2 y_1 = 83.57 \qquad \sum x_3 y_1 = 103.96$$

El coeficiente de correlación se calcula a partir de:

$$R(y_1\, x_i) = \frac{D_i}{\sqrt{G_i \cdot H_i}}$$

donde

$$D_i = n\sum x_i y_1 - \sum x_i \sum y_1$$

$$G_i = n\sum y_1^2 - \left(\sum y_1\right)^2$$

$$H_i = n\sum x_i^2 - \left(\sum x_i\right)^2$$

lo cual da

$$R(x_1\,y_1) = 0.979$$

$$R(x_2\,y_1) = 0.991$$

$$R(x_3\,y_1) = 0.989$$

Paso 3: Elegir el factor explicativo a usarse primero.

En esta primera prueba escogemos el factor explicativo con el más alto coeficiente de correlación con y_1. Aquí el número de abonados era = x_2.

Paso 4: Hacer la prueba número 1 con función de regresión

$$y_1 = a_1 + b_1 \cdot x_2$$

Las fórmulas para el cálculo de a y b en la sección 2.2 pueden escribirse de la siguiente manera:

$$a_1 n + b_1 \sum x_2 = \sum y_1$$

$$a_1 \sum x_2 + b_1 \sum x_{2i}^2 = \sum x_2 y_1$$

o con los valores numéricos:

$$10.00\ a_1 + 25.57\ b_1 = 31.88$$

$$25.57\ a_1 + 66.58\ b_1 = 83.57$$

La solución es:

$$a_1 = -\ 1.196$$

$$b_1 = \ \ 1.715$$

Paso 5: Calcular el residuo r_1.

La función encontrada en el paso 4 (prueba 1) da los siguientes valores ajustados de y_1:

Año fiscal	x_2	Valores ajustados - 1.13 + 1.677 x = y_1	Comparación con los dados de y	
			Tabla 2.2/1 y	residuo r_1
1958/59	2.06	2.34	2.31	- 0.03
1959/60	2.15	2.49	2.42	- 0.07
1960/61	2.24	2.64	2.60	- 0.04
1961/62	2.35	2.83	2.85	+ 0.02
1962/63	2.47	3.04	3.04	+ 0.00
1963/64	2.60	3.26	3.38	+ 0.12
1964/65	2.72	3.47	3.64	+ 0.17
1965/66	2.86	3.71	3.70	+ 0.01
1966/67	3.00	3.95	3.88	- 0.07
1967/68	3.12	4.15	4.06	- 0.09

Tabla 2.2/5 : Valores ajustados de *y* en el paso 4

Paso 6: Calcular los coeficientes de correlación con r_1.

Antes de la segunda prueba calcularemos los coeficientes de correlación entre los residuos de la prueba 1 y el resto de factores explicativos x_1 y x_3.

Residuos de la prueba No. 1 r_1	x_1 de la Tabla 2.2/2	x_3	Sumas calculadas Variable	Valor
- 0.03	7.44	2.53	$\sum r_1^2$	= 0.0642
- 0.07	7.47	2.64		
- 0.04	7.50	2.76	$\sum x_1^2$	= 583.31
- 0.02	7.54	2.90		
- 0.00	7.58	3.05	$\sum x_3^2$	= 102.97
+ 0.12	7.63	3.22		
+ 0.17	7.70	3.39	$\sum r_1 x_1$	= - 0.0073
- 0.01	7.77	3.57		
- 0.07	7.84	3.76	$\sum r_1 x_3$	= - 0.1199
- 0.09	7.89	3.94		
Total	76.36	31.76		

Tabla 2.2/6 : Cálculo de los coeficientes de correlación.

Los coeficientes de correlación fueron:

$R\ (r_1,\ x_1) = -\ 0.029$

$R\ (r_1,\ x_3) = -\ 0.019$

Esto indica una correlación muy baja.

Paso 7: Elegir el siguiente factor explicativo.

El más alto coeficiente de correlación es el que está entre r_1 y x_1, así que nuestro segundo intento será para usar el factor explicativo x_1.

Paso 8: Hacer el segundo intento con la función:

$r_1 = a_2 + b_2\, x_1$

Los parámetros a_1 y b_2 se dan a partir de las ecuaciones siguientes:

$$a_2 n + b_2 \sum x_1 = \sum r_1$$

$$a_2 \sum x_1 + b_2 \sum x_1^2 = \sum r_1 x_1$$

o $\quad 10.00\ a_2 + 76.36\ b_2 = -\ 0.04$

$76.36\ a_2 + 583.31\ b_2 = -\ 0.3089$

La solución es entonces:

$a_2 = 0.18 \qquad 0.113$

$b_2 = -\ 0.023 \qquad -\ 0.0154$ (cor - 0.028822)

Consecuentemente $r_1 = 0.113 - 0.0154\,x_1$

Paso 9: Los resultados de las pruebas 1 y 2 anteriores darán ahora la siguiente función de regresión para llamadas locales:

$$\begin{array}{lll} y_1 = \underset{\substack{\uparrow \\ \text{(de la prueba 1)} \\ \downarrow}}{(-1.196 + 1.715\,x_2)} & \underset{\substack{\uparrow \\ \text{(de la prueba 2)} \\ \downarrow}}{+} & r_1 \quad = \\ \quad = (-1.196 + 1.715\,x_2) & + & (0.113 - 0.0154\,x_1) \end{array}$$

consecuentemente

$$y_1 = -1.083 - 0.0154\,x_1 + 1.715\,x_2$$

Este resultado está comparado con los valores dados y_1 en la Tabla 2.2/1.

Paso 10: Calcular los residuos r_2

Año fiscal	Valores ajustados para llamadas locales - 1.083 - 0.0154 x_1 + 1.715 x_2 = y_1	Valores dados de *y* Tabla 2.2/1 x_2	Residuo r_2
1958/59	2.34	2.31	- 0.03
1959/60	2.49	2.42	- 0.07
1960/61	2.64	2.60	- 0.04
1961/62	2.83	2.85	+ 0.02
1962/63	3.04	3.04	0.00
1963/64	3.26	3.38	+ 0.12
1964/65	3.46	3.64	+ 0.18
1965/66	3.70	3.70	0.00
1966/67	3.94	3.88	- 0.06
1967/68	4.15	4.06	- 0.09

Apartir de esta función de regresión, $y\ (x_1,\ x_2)$, encontramos el siguiente número futuro de llamadas locales (en millones):

	x_1	x_2	$y\ (x_1x_2)$	$y(x_2)$
1968/69	7.96	3.21	4300	4310
1969/70	8.04	3.34	4520	4530
1970/71	8.12	3.45	4710	4720
1971/72	8.20	3.57	4910	4930
1972/73	8.28	3.67	5080	5100
1973/74	8.36	3.78	5270	5290
1974/75	8.43	3.88	5440	5460

El resultado se da en la Figura 2.2/3

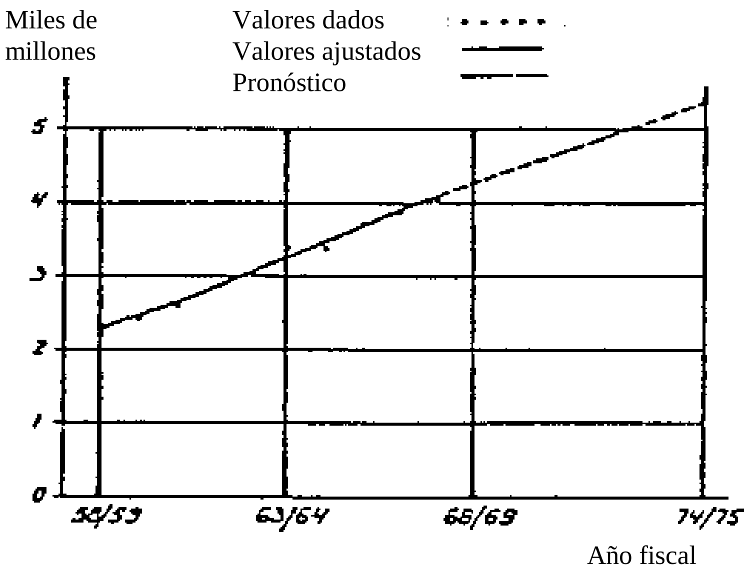


Figura 2.2/3

A partir de los residuos r_2 de deduce que la prueba 2 prácticamente no dio diferencia comparada con la primera prueba. Esto se desprende del pequeño coeficiente de correlación $R(r_1, x_1)$.

Debemos, sin embargo, detenernos en el paso 8 y usar la función de regresión.

$$y_1 = -1.083 - 0.0154\, x_1 + 1.715\, x_2$$

Método Directo

Para este método asumimos desde el principio que la variable que deseamos pronosticar depende de un número de variables explicativas.

Ejemplo: *Pronóstico del número de llamadas de larga distancia.*

El número de llamadas de larga distancia, y_2, generalmente aumenta considerablemente al introducirse el servicio automático (STD). Con frecuencia se ha observado que la introducción del STD entre dos áreas causa la duplicación del número de llamadas.

Por tanto, parece apropiado usar el factor explicativo *"grado de automatización"* (representado por x_4) para pronosticar el número de llamadas de larga distancia. Este grado se define como la relación entre el número de llamadas de larga distancia automáticas y el total de llamadas de larga distancia.

También es probable que el número de aparatos telefónicos (representado por x_3) sea un factor explicativo útil. Después de estas consideraciones, seleccionamos a priori los dos factores explicativos:

- número de aparatos telefónicos $= x_3$
- grado de automatización $= x_4$

Se supone que no necesitamos más factores explicativos, así que podemos aplicar el "*método directo*" para el cálculo de parámetros en la función de regresión:

$$y_2 = a + b_3 \cdot x_3 + b_4 \cdot x_4 \tag{2.2.17}$$

Año fiscal	Llamadas de larga distancia en miles de millones y_2	Aparatos telefónicos (en millones) x_3	Grado de automatización x_4
1958/59	0.235	2.53	0.53
1959/60	0.268	2.64	0.62
1960/61	0.315	2.76	0.70
1961/62	0.351	2.90	0.76
1962/63	0.400	3.05	0.81
1963/64	0.445	3.22	0.84
1964/65	0.500	3.39	0.89
1965/66	0.568	3.57	0.93
1966/67	0.630	3.76	0.97
1967/68	0.663	3.94	0.98
Total	4.375	31.76	8.03

Tabla 2.2/8 Llamadas de larga distancia y factores explicativos

Año fiscal		Aparatos telefónicos en millones x_3	Grado de automatización x_4
1968/69		4.11	0.99
1969/70		4.29	0.99
1970/71		4.46	1.00
1971/72		4.63	1.00
1972/73		4.79	1.00
1973/74		4.96	1.00
1974/75		5.12	1.00

Tabla 2.2/9 : Pronósticos dados para los factores explicativos

A partir de las cifras del período 1958-1968, calculamos los siguientes totales:

$$\sum x_3^2 = 102.97 \qquad \sum x_3 x_4 = 26.13$$

$$\sum x_4^2 = 6.65 \qquad \sum y_2 x_3 = 14.55$$

$$\sum y_2 x_4 = 3.\ 71$$

Los parámetros a, b_3 y b_4 se estiman a partir del sistema de ecuaciones:

$$a\,n + b_3 \sum x_3 + b_4 \sum x_4 = \sum y_2$$

$$a \sum x_3 + b_3 \sum x_3^2 + b_4 \sum x_3 x_4 = \sum y_2 x_3$$

$$a \sum x_4 + b_3 \sum x_3 x_4 + b_4 \sum x_4^2 = \sum y_2 x_4$$

ó

$$\begin{aligned} 10.00\,a + 31.76b_3 + 8.03b_4 &= 4.38 \\ 31.76\,a + 102.97b_3 + 26.13b_4 &= 14.55 \\ 8.03\,a + 26.13b_3 + 6.65b_4 &= 3.71 \end{aligned}$$

La solución del sistema de ecuaciones es:

$a = -0.481$
$b_3 = 0.222$
$b_4 = 0.267$

En consecuencia, la función de regresión será:

$$y_2 = -0.481 + 0.222 \cdot x_3 + 0.267 \cdot x_4$$

En la tabla siguiente se compara el resultado con los valores y_2 dados:

Año fiscal	$y_2 =$ $-0.481 + 0.222\ x_3 + 0.267\ x_4$	Comparación con los valores dados de	
		De la Tabla 2.2/8	Residuo *r*
1958/59	0.223	0.235	+0.012
1959/60	0.271	0.268	-0.003
1960/61	0.319	0.315	-0.004
1961/62	0.366	0.351	-0.015
1962/63	0.412	0.400	-0.012
1963/64	0.458	0.445	-0.013
1964/65	0.510	0.500	-0.010
1965/66	0.560	0.568	+0.008
1966/67	0.613	0.630	+0.017
1967/68	0.656	0.663	+0.007

Tabla 2.2/10 Valores ajustados de las llamadas de larga distancia *(y_2)*

Extrapolando estos valores al futuro, encontramos el siguiente pronóstico para el número de llamadas de larga distancia en miles de millones:

1968/69	0.695
1969/70	0.735
1970/71	0.776
1971/72	0.814
1972/73	0.887
1974/75	0.923

El resultado se presenta en la Figura 2.2/4.

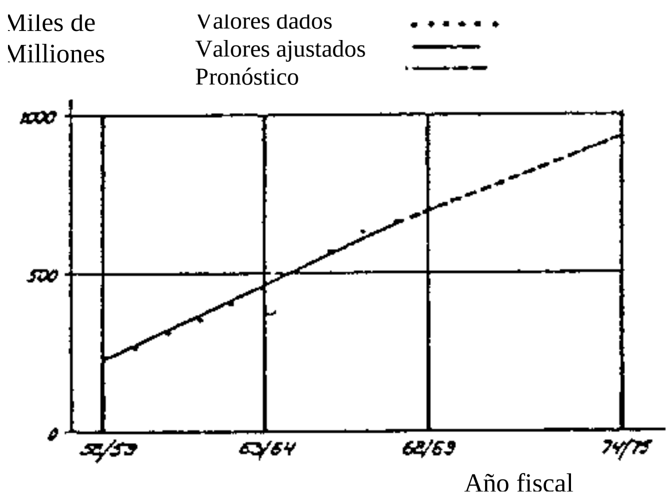


Figura 2.2/4

2.3 <u>Métodos de tendencia</u>

Una serie de tiempo es un conjunto de observaciones de una cantidad como una función de tiempo. Las observaciones pueden seguir cierto patrón y el planificador desea averiguar cómo se desarrollará en el futuro la cantidad observada, es decir, estimar su tendencia.

Esta sección tratará sobre las posibles curvas para el desarrollo, tales como la tendencia lineal, la tendencia exponencial y las tendencias hacia un nivel de saturación. La sección sobre series de tiempo trata con los métodos para explicar una variación observada y el patrón de crecimiento como resultado de cuatro factores: tendencia, variaciones cíclicas, variaciones estacionales y variaciones aleatorias.

Si los datos históricos disponibles son correctos, los métodos de tendencia pueden usarse para pronosticar el desarrollo futuro sobre un importante supuesto: en el futuro existirá la misma dependencia entre las variables. Este es el mismo supuesto que se hace para todo pronóstico basado en la extrapolación de curvas que describen el pasado.

2.3.1 *Análisis de series de tiempo*

El análisis de series de tiempo es el nombre para un número de métodos requeridos para encontrar una explicación para las fluctuaciones en los valores (y_i) estudiados.

Usualmente los valores de y se piensa que son el resultado de cuatro factores:

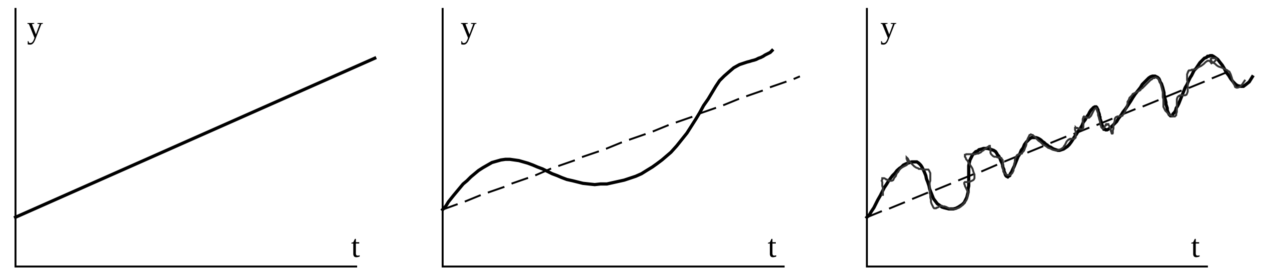
T. La tendencia T es el proceso de crecimiento fundamental (el cual puede ser, por supuesto, negativo o cero).

C. Las variaciones cíclicas C son cantidades sinusoidales con muy baja frecuencia. Un ejemplo de esto son los ciclos de negocios: siete años de prosperidad seguidos por siete años de escasez.

S. Las variaciones estacionales S también oscilan, pero con una mayor frecuencia y son generalmente más pronunciadas que las variaciones cíclicas. Usualmente se incorporan en S patrones de variaciones diarias, semanales o anuales.

I. Los movimientos irregulares o aleatorios (I) se deben a sucesos tales como huelgas, terremotos, cambios tarifarios, etc., y pueden considerarse como valores residuales no explicados de T, C y S.

La figura 2.3/1 muestra los efectos de T, C, y S.



(a) Tendencia de largo plazo (b) Tendencia de largo plazo, movimiento cíclico (c) Tendencia de largo plazo, movimientos cíclicos y estacionales

Figura 2.3/1 : Influencia de las variaciones cíclicas y estacionales en la tendencia, en una serie de tiempo ideal

Los primeros tres factores se estudian por separado. El factor 1 no puede expresarse como una función de tiempo y actúa como una variación residual.

Podemos expresar el efecto combinado de los cuatro factores como:

$$Y = T + S + C + I \qquad (2.3.1)$$

o por medio de una multiplicación, expresada por la fórmula:

$$Y = T \cdot S \cdot C \cdot I \qquad (2.3.2)$$

También es posible una combinación de suma y multiplicación, por ejemplo:

$$Y = T + S \cdot C \cdot I \tag{2.3.3}$$

Para pronósticos a largo plazo por lo general es suficiente añadir los datos dentro de los valores anuales e intentar ajustar una tendencia lineal a través de ellos.

Se puede ajustar una variedad de curvas matemáticas para analizar datos de series de tiempo, desde una extrapolación a mano alzada de un gráfico de la serie, el ajuste de una simple línea recta a la serie de datos transformada, hasta el uso de curvas polinómicas e inclusive métodos aún más complicados.

Tendencia lineal

Cualquier línea recta que proporcione la relación entre dos variables puede expresarse como:

$$y = a + b \cdot t \tag{2.3.4}$$

Cuando a y b tienen sus valores especificados, la línea está unívocamente definida y se puede calcular el valor de y para cualquier valor de t. El planificador dispone de un número de pares de valores registrados de y e t. Por lo general estos valores no caen exactamente a lo largo de una línea recta.. Su tarea es estimar a y b. Asumiendo que se dispone de n registros, los valores de a y b, calculados con el criterio de mínimos cuadrados, se obtienen resolviendo el siguiente par de ecuaciones simultáneas:

$$\sum y = na + b\sum t \tag{2.3.5}$$

$$\sum t \cdot y = a\sum t + b\sum t^2 \tag{2.3.6}$$

donde n = número de pares de valores (t, y);

Método

Calcular los siguientes valores:

$$\sum t \cdot y = t_1 y_1 + t_2 y_2 + \ldots + t_n y_n$$

$$\sum t^2 = t_1^2 + t_2^2 + \ldots + t_n^2$$

$$\sum t = t_1 + t_2 + \ldots + t_n$$

$$\sum y = y_1 + y_2 + \ldots\ y_n$$

n = número de pares = número de valores de t = número de valores de y

Aquí: $\bar{t} = \dfrac{\sum t}{n}$ $\qquad \bar{y} = \dfrac{\sum y}{n}$

Finalmente, calcular a y b para el sistema de ecuaciones (2.3.5) y (2.3.6):

$$b = \frac{\sum t \cdot y - n \cdot \bar{t} \cdot \bar{y}}{\sum t^2 - n\bar{t}^2} \tag{2.3.7}$$

y

$$a = \bar{y} - b\bar{t} \tag{2.3.8}$$

Para cada valor de t, el valor de y está dado a partir de:

$$y = a + b \cdot t$$

la cual es la ecuación para la tendencia de línea recta que estamos buscando.

Tendencia exponencial

El crecimiento exponencial se puede expresar con la ecuación

$$y = a \cdot e^{b \cdot t} \tag{2.3.9}$$

donde *y* e *t* son las variables, y *a* y *b* son los parámetros de la curva, y *e* = es la base de los logaritmos naturales.

Tomando el logaritmo de toda la ecuación:

$$ln\, y = ln\, a + b \cdot t \tag{2.3.10}$$

Introducir

$$z = ln\, y \qquad \text{y} \qquad c = ln\, a$$

La ecuación puede entonces escribirse:

$$z = c + b \cdot t \tag{2.3.11}$$

Vemos que esta ecuación es una línea recta.

Método

Tomar el logaritmo de todos los valores de y y llamarlos z.

Luego calcular:

$$\sum tz \qquad \sum t^2 \qquad \sum t \qquad \sum z$$

$$\bar{t} = \frac{\sum t}{n} \qquad \bar{z} = \frac{\sum z}{n} \qquad n\,\bar{t}^2$$

como se hizo antes !

Entonces tenemos:

$$b = \frac{\sum t \cdot z - n \cdot \bar{t} \cdot \bar{z}}{\sum t^2 - n \cdot \bar{t}^2} \tag{2.3.12}$$

y

$$\begin{aligned} c &= \bar{z} - b \cdot t \\ z &= c + b \cdot t \end{aligned} \tag{2.3.13}$$

Para cada valor de *t*, se calcula ahora el valor de y como

y = antilogaritmo de z (2.3.14)

Verificación del modelo de regresión

A *Pruebas del significado de los parámetros de tiempo*

Este enfoque de la regresión depende del supuesto de que la cantidad en consideración (un tráfico medido o una tasa de llamadas) se correlaciona con el tiempo. Esto puede o no ser verdad; los valores pueden ser bastante independientes del tiempo. No obstante, las fórmulas anteriores siempre producen un valor numérico para *b*. Por tanto es necesario probar si este valor se alcanza por puro azar y si su media real es cero. Esto puede llevarse a cabo calculando el valor

$$H = \frac{b\sum(t-\bar{t})^2}{s} \qquad \text{donde} \qquad s^2 = \frac{\sum y^2 - a\sum y - b\sum ty}{n-2} \tag{2.3.15}$$

Calcular primero s^2 , luego calcular s = raíz cuadrada de s^2.

Las notaciones anteriores son válidas para la línea recta. Para el caso exponencial, calcular s^2 como

$$s^2 = \frac{\sum z^2 - c\sum z - b\sum t \cdot z}{(n-2)} \tag{2.3.16}$$

Para una verificación precisa, es necesario usar tablas estadísticas, pero como una guía aproximada, H debe ser mayor que 2 o menor que -2 para que la descripción lineal sea válida. Si se obtienen valores en el rango de - 2 a +2, entonces el valor de b no es significativamente diferente de cero, lo cual significa que la cantidad de tráfico no cambia sistemáticamente con el tiempo y que la descripción analítica es inapropiada. El mejor pronóstico para valores futuros de y es que no cambian con el tiempo.

B *Examen de errores sistemáticos*

Aun si el parámetro b pasara la prueba anterior, el modelo podría invalidarse debido a la existencia de errores sistemáticos en el proceso de ajustar la "mejor" línea a los puntos de datos. Debe usarse el siguiente procedimiento para verificar los errores sistemáticos:

Representar los valores en la línea ajustada correspondiente a los valores registrados de y por $\hat{y}$, y luego representar las primeras lecturas por $y_1, \hat{y}_1$, las segundas por $y_2, \hat{y}_2$, y así sucesivamente.

Calcular

$$w = \sum_{i=1}^{n-1}(y_i - \hat{y}_i)(y_{i+1} - \hat{y}_{i+1}) \tag{2.3.17}$$

y

$$v = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2.3.18}$$

Calcular el valor $2 - 2\frac{w}{v}$

La estadística conocida como: _Estadística de Durbin-Watson_, debe encontrarse entre 1.5 y 2.5 y de preferencia entre 1.7 y 2.3. Si el valor está fuera de este rango, deben examinarse los datos por indicios de desviación sistemática de la curva ajustada. Son causas comunes la forma errónea de la curva (ver figura 2.3/2, gráfico 3) o la discontinuidad en los datos (ver la misma figura, gráfico 4). Si la *Estadística de Durbin-Watson* da un resultado de prueba insatisfactorio, deben hacerse los ajustes convenientes a los datos y repetirse el análisis. De otra manera los pronósticos pueden estar sujetos a un error serio.

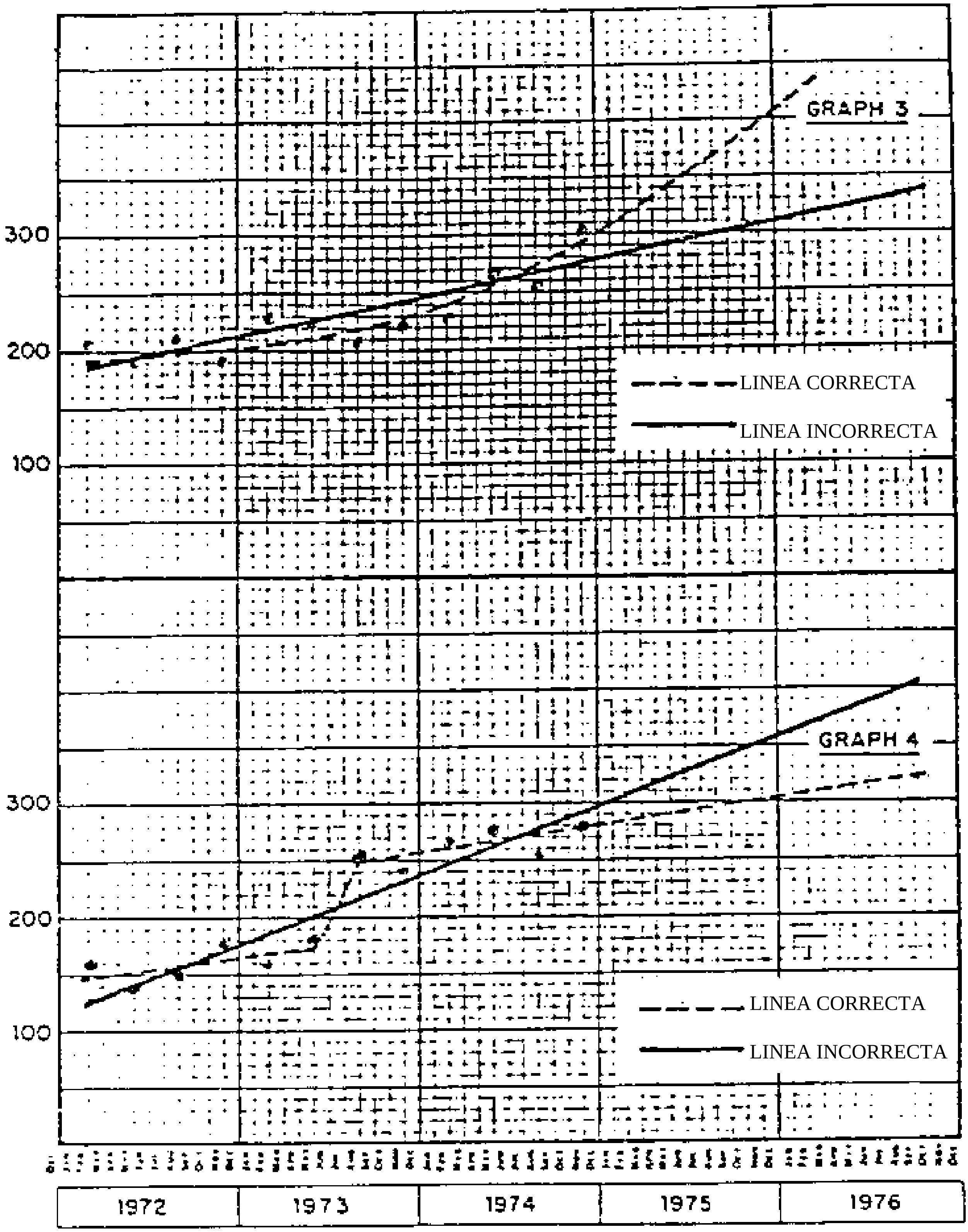

GRAPH 3
300
200
100
LINEA CORRECTA
LINEA INCORRECTA
GRAPH 4
300
200
100
LINEA CORRECTA
LINEA INCORRECTA
1972
1973
1974
1975
1976

Figura 2.3/2

Pronóstico de futuros niveles

Asumiendo que la descripción analítica del pasado logró aprobar las pruebas anteriormente descritas, el siguiente paso es proporcionar los pronósticos requeridos.

Todo lo que se requiere es calcular el valor de y desde:

$$y = a + b \cdot t$$

(Si una curva de crecimiento exponencial se ajusta a los datos históricos, el actual nivel de pronóstico se obtiene tomando el antilogaritmo natural del valor de y).

Una idea general de la precisión del pronóstico puede obtenerse diseñando un intervalo de confianza para el valor del pronóstico, dentro del cual pueda establecerse *el valor futuro con un grado predeterminado de probabilidad*.

Por ejemplo, si el pronóstico para un tiempo futuro t_0 es y_0 , entonces hay aproximadamente un 95% de probabilidad de que el valor logrado de y esté dentro del rango.

$$y_0 \pm 2\sqrt{u}$$

donde $$u = s^2\left[1 + \frac{1}{n} + \frac{(t_0 - \bar{t})^2}{\sum(t - \bar{t})^2}\right] \qquad (2.3.19)$$

y s^2 se calcula a partir de (2.3.15).

Debe destacarse (ver Figura 2.3/3) que cuanto más lejano sea el pronóstico futuro, más amplio será el intervalo de confianza y, por consiguiente, el pronóstico será más incierto.

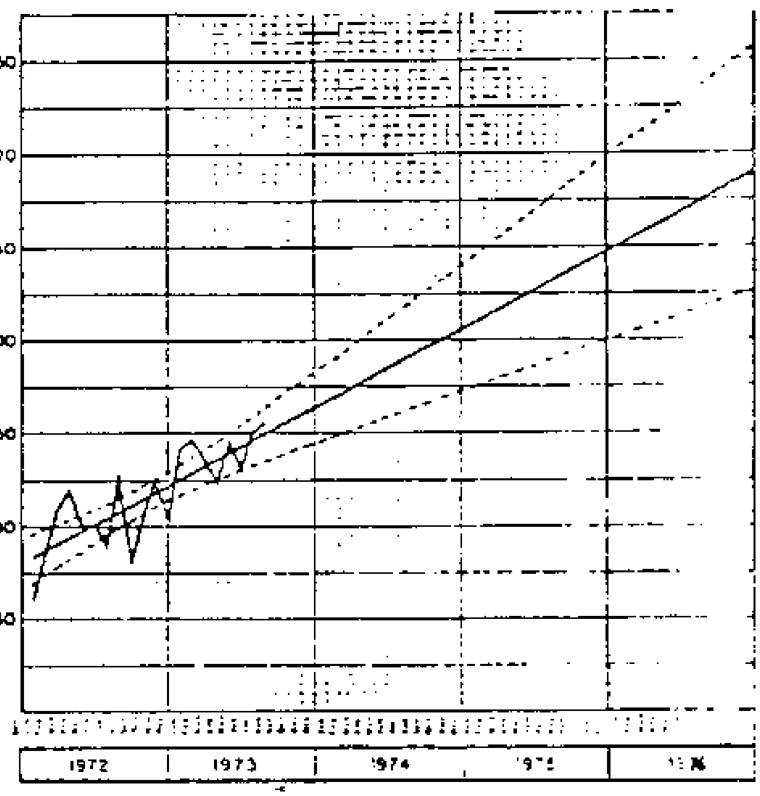


Figura 2.3/3 : Intervalo de confianza (95 %) para el desarrollo lineal pronosticado

Si las cantidades expresadas en un gráfico muestran un patrón consistente a lo largo del tiempo es posible dibujar la curva que mejor ajuste. Entonces se prevén los valores futuros por extrapolación de esta curva.

El primer problema es encontrar la ecuación para la curva que mejor describa la relación entre cantidad y tiempo.

Para muchos tipos de equipos técnicos, la historia de su uso puede describirse en tres fases. En la primera fase observamos un uso acelerado, que en la segunda fase cambia a un incremento lineal en el tiempo. En la tercera fase el uso disminuye y eventualmente alcanza un nivel de saturación. Ejemplos de tal fenómeno pueden ser el número de hogares con aparatos de radio, televisión y teléfonos.

Para otros casos el nivel de saturación puede llegar a cero. Por ejemplo: el uso de carruajes en Londres, el empleo del servicio telegráfico, la utilización de embarcaciones convencionales para transportar mercadería, el uso de hachas de piedra en las batallas etc.

No vamos a tratar con este último tipo de curva de desarrollo.

Para una cantidad que en el futuro pueda alcanzar un nivel de saturación, es posible describir la historia completa con una sola expresión matemática. No obstante, puede ser más simple y más exacto describir cada fase por separado.

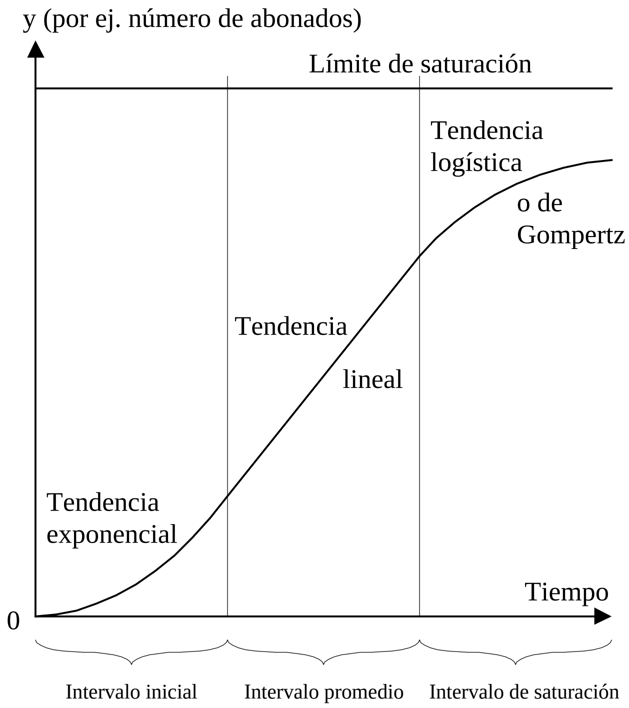


Figura 2.3/4 : Desarrollo en el tiempo de un servicio de telecomunicaciones

Las curvas que describen la tendencia de desarrollo en el tiempo a menudo son llamadas curvas de crecimiento, aun si el "crecimiento" es a veces realmente una disminución en la cantidad. Aquí hay algunos tipos comunes de curvas de tendencia.

Lineal: $y = a + b \cdot t$ (2.3.20)

Parabólica: $y = a + b \cdot t + c \cdot t^2$ (2.3.21)

Exponencial: $y = a\, e^{b \cdot t}$ (2.3.22)

Gompertz: $y = e^{a - b \cdot r(t)}$ (2.3.23)

<u>Notaciones</u>

t =	punto de tiempo (variable independiente)
a, b, c, r =	parámetros a ser calculados a partir de los datos históricos
y =	ítem a ser pronosticado (variable dependiente)
e =	base del logaritmo natural

A continuación damos algunos ejemplos numéricos simples sobre tendencias:

1) *Tendencia lineal* $y = a + b \cdot t$

La fórmula tiene como parámetros desconocidos *a* y *b,* los que han de calcularse a partir de los datos históricos dados. Para calcular dos parámetros necesitamos dos ecuaciones.

Estas dos ecuaciones pueden obtenerse de dos puntos en un diagrama a través del cual pasará la línea recta.

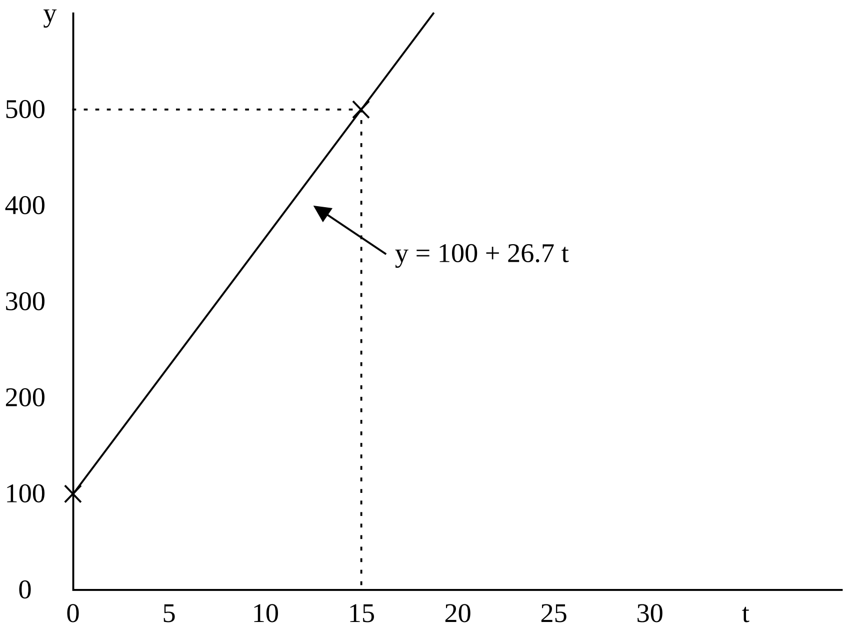


Figura 2.3/5 : Tendencia lineal, estimada desde 2 puntos.

Asumir los siguientes puntos: $t = 0$, $y = 100$ $t = 15$, $y = 500$

Al insertar los dos puntos en la ecuación de la línea recta (2.3.20) se obtiene:

$$100 = a + b \cdot 0$$

$$500 = a + b \cdot 15$$

lo cual da: $a = 100$ y $b = \dfrac{500 - 100}{15} = 26.7$

Esto da la tendencia:

$$y = 100 + 26.7 \cdot t$$

2) *Tendencia exponencial* $y = a\,e^{bt}$

Asumir los mismos puntos que en el caso anterior:

$t = 0$ $t = 15$

$y = 100$ $y = 500$

Las dos ecuaciones requeridas serán

$100 = a \cdot e^{b \cdot 0}$

$500 = a \cdot e^{b \cdot 15}$ Por lo que: $a = 100$ y b se calcula a partir de:

$500 = 100 \cdot e^{b \cdot 15}$

$5 = e^{b \cdot 15}$

$ln5 = 1.609 = 15 \cdot b$ lo cual da $b = 0.1073$

La tendencia es: $y = 100\,e^{0.1073 \cdot t}$

La curva se traza en el diagrama siguiente:

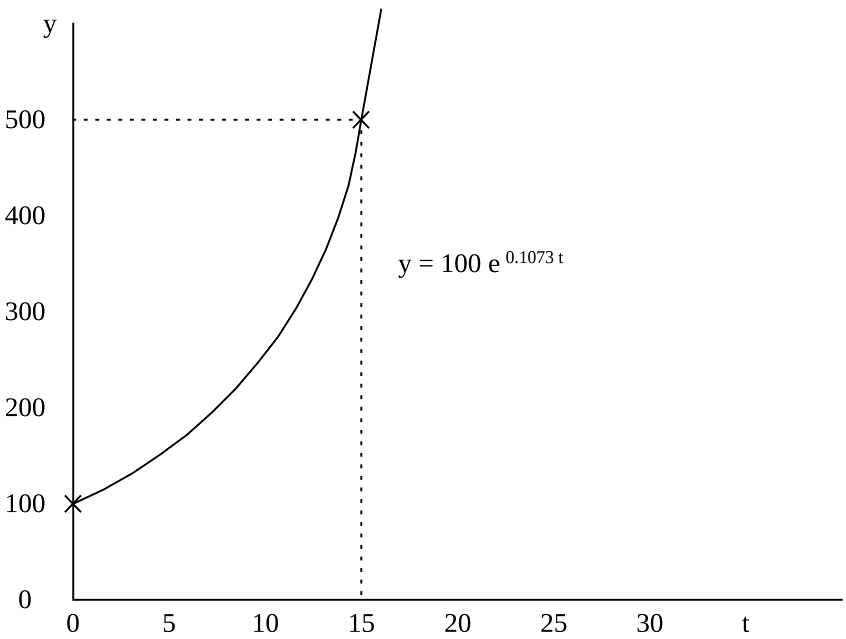


Figura 2.3/6 : Tendencia exponencial determinada a partir de 2 puntos

3) *Tendencia Gompertz'* $y = e^{a - b \cdot r^t}$

En este caso tenemos tres parámetros: *a, b* y *r*, y el cálculo requiere de tres ecuaciones para la determinación de los parámetros. Usamos los mismos dos puntos que antes.

$t = 0$ $t = 15$

$y = 100$ $y = 500$

Puede obtenerse una tercera ecuación asumiendo un valor de <u>*saturación*</u> en el infinito, es decir, el punto:

$t = \infty$ $y = 3000$ (es decir, el valor 3000 será alcanzado después de un número infinito de años)

Los parámetros se calculan a partir de:

$$100 = e^{a-b \cdot r^{0}} \qquad\qquad 500 = e^{a-b \cdot r^{15}}$$

La tercera ecuación da desde que $r < 1$ y t = ∞

$$3000 = e^{a} \qquad \text{por tanto} \qquad a = 1n3000 = 8.006$$

Si $a = 8.006$ se inserta en la primera ecuación para,

$$4.605 = 8.006 - b; \qquad \text{lo cual da} \qquad b = 3.401$$

Entonces r se calcula desde la segunda ecuación para

$$a = 8.006 \qquad \text{y} \qquad b = 3.401$$

$$6.215 = 8.006 - 3.401\, r^{15}$$

ó $$r = \left(\frac{8.006 - 6.215}{3.401} \right)^{1/15} \qquad\qquad r = 0.958$$

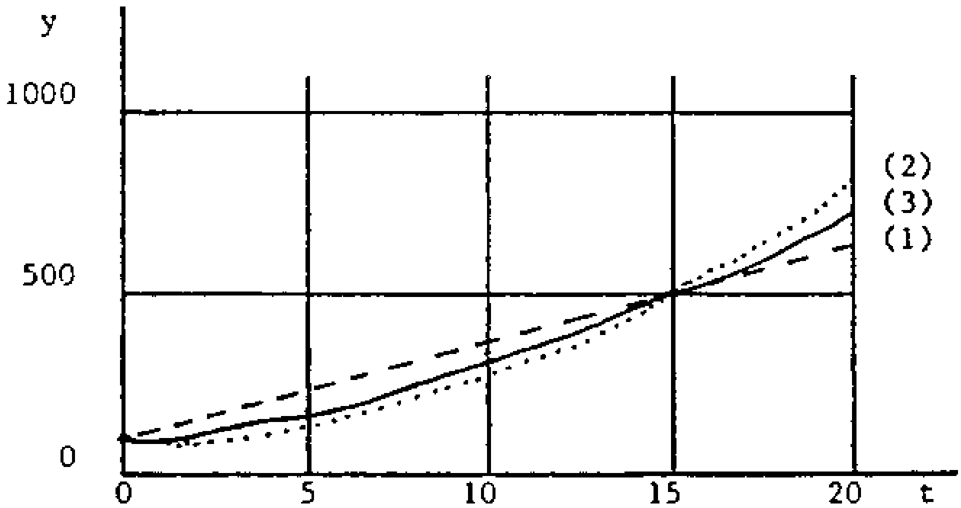


Figura 2.3/7 : Ejemplos numéricos sobre tendencias

1) Tendencia lineal: $y = 100 + 26.7 \cdot t$

2) Tendencia exponencial: $y = 100 \cdot e^{0.1073 \cdot t}$

3) Tendencia Gompertz': $y = e^{8.006 - 3.401 \cdot (0.958)^{t}}$
(valor de saturación = 3000)

Otras curvas de crecimiento:

Considerar

$$y = M - a \cdot e^{\cdot t} \tag{2.3.24}$$

Hay cuatro configuraciones de esta curva de acuerdo a los signos de *a* y *b*. Estas se ilustran a continuación.

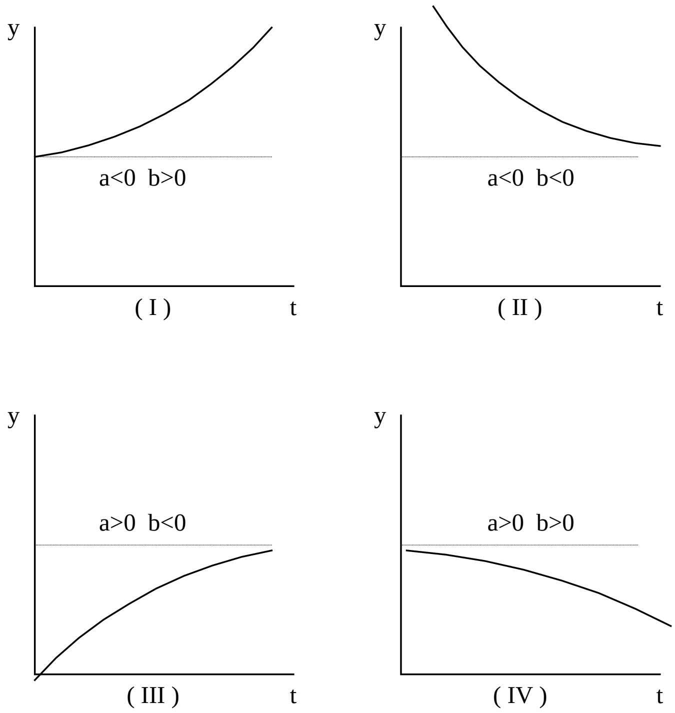


Figura 2.3/8 : Curvas exponenciales modificadas

Cuando se prevé la demanda en telecomunicaciones, sólo las configuraciones (ii) y (iii) se aplican en los pronósticos a largo plazo, de las cuales (ii) es la que se usa con menos frecuencia. Para variables como el número de llamadas manuales por abonado o el número de telegramas per capita, puede usarse un descenso hacia un mínimo como se muestra en (ii). Sin embargo, pocos servicios de telecomunicaciones se caracterizan por una declinación sostenida. Por lo tanto, (iii) es la curva más aplicable. Por supuesto, la aproximación implicada hacia un nivel de saturación como el que se muestra en (iii) puede cubrir un período que se extiende sobre varias décadas. Posiblemente ésta pueda aplicarse para el pronóstico de densidad telefónica en países con muy alta penetración, pero es inaplicable donde el desarrollo de la telefonía está en una etapa temprana. Una alternativa a la curva de Gompertz es la curva logística, que se describe a continuación.

<u>Curva logística</u>

La curva logística simple está dada por la fórmula:

$$y = F + \frac{M - F}{1 + a \cdot e^{-b \cdot t}} \qquad (2.3.25)$$

La curva tiene forma de *S* con un valor mínimo *F*, y un máximo *M*. La curva aumenta a su máximo a lo largo del tiempo para $a > 0,\ b > 0$.

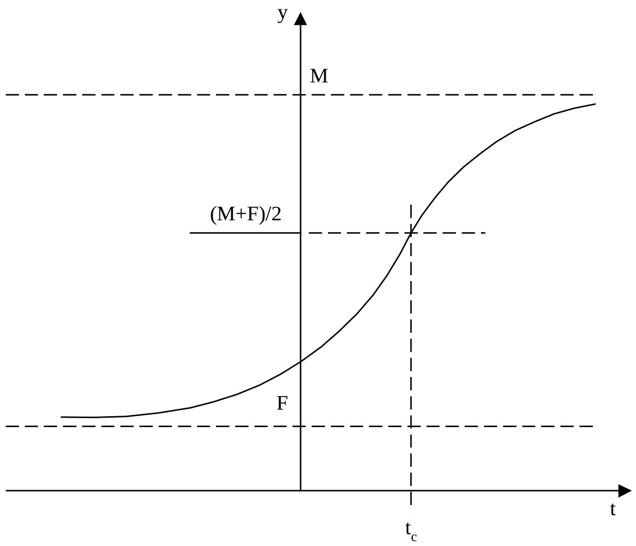


Figura 2.3/9 : Curva Logística

La curva tiene las siguientes propiedades:

i. El crecimiento aumenta constantemente hasta un pico y luego disminuye.

ii. El punto de inflexión para el cual la tasa de crecimiento absoluto cambia de un aumento a un descenso, ocurre a mitad de camino entre el máximo y el mínimo, en el tiempo $t = t_c$ donde t_c está dado por $1/b \ 1n\ (1/a)$.

Comentarios

El propósito de un análisis de tendencia no siempre es encontrar cómo el *promedio* de todos los valores de tráfico cambia con el tiempo. Para fines de dimensionamiento estamos particularmente interesados en cómo el extremo superior de la distribución anual de tráfico cambia con el tiempo, a fin de obtener un valor representativo para la hora de tráfico pico. Esta hora de tráfico pico determina luego el número de circuitos y conmutadores requeridos en el futuro. Para llevar a cabo una extrapolación significativa de una serie de tiempo es, por tanto, esencial que los datos utilizados se refieran a valores realmente representativos para nuestro propósito. Un grupo de datos históricos basados en minutos pagados por mes, por ejemplo, no merecería el esfuerzo de ningún análisis serio de series de tiempo. Ello es así porque la relación entre la hora tráfico pico y los minutos pagados mensualmente generalmente no está establecida por la experiencia, sino que tiene que calcularse como un producto de varios factores inexactos.

Otro caso en el que un análisis de series de tiempo no funcionará es cuando cambian las funciones operacionales. Por ejemplo, si el servicio de repente mejora notablemente el tráfico usualmente cambia de una forma drástica. Tales saltos en la mejora del servicio pueden significar un aumento en el tráfico de un 100% o más. Para estimar la magnitud de un aumento tan repentino en el tráfico se requiere información sobre el comportamiento del servicio antes y después del cambio, cuánto tráfico se suprimía antes del cambio, etc. Dichos datos generalmente no están disponibles o son más bien difíciles de reunir.

2.4 Juicio individual

El método de pronóstico matemáticamente más simple es el criterio personal. El pronóstico se basa en la experiencia y en la información recogida. No se hace ningún análisis sistemático.

Sin embargo, la calidad de un pronóstico depende directamente de la calidad de la información básica. A menos que se disponga de datos básicos fiables, el planificador debe establecer un proceso adecuado de recopilación de datos antes de comenzar a trabajar en su pronóstico, o más bien, asegurarse de contar con datos más fiables para su trabajo futuro.

Un ejemplo típico de pronósticos basados en juicio individual es estimar el número futuro de abonados en un área determinada. Generalmente estos pronósticos están relacionados con los edificios existentes y futuros. Cada tipo de edificio tiene su densidad telefónica típica. Los pronósticos a corto plazo pueden referirse a los edificios existentes, edificios en construcción y en planes de construcción. Los pronósticos a mediano y largo plazo pueden referirse también a diferentes clases de áreas, donde las áreas con edificios de oficinas requerirán más teléfonos que las áreas residenciales. El planificador puede tener como una ayuda tablas que muestren el número promedio de abonados en los diferentes tipos de edificios y áreas. Para llevar a cabo este tipo de investigación, es esencial que cuente con información correcta de las autoridades de planificación de la ciudad, así como su experiencia pasada.

En la Tabla 2.4/1 se muestra un ejemplo del número promedio de abonados para diferentes tipos de edificios. Naturalmente, la tabla se ajustará a la situación del país o ciudad considerados. La tabla 2.4/2 muestra valores típicos de áreas con diferente tipo de construcción para pronósticos a mediano y largo plazo. Aquí también deben ajustarse los valores de acuerdo al país en consideración. Las cifras que se dan en las tablas necesitan actualizarse cuando la densidad telefónica aumenta.

Tipo de edificio	Número de abonados
Edificios oficiales y oficinas, bancos, compañías aseguradoras, grandes hoteles, clubes, grandes restaurantes, hospitales, grandes almacenes.	investigar
Hoteles pequeños, restaurantes, tiendas de comida, casas de huéspedes.	1 - 2
Farmacias, médicos, abogados, etc.	1 - 1.5
Tiendas	0.5 - 1
Fábricas grandes	investigar
Fábricas pequeñas, talleres	0.5 - 1.5
Cines, estaciones de gasolina	1 - 2
Casas privadas de clase alta	1
Casas privadas de clase más baja	0.3 - 0.5
Viviendas unifamiliares	0.3 / apartamento
Bloques de apartamentos de clase alta	0.5 - 1 / apartamento
Bloques de apartamentos de clase más baja	0.2 / apartamento

Tabla 2.4/1 : Distribución del número de abonados por diferente clase de edificio. Este ejemplo es válido para un país en desarrollo con 4 a 8 abonados por cada 100 habitantes.

Tipo de área edificada		Abonados por hectárea
A.	Barrios marginales	0.25
B.	Parques, jardines, etc.	0.50
C.	Antiguas casas privadas con grandes jardines	1.00
D.	Distritos residenciales de trabajadores pobres	1.50
E.	Distritos residenciales de trabajadores más acomodados	2.00
F.	Casas modernas privadas con jardines grandes	3.00
G.	Distritos residenciales de trabajadores modernos	4.00
H.	Areas industriales	5.00
I.	Casas privadas modernas con jardines pequeños	7.00
J.	Casas no dispersas de tipo antiguo	8.00
K.	Areas de viviendas de clase trabajadora y pequeños talleres	10.0
L.	Casas modernas no dispersas	13.0
M.	Casas residenciales no dispersas de 1 a 2 plantas y pequeñas tiendas	18.0
N.	Bloques de apartamentos de hasta 4 plantas	25.0
O.	Bloques de apartamentos y tiendas de hasta 4 plantas	28.0
P.	Centros comerciales en áreas residenciales	30.0
Q.	Bloques de apartamentos de más de 4 plantas	40.0
R.	Edificios de oficinas de más de 3 plantas	80.0
S.	Edificios de oficinas de 4 a 6 plantas	150
T.	Edificios de oficinas de más de 6 plantas	250

Tabla 2.4/2 : Distribución del número de abonados por hectárea en áreas con diferentes clases de construcción : Este ejemplo es válido para un país en desarrollo, con 4-8 abonados por cada 100 habitantes.

2.5 Otros Métodos

El siguiente es un estudio de algunos otros métodos usados para pronósticos.

Comparaciones Analíticas

Este método se basa en los datos de otro método, elaborado de preferencia para la comparación de áreas con desarrollo más avanzado. El pronóstico se hace entonces por comparación y asumiendo que el desarrollo será el mismo.

Este método se utiliza cuando la serie de datos históricos adecuados es demasiado corta para permitir un análisis de tendencia fiable. Este es un caso frecuente de los países en desarrollo. Se asume entonces que el desarrollo seguirá el mismo patrón que las regiones o áreas más desarrolladas.

La Tabla 2.5/1 muestra el crecimiento de la penetración telefónica en el período 1960-1975, expresado en el número de teléfonos por cada 100 habitantes, para ciertos países. Este muestra también en qué período el Reino Unido se desarrolló con la misma tasa y cuántos años le tomó alcanzar la misma penetración que tuvieron los otros países de 1960 a 1975.

País	Teléfonos por 100 habitantes		Penetración correspondiente en el R. U.	
	1960	1975	Fechas	Período
Argentina	5.99	9.41	12 años	1937-1949
Chile	2.43	4.26	8 años	1923-1931
Costa Rica	1.34	5.02	25 años	1910-1935
Chipre	2.95	10.65	25 años	1925-1950
Fiji	1.71	4.53	18 años	1914-1932
Grecia	2.30	20.71	45 años	1922-1967
Irlanda	5.13	12.78	20 años	1935-1955
Japón	5.21	37.88	41 años	1935-1976
México	1.46	4.37	20 años	1911-1931
España	5.47	12.46	18 años	1936-1954

Tabla 2.5/1 : Número de teléfonos por cada 100 habitantes para ciertos países

Tabla 2.5/1 muestra que no hay una clara conexión entre el tiempo que demoró a un país particular desarrollar desde el valor de 1960 hasta el de 1975 y el tiempo que necesitó el Reino Unido para alcanzar el mismo grado de desarrollo (crecimiento). Por lo tanto no es recomendable comparar países con gran diferencia en el desarrollo. La tabla 2.5/2 da la misma información que la tabla anterior, pero para un grupo de países con aproximadamente el mismo desarrollo económico.

País	Teléfonos por 100 habitantes		Fechas equivalentes para el R. U.	
	1960	1975		
Australia	20.88	37.49	9 años	1969-1976
Dinamarca	22.17	42.48	10 años	1968-1978
Finlandia	12.89	35.78	20 años	1955-1975
República Federal de Alemania	9.98	30.25	24 años	1949-1973
Islandia	21.42	40.41	10 años	1967-1977
Holanda	13.15	34.41	19 años	1955-1974
Noruega	19.50	33.90	8 años	1966-1974

Tabla 2.5/2 : Penetración telefónica 1960-1975 en algunos países con aproximadamente el mismo desarrollo económico.

Esta tabla muestra una mejor armonía con el desarrollo en el Reino Unido que la tabla anterior.

Ahora podemos concluir lo siguiente:

1. La proyección a corto plazo de la densidad telefónica de un país desde un nivel dado, en comparación con el crecimiento de algún país más avanzado que estuvo antes en el mismo nivel, no siempre puede ser recomendable debido a que circunstancias especiales pueden causar que el desarrollo no siga la misma tendencia. Tales circunstancias son, por ejemplo: recesión o expansión económica, limitaciones en los suministros, diferentes políticas tarifarias, etc.

2. Para proyecciones a largo plazo es necesario escoger un país que esté por lo menos 15 años más avanzado, a fin de tener una serie suficientemente larga para el posible crecimiento futuro.

3. Si tal comparación se lleva a cabo, no obstante, se asume que es válida una comparación entre las características de un país en los últimos 15 años y las de otro país en los 15 años anteriores. Esto es improbable por el rápido cambio en la economía mundial y debido a la posibilidad de sucesos desestabilizadores tales como guerras y grandes recesiones, que distorsionan la comparación. Los datos confirman esta afirmación.

Otros Métodos:

Método de la inspección ocular

El método consiste simplemente en que el planificador inspecciona los datos graficados a fin de darse una idea de alguna posible tendencia o relación entre las variables y luego trata de dibujar manualmente la curva correspondiente.

Promedios móviles

El nuevo valor se calcula como la media de un número de valores observados. El propósito de este enfoque es reducir las irregularidades causadas, por ejemplo, por variaciones estacionales.

Suavizado exponencial

Es un método similar al de los promedios móviles, pero con ponderación, de modo que las observaciones más recientes son las que mayor peso tienen.

Investigación de mercado

Este método consiste en estudiar el mercado por medio de preguntas a los clientes potenciales.

Métodos combinados (incluyendo análisis estructural)

Una combinación de los diferentes métodos tales como: investigación de mercado, curvas de crecimiento, modelos de proyección y econométricos.

3. PRONOSTICO PARA LA PLANIFICACION DE CENTRALES

Para la planificación de centrales, la demanda debe pronosticarse para todos los ítems particulares de cada central por separado. Esto incluye lo siguiente, por categoría de abonado:

- número de líneas directas de abonados (DEL);
- tráfico de origen;
- tráfico de destino;
- tráfico local;
- tráfico de larga distancia;
- tráfico internacional;
- servicios especiales.

Una vez que estos pronósticos se han efectuado, pueden sumarse las cifras correspondientes de las diferentes centrales para producir resultados para áreas más grandes o para todo el país. Esto se conoce como “pronóstico de abajo hacia arriba”.

Cuando se tiene que hacer un pronóstico para determinada área, debe ser reunirse toda la información disponible acerca de futuros cambios en dicha área. La información relevante que se obtiene de esta forma puede hacer posible prever discontinuidades en el desarrollo de la demanda. Por otro lado, las fluctuaciones de las cantidades de interés son relativamente grandes cuando analizamos un área pequeña. El número de registros históricos fiables también es usualmente muy limitado.

Estos dos factores combinados hacen que el pronóstico de las pequeñas áreas sean menos fiables. La suma de los resultados para obtener cifras que correspondan a un área más extensa puede conducir a graves errores. Para evitarlo se compara el pronóstico "*de abajo hacia arriba*" con un pronóstico *"de arriba hacia abajo"*.

Comenzando por el pronóstico para todo el país o para un área grande, uno puede llegar a pronósticos para regiones más pequeñas, dividiendo las cantidades pronosticadas. Usualmente se dispone de información más amplia y fiable para el pronóstico de grandes áreas. Por tanto, el pronóstico *"de arriba hacia abajo"* generalmente se considera más fiable que el pronóstico *"de abajo hacia arriba"*.

Hay dos formas para lograr la concordancia entre los dos pronósticos:

1. Cambiando la división de la cantidad total entre las diferentes áreas.

2. Reduciendo los pronósticos totales por un factor dado, de manera uniforme o en base a algún sistema.

Es evidente que este procedimiento de ajustar los dos pronósticos es más bien subjetivo. También es notorio que un área que haya tenido un alto pronóstico tendrá que reducir sus cifras, a menos que haya una razón satisfactoria para este alto pronóstico. Tal razón puede ser un plan de desarrollo conocido para el área, que motivaría un aumento súbito en la demanda. El pronóstico de una área grande no siempre toma en cuenta tales peculiaridades locales.

4. PRONOSTICO PARA LA PLANIFICACION DE REDES

4.1 *Introducción*

La planificación de una red telefónica se basa en las estimaciones sobre la necesidad de tráfico futuro. Se necesita un *pronóstico a largo plazo* para que el plan de desarrollo asegure una ampliación coordinada durante un período de 15 a 25 años. El plan de desarrollo debe actualizarse aproximadamente cada 2 a 4 años.

Dentro del plan de desarrollo se necesitan *pronósticos a corto plazo* para proporcionar los datos básicos para planificar los pasos de las ampliaciones actuales. Ellos deben contener estimaciones de las necesidades de tráfico para los próximos 4 a 6 años, a ser satisfechos por los tiempos de entrega, etc. Los pronósticos a corto plazo deben actualizarse cada año.

Las necesidades de tráfico se expresan en erlangs. Para pronosticar estas necesidades se estima el tráfico entre cada par de centrales, generalmente por separado para cada dirección. También se estima el tráfico dentro cada área de central.

4.2 *Matriz de tráfico*

Para especificar las necesidades de tráfico en una región con n centrales se requieren n^2 valores de tráfico. Una manera estandarizada de especificar estos tráficos es presentarlos en una matriz, la llamada matriz de tráfico.

hacia desde	1	i	j	n	SO
1	A(11)			A(1n)	O(1)
i		A(ii)	A(ij)		O(i)
j		A(ji)	A(jj)		O(j)
n	A(n1)			A(nn)	O(n)
ST	T(1)	T(i)	T(j)	T(n)	A(11)

Aquí:

A(ij) es el tráfico de i a j;

A(ji) es el tráfico de *j* a *i;*

A(ii) es el tráfico local en la central *i*;

O(i) es la suma de todo el tráfico que se origina en *i*;

T(j) es la suma de todo el tráfico terminando en *j*.

Sumando los totales de las filas O(i), es decir, las entradas en la columna SO (suma de tráfico de origen) obtenemos el tráfico total A. El mismo resultado se obtiene sumando los totales de todas las columnas T(j), esto es, las entradas en la fila ST (suma del tráfico de destino). En resumen:

$$\sum_i O(i) = \sum_j T(j) = A..$$

Mientras no haya confusión, se puede usar el símbolo *A(i, j)*. Sin embargo, con frecuencia será necesario distinguir entre el tráfico actual desde *i* a *j A(i, j/0)* y el tráfico estimado en alguna fecha futura *t*: *A(i, j/t)*. Luego,

$$O(i/t) = \sum_j A(ij/t) \qquad \text{y} \qquad T(j/t) = \sum_i A(ij/t)$$

4.3 <u>*Pronóstico punto a punto*</u>

Existen varios métodos para predecir *A(i,j/t)* basados en el crecimiento esperado del número de clientes en las áreas *(y)* y *(j),* en los cambios esperados en el tráfico por abonado, etc. Entonces puede completarse la matriz de tráfico añadiendo las entradas por filas para obtener O(i/t) y por columnas para obtener T(j/t).

Para estimar los tráficos futuros punto a punto en una red, usualmente uno se puede basar en los cálculos del crecimiento previsto en las líneas de abonados y en la matriz de tráfico actual. Se usan fórmulas diferentes, de las cuales las más comunes se dan a continuación. No se puede sostener que una fórmula sea más exacta que la otra. Solamente la realimentación de los registros futuros puede señalar cuál fórmula es la mejor para un caso particular. No obstante esto se haya descubierto, no hay garantía de que siempre será así .

Estimación del tráfico total

Tomando en cuenta que categorías diferentes de abonados generan cantidades diferentes de tráfico, a veces será posible estimar el tráfico futuro a partir de:

$$A(t) = N_1(t)\cdot\alpha_1 + N_2(t)\cdot\alpha_2 + \mathrm{K} \tag{4.3.1}$$

donde $N_1(t)$, $N_2(t)$, etc., son el número de abonados en los tiempos 1, 2, etc., y α_1, α_2, etc., es el tráfico por cliente por categoría 1, 2, etc.

Si no es posible separar los abonados en categorías con tráfico diferente, el tráfico futuro puede estimarse sencillamente como:

$$A(t) = A(0)\cdot\frac{N(t)}{N(0)} \tag{4.3.2}$$

donde *N(t)* y *N(0)* son el número de abonados en los tiempos *t* y cero.

Estimación del tráfico punto a punto

Para estimar el tráfico de una central a otra se pueden aplicar varias fórmulas. La idea principal es tomar en cuenta el incremento de abonados en ambas centrales y aplicar ciertos factores de ponderación a estos crecimientos:

$$A_{ij}(t) = A_{ij}(0)\frac{W_iG_i + W_jG_j}{W_i + W_j} \tag{4.3.3}$$

donde W_i y W_j son las ponderaciones y G_i es el crecimiento de abonados en la central *i,* y G_j en la central *j*.

$$G_i = \frac{N_i(t)}{N_i(0)} \qquad G_j = \frac{N_j(t)}{N_j(0)} \tag{4.3.4}$$

Existen métodos diferentes para estimar W_i y W_j .

Primera Fórmula de Rapp

$$W_i = N_i(t) \qquad W_j = N_j(t) \tag{4.3.5}$$

El supuesto aquí es que el tráfico por abonado de la central *i* a la central *j* es proporcional al número de abonados en la central *j*.

Segunda Fórmula de Rapp

$$W_i = N_i(t)^2 \qquad W_j = N_j(t)^2 \tag{4.3.6}$$

Esta fórmula supone que el cambio del tráfico de origen y terminado por abonado es tan pequeño como sea posible.

Fórmula de Australian Telecom

$$W_i = \frac{N_i(0) + N_i(t)}{2} \qquad W_j = \frac{N_j(0) + N_j(t)}{2} \tag{4.3.7}$$

Esta fórmula es una modificación de la Primera Fórmula de Rapp.

Una cuarta fórmula se deduce del siguiente supuesto: el tráfico de un abonado en la central *i* a todos los abonados en la central *j* es constante.

$$\begin{gathered} \frac{A_{ij}(t)}{N_i(t) \cdot N_j(t)} = \frac{A_{ij}(0)}{N_i(0) \cdot N_j(0)} \\ A_{ij}(t) = A_{ij}(0) \cdot G_i \cdot G_j \end{gathered} \tag{4.3.8}$$

Estas cuatro fórmulas pueden ajustarse posteriormente introduciendo *N* en los factores de ponderación.

Modelo Gravitacional

El tráfico entre dos centrales se puede expresar como:

$$A_{ij} = K(d_{ij}) \cdot N_i \cdot N_j \tag{4.3.9}$$

donde $K(d_{ij})$ = factor de comunidad de interés.

Se ha encontrado que este factor depende, hasta cierto punto, de las distancias entre las centrales d_{ij} Podemos, entonces, escribir $K(d_{ij})$ así:

$$K(d_{ij}) = e^{-\gamma \cdot d_{ij}} \tag{4.3.10}$$

ó

$$K(d_{ij}) = d_{ij}^{-g} \tag{4.3.11}$$

El valor de los parámetros γ ó g puede calcularse a partir de una matriz de tráfico conocida. Puede ser necesario ajustar la expresión para A_{ij} para pares de centrales con relaciones especiales entre ellas; por ejemplo: una fábrica grande en una parte del país y la oficina principal en otra parte.

Las fórmulas dadas aquí son todas deducidas bajo ciertos supuestos. Un supuesto es que el tráfico por abonado permanece constante y otro es que un abonado tiene el mismo interés en llamar a cada uno de los otros abonados. Sin embargo, en países desarrollados se ha encontrado que los nuevos abonados pueden tener menos tráfico que los viejos. Esto se ha tomado en cuenta parcialmente en la fórmula australiana.

En principio es imposible decir que una fórmula es más fiable que otra. El planificador debe intentar fórmulas diferentes y tratar de encontrar cuál le da el resultado mas creíble para su caso en particular.

4.4 *Método del doble factor de Kruithof*

El método de Kruithof nos permite estimar los valores individuales del tráfico futuro, $A(i,j)$ en una matriz de tráfico. Los valores actuales se asumen conocidos, así como los valores futuros de las sumas de filas y de columnas.

El procedimiento es ajustar los valores individuales $A(i,j)$ de modo que estén de acuerdo con las nuevas sumas de filas y de columnas, es decir:

$$A(i,j) \text{ se cambia a } A(i,j)\cdot\frac{S_1}{S_0}$$

donde S_0 es la suma actual y S_1 es la nueva suma para la fila o columna individual.

Si empezamos con ajustar los valores $A(i, j)$, con respecto a las nuevas sumas de filas, S_j , estas sumas coincidirán, pero no así las sumas de columnas. El próximo paso será entonces ajustar los valores encontrados de $A(i,j)$ para que coincidan con las sumas de columnas. Esto hará que las sumas de filas no concuerden, entonces el siguiente paso será ajustar los nuevos valores de $A(i,j)$ para que coincidan con las sumas de filas. Este procedimiento seguirá hasta alcanzar la exactitud suficiente tanto de las sumas de filas como de columnas. La iteración es más bien rápida y proporciona generalmente un resultado satisfactorio luego de más o menos tres correcciones.

Este procedimiento se comprenderá mejor mediante el ejemplo numérico que se da a continuación

El método es aplicable cuando no se espera mucho cambio en las proporciones entre los valores individuales $A(i,j)$, y también en casos donde no es posible predecir los valores individuales $A(i,j)$ de otra manera.

Ejemplo del uso del Método del Doble Factor de Kruithof

Considere una red telefónica con dos centrales:

Dado:

1. Los intereses de tráfico actuales $A_{ij}(0)$

i j	1	2	sum
1	10	20	30
2	30	40	70
suma	40	60	100

(4.4.4)

2. Pronosticar los tráficos totales de origen y de destino futuros por cada central:

$A_{i.}(t)$ y $A_{.j}(t)$:

i j	1	2	suma
1			45
2	?		105

suma	50	100	150	(4.4.5)

Problema:

Estimar los valores de tráfico *A(i, j/t)* con el método de Kruithof.

Solución:

Iteración 1: Multiplicación de filas.
$A_{i.}$ se distribuye según el interés de tráfico actual.

Resultado:

i j	1	2	suma
1	15	30	45
2	45	60	105
suma	60	90	150

$$A_{ij}(1) = \frac{A_{ij}(0)}{A_{i.}(0)} \cdot A_{i.}(t) \qquad (4.4.6)$$

Después de la multiplicación de filas, las sumas de columnas difieren del pronóstico. La próxima iteración será la multiplicación por columnas.

Iteración 2: Multiplicación de columnas.
$A_{.j}(2)$ se distribuye según indica la iteración 1.

Resultado:

i j	1	2	suma
1	12.5	33.33	45.83
2	37.5	66.67	104.17
suma	50	100	150

$$A_{ij}(2) = \frac{A_{ij}(1)}{A_{.j}(0)} \cdot A_{.j}(t) \qquad (4.4.7)$$

Después de la multiplicación de columnas, las sumas de las filas difieren de los valores pronosticados. La siguiente iteración será la multiplicación de filas.

Iteración 3: Multiplicación de Filas.
$A_{i.}$ se distribuye como indica la iteración 2.

Resultado:

i j	1	2	suma
1	12.27	32.73	45
2	37.80	67.20	105
suma	50.07	99.93	150

$$A_{ij}(3) = \frac{A_{ij}(2)}{A_{i.}(0)} \cdot A_{i.}(t) \qquad (4.4.8)$$

Iteración 4: Multiplicación de Filas.
$A_{.j}$ se distribuye como indica la iteración 3.

Resultado:

i j	1	2	suma
1	12.25	32.75	45
2	37.75	67.25	105
suma	50	100	150

$$A_{ij}(4) = \frac{A_{ij}(3)}{A_{.j}(0)} \cdot A_{.j}(t) \tag{4.4.9}$$

Después de 4 iteraciones las sumas de las filas y las columnas son iguales a los valores previstos. Ahora podemos escribir:

$$A_{ij}(t) = A_{ij}(4) \tag{4.4.10}$$

5. CONCLUSIONES

Este capítulo ha tratado principalmente con las matemáticas que están detrás de todos los métodos de pronóstico. Esto hará posible al lector colaborar en el trabajo de pronóstico, especialmente en lo concerniente a los tratamientos preliminares de datos históricos.

Además de explicar cuándo y por qué los pronósticos son necesarios, este capítulo contiene algunos métodos frecuentemente usados para la descripción de datos históricos en fórmulas matemáticas. Nuevamente se debería de hacer énfasis en que *el tratamiento de los datos históricos no es pronóstico*. Sólo es tratamiento de datos históricos. Por otra parte, el pronóstico es el arte de predecir el desarrollo futuro. Sólo en ciertos casos para propósitos de corto plazo, la extrapolación de curvas obtenidas de los datos históricos proporcionarán pronósticos fiables. En la mayoría de los casos debe juzgarse si las condiciones futuras y supuestos son los mismos del pasado. Los ajustes de los valores extrapolados pueden ser necesarios por varias razones. La mejora del funcionamiento del servicio puede duplicar el tráfico dentro de un mes o dos. Nuevas actividades en la sociedad (por ej. industrias) pueden causar cambios drásticos en la dispersión del tráfico, etc.

El pronóstico generalmente concierne a la *demanda* para suscripciones y tráfico ocasionado por varios servicios de telecomunicaciones. Ya que las demandas de cantidades son difíciles o imposibles de medir, los datos históricos dados al pronosticador son generalmente cantidades mensurables, tales como el tráfico cursado. Por eso, el planificador antes que nada debe primero transferir los datos dados en datos de demanda. Si ellos se refieren a un país y a un servicio donde la demanda está satisfecha, ésta puede estimarse con precisión aceptable.

Si la demanda es mucho mayor que la oferta es bien difícil estimar la demanda real. El planificador tiene que tomar en cuenta esta dificultad mientras las condiciones no permitan un equilibrio entre la oferta y la demanda.

Otra dificultad en el uso de datos históricos es que ellos no son suficientemente fiables como para permitir un análisis. Una razón para esto es que los datos se hayan recogido bajo diferentes fundamentos, para diferentes propósitos y en diferentes tiempos. Inclusive puede suceder que no se disponga de ningún dato histórico. El pronóstico entonces debe basarse en otros indicadores indirectos, lo cual generalmente no mejora la fiabilidad del pronóstico.

6. REFERENCIAS

BEAR, D. Algunas teorías sobre Distribución de Tráfico Telefónico, Un estudio crítico (Documento ITC, Estocolmo, 1973).

LEIJON, K.H. Introducción a Ingeniería de Teletráfico Práctica. Sección F. Pronóstico (UIT) REM/72/038, Benghazi, 1976.

CCITT (GAS 5) Métodos Usados en el Pronóstico de Largo Plazo de la Demanda Doméstica de Telecomunicaciones y Recursos Requeridos (1980).

LONNQUIST, I.	Teoría y Métodos de Pronóstico de Tráfico (TELE, Estocolmo, 1968).
LONNQUIST, I.	Seminario sobre Pronóstico en Ingeniería de Teletráfico, Estambul, 1980 (UIT).
LONNSTROM, M.M.	Un proyecto de Desarrollo Telefónico, Estocolmo, 1967.
MARKLUND, F.	
MOO, I.	
NYEAARD, J.	Una Herramienta en el Pronóstico de Tráfico Telefónico Integrado para Uso en Planificación, I. (Teleteknik Vol. XIV, 1970, No. 1).
RAPP, Y.	Algunos aspectos Económicos sobre la Planificación a largo plazo de Redes Telefónicas. (Estocolmo, 1969).
TURNER, W.M.	Pronóstico. Seminario UIT/SIDA, Nueva Delhi, 1975 SEM/ASIA/75.
CCITT (GAS 5)	Estudios Económicos a Nivel Nacional en el Campo de las telecomunicaciones.

Nota: Este capítulo ha sido editado de un manuscrito preparado por K.H. Leijon, L.M. Ericsson, Suecia.

7. EJERCICIOS

TECNICAS DE PRONOSTICO

1. Un objeto va a ser pronosticado. El desarrollo histórico ha sido el siguiente:

Año (al final de)	Escala de Tiempo (t)	Stock (y)	Aumento Absoluto	%
1968	1	583		
1969	2	615	32	5.5
1970	3	646	31	5.0
1971	4	697	51	7.9
1972	5	738	41	5.9
1973	6	802	64	8.7
1974	7	844	42	5.2

Hacer un pronóstico del stock esperado dentro de 5 y 10 años, de cuatro diferentes maneras:

a. Asumiendo un aumento *absoluto* constante e invariable por año, basado en el incremento histórico promedio.

b. Asumiendo un *porcentaje* de incremento anual constante e invariable, basado en el incremento total de los pasados seis años.

c. Adaptando una línea de tendencia lineal a los datos históricos.

d. Adaptando una tendencia exponencial a los datos históricos.

Representar los datos históricos en un diagrama y mostrar en los mismos diagramas las dos líneas de tendencia c) y d) ampliadas al año 16 (6 + 10), y también los otros dos pronósticos a) y b).

Comparar el resultado de los dos pronósticos.

Verificar las líneas de tendencia en lo que concierne a la fiabilidad estadística. (Coeficiente de correlación, prueba T y prueba de Durbin - Watson).

Discutir la credibilidad de los pronósticos.

2. Calcular el pronóstico de todas las conexiones para una red telefónica local para 1983 (31.3.83)

Dados:

CUADRO DE CRECIMIENTO DE DEMANDAS EN UNA RED TELEFONICA LOCAL.

Fecha	Capacidad equipada	Conexiones en funcionamiento	Lista de espera	Demanda total
31/3/72	11800	9825	1901	11,726
31/3/73	12000	11114	1781	12,925
31/3/74	12100	11458	3555	15,013
31/3/75	12500	11530	5106	16,636
30/9/75	12500	11653	5662	17,315
31/3/76	12800	12275	2587	14,862
31/3/77	14150	13683	1624	15,307
31/3/78	15500	14437	1893	16,330

3. Pronosticar el tráfico total de origen para diciembre de 1986 en una central telefónica local.

Dados:

a. Registros de tráfico desde enero de 1979 hasta diciembre de 1981 (Tabla A).

b. Registros de tasa de llamadas desde enero de 1979 hasta diciembre de 1981 (Tabla B).

c. Pronóstico de conexiones para diciembre de 1986 = 2969
(1170 abonados en enero 1979 y 1600 en diciembre 1981)

A. TRAFICO TOTAL DE ORIGEN

MES	1979	1980	1981
ENERO	38.6	39.4	45.6
FEBRERO	37.9	43.7	46.2
MARZO	42.1	48.7	47.2
ABRIL	40.6	43.8	46.2
MAYO	40.1	40.2	45.6
JUNIO	38.1	42.6	48.5
JULIO	37.7	41.1	44.4
AGOSTO	39.9	44.2	47.4
SEPTIEMBRE	40.4	41.0	49.1
OCTUBRE	40.7	43.8	48.7
NOVIEMBRE	40.8	41.8	45.0
DICIEMBRE	42.2	49.5	49.5

B. TASAS TOTALES DE LLAMADAS DE ORIGEN (Erlang/Abonado)

MES	1979	1980	1981
ENERO	.033	.031	.033
FEBRERO	.033	.034	.033
MARZO	.036	.038	.034
ABRIL	.035	.033	.033
MAYO	.034	.030	.032
JUNIO	.032	.032	.035
JULIO	.031	.031	.032
AGOSTO	.033	.032	.034
SEPTIEMBRE	.034	.030	.035
OCTUBRE	.034	.032	.032
NOVIEMBRE	.033	.030	.028
DICIEMBRE	.034	.036	.031

4. Considere una red telefónica con dos centrales.

Dados:

a. Los intereses de tráfico actuales $[A_{ij}(0)]$

j i	1	2	0_2 (o)
*	10	20	30
2	30	40	70
T_j (o)	40	60	100

b. Los valores futuros previstos para los tráficos totales de origen y de destino por central [0; (t) T; (t)]

j i	1	2	o_i (E)
1	?	?	45
2	?	?	105
T_j (t)	50	100	150

Estimar los valores de tráfico $A_{ij}^{(t)}$ usando el método de Kuithof.

5. Calcular los intereses de tráfico entre dos centrales en un área local.

Dados:

a. Un área local está dividida en *áreas de tráfico* No. 1,2, 3, 4... Los tráficos futuros entre todas las *áreas de tráfico* están pronosticados.

b. El área local se divide en áreas de central. Las *áreas de central* no coinciden con las *áreas de tráfico*.

c. Tratar de calcular el tráfico futuro esperado desde una central A hasta la central B.

d. Tenemos la siguiente información:

La central *A* tendrá:

- 5000 líneas de abonados del área de tráfico I, que tiene en total 10,000 líneas de abonados.
- 8000 líneas de abonados del área de tráfico II, que tiene en total 12,000 líneas de abonados.

La central *B* tendrá:

- 9000 líneas de abonados del área de tráfico III, que son todas las líneas de abonado que hay allí.
- 2000 líneas de tráfico del área de tráfico IV, que tiene un total de 6000 líneas de abonados.

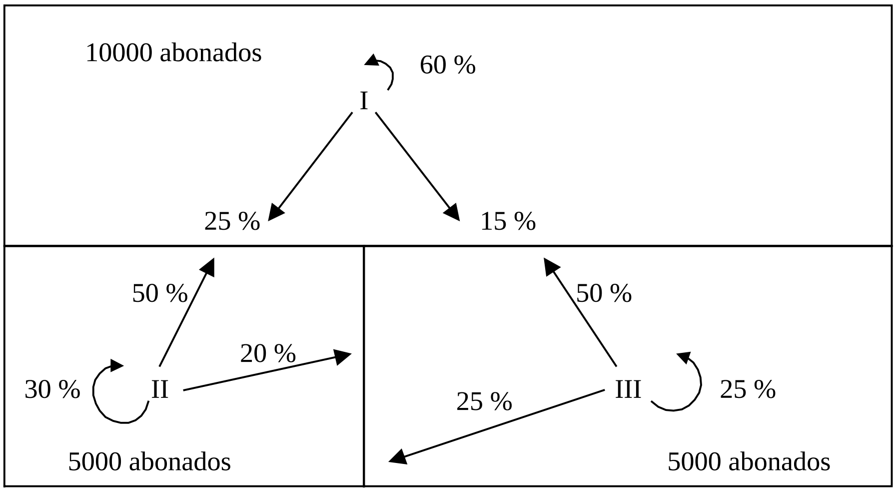
e. Por el pronóstico también sabemos que el tráfico total esperado desde cada área de tráfico hacia cada una de las áreas involucradas es el siguiente:

Del área de tráfico No.	Al área de tráfico No.	Tráfico total esperado
I	III	100 erl.
I	IV	90 erl.
II	III	105 erl.
II	IV	95 erl.

6. Estimar el interés de tráfico entre centrales: El objetivo del ejemplo es ilustrar cómo una matriz de tráfico para centrales puede calcularse desde una matriz de tráfico dada para *áreas de tráfico*.

Dada:

Un área consistente en tres áreas de tráfico I, II, III:



Area de Tráfico 1

Número de abonados	10,000	
Tráfico total originado/abonado	0.06 erl.	
Distribución de este tráfico	60 %	I
	25 %	II
	15 %	III

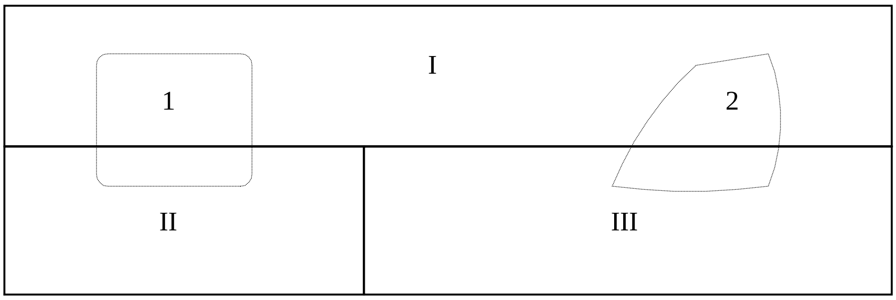
Area de Tráfico II

Número de abonados	5,000	
Tráfico total originado/abonado	0.05 erl.	
Distribución de este tráfico	50 %	I
	30 %	II
	20 %	III

Area de Tráfico III

Número de abonados	5,000	
Tráfico total originado/abonado	0.04 erl.	
Distribución de este tráfico	50 %	I
	25 %	II
	25 %	III

El área está servida por un número de centrales 1, 2, ... etc.



Central 1

Número total de abonados	8,000
No. de abonados pertenecientes al área de tráfico I	5,000
No. de abonados pertenecientes al área de tráfico II	3,000

Central 2

Número total de abonados	6,000
No. de abonados pertenecientes al área de tráfico I	4,000
No. de abonados pertenecientes al área de tráfico III	2,000

Tarea:

Calcular el flujo de tráfico total esperado de la central 1 a central 2.

Sugerencia:

Comenzar con el cálculo de tráfico de *un* abonado en el área de tráfico I a *un* abonado en el área de tráfico II, etc.

7. Se ha estimado la matriz de tráfico actual $A_{ij}^{(0)}$:

Hacia Desde	Central No. j			
	1	2	3	Suma
1	25	30	45	100
2	35	55	110	200
3	60	85	155	300
Suma	120	170	310	600

Se ha pronosticado el número de líneas principales por central en el año *t* :

Central No.	N_i (0)	N_i(t)
1	2000	3000
2	3500	3500
3	6800	7500

Las líneas principales no se han clasificado en diferentes categorías ya que se espera que la proporción de los diferentes abonados sea la misma en el futuro.

Por eso el tráfico total de origen y de destino por central se ha previsto mediante las fórmulas siguientes:

$$A_{i.}^{(t)} = N_i^{(t)} \cdot \frac{A_{i.}^{(0)}}{N_i^{(0)}}$$

$$A_{.j}^{(t)} = N_j^{(t)} \cdot \frac{A_{.j}^{(0)}}{N_j^{(0)}}$$

Central No.	$A_i^{(t)}$	$A_j^{(t)}$
1	150.0	180.0
2	200.0	170.0
3	331.9	341.9
Suma	681.9	691.9

Ya que la suma de $A_{i.}^{(t)}$ y $A_{.j}^{(\ )}$ difieren, podemos usar el valor medio de estas sumas como una estimación de $A_{..}^{(t)}$ y ajustar $A_{i.}^{(t)}$ y $A_{.j}^{(\ )}$. Esto da:

Central No.	$A_i^{(t)}$	$A_{.j}^{(t)}$
1	151.1	178.7
2	201.5	168.8
3	334.3	339.4
Suma	686.9	686.9

a. Dibujar ahora una matriz de tráfico para el año *t* futuro y rellenarla con los tráficos totales calculados anteriormente.

b. Calcular los diferentes tráficos punto a punto usando el método de crecimiento ponderado, a partir de los datos conocidos acerca del número actual y futuro de líneas principales y los tráficos actuales punto a punto. El tipo de ponderación es de su elección.

c. Colocar estos valores en una nueva matriz de tráfico y calcular las sumas de las columnas y filas. Se podrá apreciar que estas sumas no concuerdan con los valores de la primera matriz.

d. Si consideramos que el tráfico por línea principal será constante durante el período de pronóstico, entonces los tráficos punto a punto obtenidos deben corregirse, de modo que sus sumas coincidan con los tráficos totales pronosticados.

Si el tiempo lo permite, hacer esa corrección empleando el Método del Doble Factor de Kruithof.

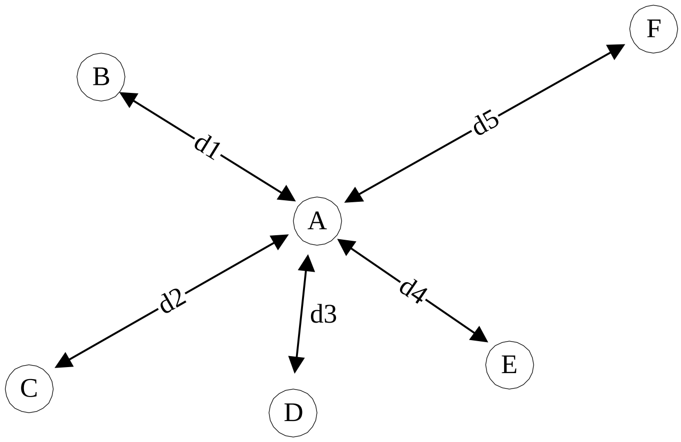
8. Suponer que un pronóstico de tráfico de troncales durante la hora pico para una central "A" es de 75 erlangs, que se distribuyen en cinco rutas hacia las centrales "B", "C", "D", "E" y "F", para las cuales no se dispone de datos de tráfico. Suponer también que los pronósticos de conexiones para las centrales son:

A	10 000	(c_0)
B	5 000	(c_1)
C	7 000	(c_2)
D	4 000	(c_3)
E	2 000	(c_4)
F	10 000	(c_5)

Y que las distancias desde "A" son:

B	20 millas	(d_1)
C	30 millas	(d_2)
D	10 millas	(d_3)
E	10 millas	(d_4)
F	50 millas	(d_5)

La situación se ilustra en el siguiente diagrama.



Estimar la distribución del tráfico troncal total en las cinco diferentes rutas de *A* hacia *B*, de *A* hacia *C*, ...,etc.

Sugerencia:

El pronóstico de tráfico para una ruta puede obtenerse a partir de:

$$t_i = \frac{\dfrac{c_0 \cdot c_i}{d_i^2}}{\displaystyle\sum_{j=1}^{n} \frac{c_0 \cdot c_j}{d_j^2}}$$

donde *T* es el tráfico total.