

TRABAJOS DE
INVESTIGACION OPERATIVA
Vol. 1. Núm. 1, 1986, pp. 61 a 72

**ESTUDIO BAYESIANO DE LOS PERIODOS
DE OCUPACION Y DESOCUPACION EN UNA COLA
M/M/1/∞/FIFO EN EQUILIBRIO (*)**

Carmen Armero Cervera
*Departamento de Bioestadística (**)*
Facultad de Medicina
de la Universidad de Valencia
Avda. Blasco Ibáñez, 17
Valencia 10

RESUMEN

En este artículo, para un modelo de colas $M/M/1/\infty/FIFO$ en equilibrio, se obtiene la distribución predictiva del tiempo de duración de un período de ocupación, y de desocupación de la cola, así como la distribución predictiva final del número de personas atendidas en un período de ocupación, y la probabilidad de que éste sea finito. Finalmente, dichos resultados se aplican en una línea de espera concreta.

Palabras clave: Inferencia Bayesiana; Intensidad de tráfico; Número de personas atendidas en un período de ocupación; Períodos de ocupación y desocupación de la cola; Tiempos de servicio; Tiempos transcurridos entre dos llegadas consecutivas a la cola.

Clasificación AMS: 60K25-90B22.

SUMMARY

We provide the predictive distributions for the length of an idle and busy period in an $M/M1/\infty/FIFO$ system at steady-state. We also obtain the posterior distribution of the number of customers served in a busy period, and the probability that a busy period is finite. We illustrate these results with an example.

(*) Recibido, marzo 1985. Aceptado, febrero 1986.

(**) Actualmente Departamento de Estadística e Investigación Operativa Facultad de Matemáticas/Universitat de Valencia/Doctor Moliner, 50/46100 Burjassot/Valencia.

Key words: Bayesian Inference; Busy and Idle period; Interarrivals times; Number of customers served in a busy period; Service times; Traffic intensity.

AMS Classification: 60K25-90B22.

1. INTRODUCCION

Desde un punto de vista práctico, la Teoría de la Decisión y la Inferencia Estadística son herramientas imprescindibles para el análisis y diseño de un sistema en cuyo funcionamiento está involucrada una o varias líneas de espera.

Adoptando la Metodología Bayesiana, se hace necesaria, para su posterior análisis, la obtención de la distribución predictiva de las distintas variables que pueden observarse en el modelo de colas seleccionado.

En este artículo se estudian los períodos de ocupación y desocupación para una cola $M/M/1/\infty/FIFO$, con solución estacionaria para el proceso de nacimiento-muerte que la describe. En concreto, y a partir de las hipótesis iniciales descritas en la subsección 2.1, se obtiene en la subsección 2.2 la distribución predictiva final del tiempo de duración de un período de desocupación, y la distribución predictiva final del tiempo de duración de un período de ocupación; en la subsección 2.3 se determina la distribución predictiva final del número de personas, en general, clientes de la cola, atendidos en un período de ocupación, así como la probabilidad de que éste sea finito.

En la sección 3 se aplican los resultados obtenidos a una de las colas del servicio de urgencias del Hospital Clínico de Valencia, comparándolos con las correspondientes distribuciones obtenidas desde el punto de vista de la Inferencia Clásica.

2. PROCESO INFERENCIAL

2.1. Hipótesis previas

Sea $z_1 = \{x_1, x_2, \dots, x_n\}$ una muestra aleatoria de los tiempos de servicio de la cola, obtenida a partir del modelo exponencial $Ex(x|\mu) = \mu e^{-\mu x}$, $x > 0$, y sea $z_2 = \{y_1, y_2, \dots, y_m\}$ una muestra aleatoria de

los tiempos transcurridos entre dos llegadas consecutivas a la cola, obtenida a partir del modelo exponencial $Ex(x|\lambda)$.

La función de verosimilitud del número medio de entradas y de salidas del sistema será:

$$p(z_1, z_2|\lambda, \mu) = \lambda^m \mu^n \exp \{ -\mu r_1 - \lambda r_2 \} \quad (1)$$

siendo $r_1 = \sum x_i$ y $r_2 = \sum y_j$ los estadísticos suficientes relevantes. Suponemos que la información inicial sobre los parámetros λ y μ puede ser descrita como producto de dos distribuciones Gamma independientes

$$p(\lambda, \mu) = Ga(\mu|a_0, b_0) \cdot Ga(\lambda|\alpha_0, \beta_0) \quad (2)$$

dónde $Ga(x|a, b) = (b^a/\Gamma(a))x^{a-1}e^{-bx}$, $x > 0$.

Usando el método propuesto por Bernardo (1979a) para obtener distribuciones iniciales de referencia, la distribución inicial mínimo informativa de los parámetros λ, μ , que además coincide en este caso con la que se obtiene mediante la regla de Jeffreys, es:

$$\Pi(\lambda, \mu) \propto |J(\mu, \lambda)|^{1/2} = \lambda^{-1} \cdot \mu^{-1} \quad (3)$$

siendo $J(\lambda, \mu)$ la matriz de información de Fisher para el modelo (1) considerado. Es claro que la distribución inicial (3) es un caso particular de (2) cuando todos sus parámetros tienden a cero.

La distribución final de la intensidad de tráfico del sistema, $\rho = \lambda/\mu$, a partir de las densidades finales de λ y μ es:

$$p(\rho/z_1, z_2) = \frac{\beta a}{b \alpha} F(\beta a \rho / b \alpha | 2\alpha, 2a) \quad (4)$$

siendo $a = a_0 + n$, $b = b_0 + r_1$ los parámetros de la densidad final de μ , $\alpha = \alpha_0 + m$, $\beta = \beta_0 + r_2$ los correspondientes a la distribución final de λ , y representando $F(x|2u, 2v) = Cx^{u-1}/(1 + (u/v)x)^{(u+v)}$ con $C = (u/v)^u \Gamma(u+v)/\Gamma(u)\Gamma(v)$, $x > 0$, la función de densidad de una variable aleatoria con distribución F de Snedecor con $2u$ y $2v$ grados de libertad.

Finalmente, suponemos que se ha resuelto en favor de la hipótesis nula el contraste

$$\begin{aligned} H_0: \rho < 1 \\ H_1: \rho \geq 1 \end{aligned} \quad (5)$$

que garantiza el estado de equilibrio para la cola considerada.

Dicha condición supone la restricción del espacio paramétrico general al subespacio $\Omega = \{\rho, 0 < \rho < 1\}$, quedando entonces la función de dicha variable como

$$p(\rho|z_1, z_2) = \frac{p(\rho|z_1, z_2)}{P(\rho < 1|z_1, z_2)} \chi_{\Omega} \quad (6)$$

siendo $p(\rho|z_1, z_2)$ la densidad (4) y χ_{Ω} la función indicatriz de Ω .

2.2. Distribución predictiva del tiempo de duración de un período de desocupación y de ocupación del sistema

En esta subsección obtenemos en primer lugar la distribución predictiva final del tiempo de duración de un período de desocupación del sistema, para hallar posteriormente la densidad predictiva final del tiempo de duración de un período de ocupación del servicio, ambas a partir de las hipótesis establecidas en el apartado anterior.

Para un modelo de colas $M/M/1$ la distribución del tiempo de duración, t , de un período en el que el servidor de la cola permanece desocupado depende únicamente del parámetro λ que gobierna el proceso de llegadas a la cola en la forma $p(t|\lambda) = \text{Ex}(t|\lambda)$ (Kleinrock, 1976, p. 208). Consecuentemente la distribución predictiva de t , obtenida como una mixtura de $p(t|\lambda)$ y la distribución final de λ es de la forma:

$$p(t|z_2) = \alpha\beta^\alpha/(\beta + t)^{\alpha+1} \quad , \quad t > 0 \quad (7)$$

Únicamente existe momento absoluto de orden k para la variable t cuando $k < \alpha$; y para estos casos su expresión es:

$$E(t^k|z_2) = \alpha\beta^\alpha B(k + 1, \alpha - k) \quad (8)$$

(ver Gradshteyn y Ryzhik, 1965, fórmula 3.194.3), donde $B(x, y)$ es la notación usual para funciones Beta.

Es sabido (Kleinrock, 1976, p. 215) que la densidad del tiempo de duración de un período en el que el servidor permanece atendiendo a los distintos clientes que lo solicitan, condicionada por los dos parámetros λ, μ que rigen el proceso de entradas y salidas de la cola es:

$$g(t|\rho, \mu, \rho < 1) = \frac{e^{-\mu(\rho+1)t}}{t(\rho)^{1/2}} I_1\{2t\mu\rho^{1/2}\} \quad (9)$$

siendo $I_1(x)$ la función modificada de Bessel de primera clase de orden 1.

Considerando el cambio de variables, $\rho = \lambda/\mu$ y $\psi = \mu$, la distribución de μ condicionada a ρ es:

$$p(\mu|\rho) \sim \text{Ga}(\mu|a + \alpha, b + \rho\beta) \quad (10)$$

con lo cual, la densidad del tiempo de ocupación del servicio condicionada por ρ será:

$$\begin{aligned} g(t|\rho, z_1, z_2, \rho < 1) &= E_{\mu|\rho}[g(t|\rho, \mu, \rho < 1)] = \\ &= C \sum_{m=0}^{\infty} \frac{t^2 \rho \Gamma(a + \alpha + 2m + 1)}{(m + 1)! m! [\rho(t + \beta) + (t + b)]^{(a + \alpha + 2m + 1)}} \end{aligned} \quad (11)$$

con $C = (b + \rho\beta)^{-(a + \alpha)} / \Gamma(a + \alpha)$ (ver Gradshteyn y Ryzhik, 1965, fórmula 3.381.4), y por tanto, la distribución predictiva final del tiempo de duración de un período de ocupación será:

$$\begin{aligned} g(t|z_1, z_2, \rho < 1) &= E_{\rho|\rho < 1}[g(t|\rho, z_1, z_2, \rho < 1)] = \frac{C}{(t + b)^{a + \alpha + 1}} \sum_{m=0}^{\infty} \\ &\cdot \frac{t^{2m} \Gamma(a + \alpha + 2m + 1) {}_2F_1(a + \alpha + 2m + 1, a + \alpha, a + \alpha + 1; \zeta(t))}{(t + b)^{2m} (\alpha + m)(m + 1)! m!} \end{aligned} \quad (12)$$

siendo $C = b^a \beta^\alpha / P(\rho < 1|z_1, z_2) \Gamma(a) \Gamma(\alpha)$ y $\zeta(t) = -(t + \beta)/(t + b)$ (ver Gradshteyn y Ryzhik, 1965, fórmula 3.194.1), y representando ${}_2F_1(\alpha, \beta, \gamma; z)$ la notación usual para funciones hipergeométricas de Gauss.

A nivel práctico es complicado trabajar con funciones hipergeométricas, que como la involucrada en la expresión (10) no admiten cuando $b \leq \beta$, representación como serie convergente en el círculo unidad. Para la tabulación de $g(t|z_1, z_2, \rho < 1)$ se ha diseñado un algoritmo de cálculo basado en el método ara sumar series convergentes cuya razón, r_m , no siempre es decreciente con m , desarrollado por Ferrándiz (1981).

Por otra parte, los cálculos necesarios para la obtención de la media y varianza de t implican la resolución de integrales no convergentes en el intervalo $(0, 1)$, y consecuentemente, no existirá ni la esperanza ni la varianza de t .

2.3. Distribución predictiva del número de personas atendidas en un período de ocupación, y probabilidad de que un período de ocupación sea finito

Sea N_{po} la cantidad aleatoria que describe el número de personas que pueden ser atendidas durante un período de ocupación del sistema. La función de probabilidad de N_{po} condicionada a ρ con $\rho < 1$, es (Kleinrock, 1976, p. 208):

$$P(N_{po} = k | \rho, \rho < 1) = \frac{1}{k} \binom{2k-2}{k-1} \rho^{k-1} (1+\rho)^{1-2k}, \quad k \geq 1 \quad (13)$$

Utilizando la distribución final de ρ , (4), la distribución predictiva final de N_{po} , es:

$$P(N_{po} = k | z_1, z_2, \rho < 1) = E_{\rho|\rho < 1} \{ P(N_{po} = k | \rho, \rho < 1) \} = \frac{C \binom{2k-2}{k-1}}{k} I_0(\alpha + k - 2, 2k - 1, \beta/b, a + \alpha; \rho) \quad (14)$$

con

$$I_0(x, y, z, u; \rho) = \int_0^1 \frac{\rho^x}{(\rho+1)^y (1+\rho z)^u} d\rho \quad \text{con } x > 0, y > 0, z > 0, u > 0$$

y siendo $C = \beta^x \Gamma(a + \alpha) / P(\rho < 1 | z_1, z_2) \Gamma(a) \Gamma(\alpha) b^x$.

La expresión de la media de N_{p_0} incluye la resolución de una integral no convergente en $(0, 1)$, y por lo tanto, no existirá ni media ni varianza de N_{p_0} .

Por otra parte, es evidente que la probabilidad de que un período de ocupación sea finito (*POF*) o no lo sea, depende de la intensidad de tráfico de la cola. Puede verse en Kleinrock, 1976, p. 215, que dicha probabilidad es 1 si $\rho < 1$ y $1/\rho$ cuando $\rho \geq 1$.

Si se acepta como válida la hipótesis nula del contraste planteado en (5) cualquier período de ocupación tiene longitud finita con total seguridad. Si por el contrario, dicha hipótesis se rechaza en favor de la hipótesis alternativa la densidad de la intensidad de tráfico queda como

$$p(\rho|z_1, z_2) = \frac{p(\rho|z_1, z_2)}{P(\rho \geq 1|z_1, z_2)} \chi_{\Omega} \quad (15)$$

con $\Omega = \{\rho, 1 \leq \rho \leq \infty\}$, siendo entonces tal probabilidad de la forma:

$$P(POF|z_1, z_2, \rho \geq 1) = \frac{\beta^{a+1} \Gamma(a+\alpha) {}_2F_1(a+\alpha, a+1, a+2; -b/\beta)}{b^a(a+1)P(\rho \geq 1|z_1, z_2)\Gamma(a)\Gamma(\alpha)} \quad (16)$$

(ver Grashteyn y Ryzhik, 1965, fórmula 3.194.2), o bien la expresión más sencilla:

$$P(POF|z_1, z_2, \rho \geq 1) = \frac{\beta^{2a+1} P(k_1 \rho_1 > k_1) a}{b^{2a+1} (\alpha - 1) P(\rho \geq 1|z_1, z_2)} \quad (17)$$

ya que:

$$P(k_1 \rho_1 > k_1) = \frac{b^{a+1} \Gamma(a+\alpha) {}_2F_1(a+\alpha, a+1; a+2; -b/\beta)}{\beta^{a+1} (\alpha - 1) \Gamma(a+2)} \quad (18)$$

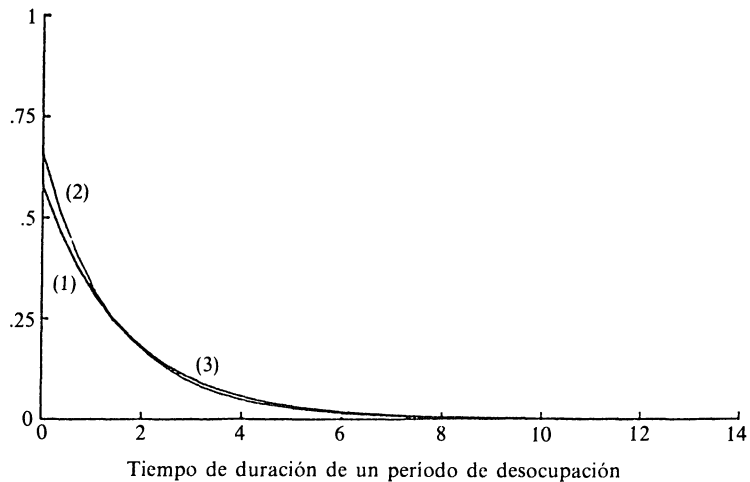
donde $k_1 \rho_1$ es una cantidad aleatoria con función de densidad F de Snedecor de parámetros $2\alpha - 2$ y $2a + 2$, con $k_1 = \beta(a + 1)/b(\alpha - 1)$.

3. EJEMPLO

Los datos que utilizamos en este apartado se han obtenido de uno de los servicios de urgencias del Hospital Clínico de Valencia. Se han tomado 25 observaciones de los tiempos de servicio y 20 de los tiem-

pos transcurridos entre dos llegadas consecutivas a la cola, siendo los estadísticos suficientes $r_1 = 32.17$ horas para los tiempos de servicio y $r_2 = 34.45$ horas para los tiempos entre llegadas. Consecuentemente, la distribución final de referencia de la intensidad de tráfico es $\rho \sim \sim 1.34F(40, 50)$, y 0.84 la probabilidad de que la cola tenga solución estacionaria.

En la gráfica siguiente se encuentran respectivamente descritas, la distribución predictiva de referencia del tiempo de duración de un período de ocupación (1), la correspondiente distribución predictiva final obtenida a partir de las opiniones iniciales del servidor de la cola ($p(\lambda, \mu) = \text{Ga}(\lambda|8,8) \text{Ga}(\mu|10, 12)$), y consecuentemente, con distribución final $p(\lambda, \mu|r_1, r_2) = \text{Ga}(\lambda|28, 42.45) \text{Ga}(\mu|25, 44.17)$, siendo entonces $\rho \sim 0.86F(56, 50)$ la densidad final de la intensidad de tráfico del servicio, y que como puede observarse indica una mayor congestión en la cola de la puede deducirse de la información muestral, (2), y finalmente la distribución de la misma variable obtenida según la metodología clásica sustituyendo λ por su estimador máximo verosímil $\hat{\lambda} = .5806$ en la distribución del tiempo de duración de un período de desocupación condicionada a λ , (3).



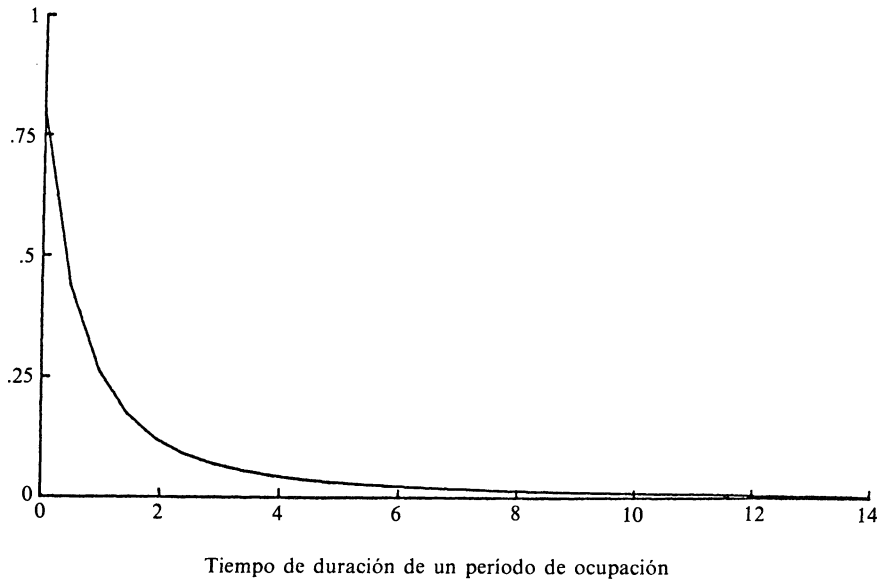
Como puede observarse, las distribuciones (1) y (3) parecen coincidir casi perfectamente. Realmente, esto no es así ya que ni su esperanza es igual (1.81 para (1) y 1.72 para (2)), ni mucho menos su varianza (3.63 para (1), 2.95 para (2)).

En la tabla que viene a continuación, y para valores del número de personas atendidas en un período de ocupación $N_{po} = 1, 2, \dots, 10, > 10$, se obtiene la distribución predictiva final de N_{po} , de referencia (1), con la información inicial considerada anteriormente (3), así como la correspondiente distribución obtenida desde la perspectiva de la Inferencia Clásica sustituyendo en la expresión (13) ρ por su estimador máximo verosímil $\hat{\rho} = 0.747$, (2).

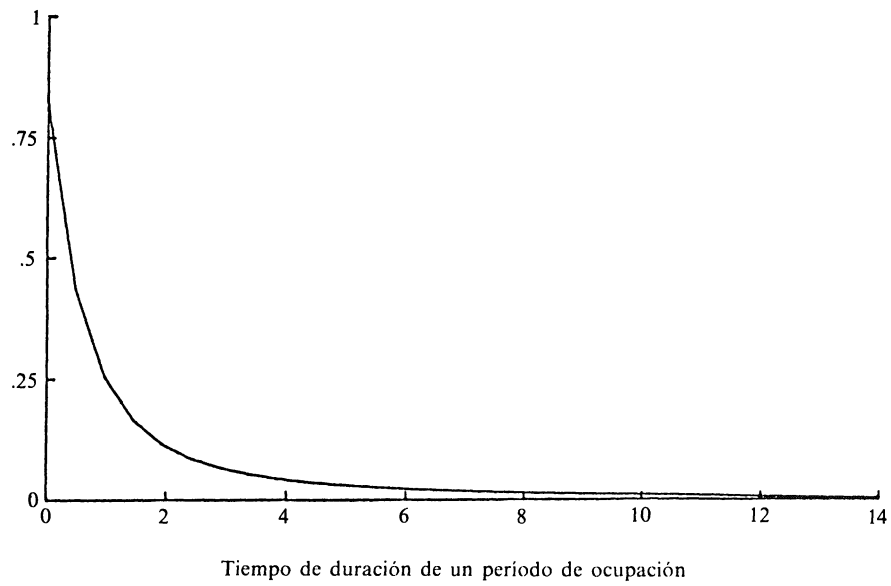
N_{po}	(1) referencia	(2) clásica	(3) inf. ini.
1	0.5938	0.5724	0.5708
2	0.1406	0.1401	0.1383
3	0.0668	0.0686	0.0671
4	0.0398	0.0420	0.0407
5	0.0266	0.0288	0.0277
6	0.0191	0.0211	0.0202
7	0.0144	0.0162	0.0155
8	0.0128	0.0129	0.0123
9	0.0091	0.0105	0.0100
10	0.0074	0.0088	0.0083
> 10	0.0696	0.0786	0.0891

Los resultados obtenidos utilizando la distribución inicial mínimo informativa de los parámetros del modelo y la obtenida mediante la Metodología Clásica son muy parecidos. Aunque la tabla expuesta sugiere que la segunda distribución es más difusa que la primera, esto no es así como podría observarse si se calculan las correspondientes probabilidades para valores superiores de la variable N_{po} , o se piensa que la estimación de los parámetros de una distribución mediante métodos clásicos produce distribuciones menos difusas que las correspondientes a las obtenidas utilizando los métodos bayesianos.

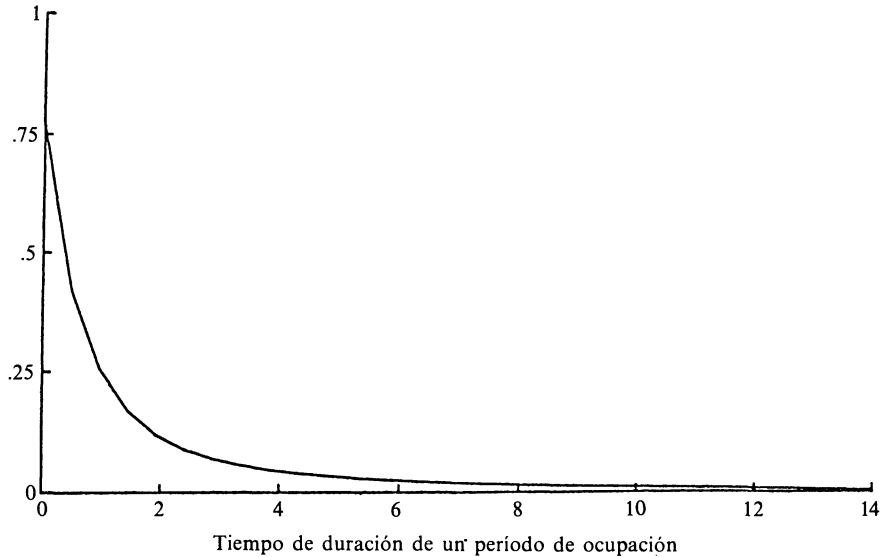
Las gráficas que siguen corresponden a, la distribución predictiva de referencia del tiempo de duración de un período de ocupación del sistema (gráfica 1), la correspondiente distribución predictiva obtenida usando la información inicial del servidor de la cola expuesta anteriormente, (gráfica 2), y la distribución obtenida desde la perspectiva Clásica, sustituyendo en la expresión (7) los parámetros λ y μ por sus estimadores máximo verosímiles.



Gráfica 1



Gráfica 2



Gráfica 3

Los comentarios hechos a propósito de los resultados obtenidos para la cantidad aleatoria N_{po} son válidos para las distintas distribuciones del tiempo de duración de un período de ocupación de la cola obtenidas: resultados y la distribución de referencia, y robustez de la distribución predictiva final de dicha cantidad aleatoria frente a la especificación de distribuciones iniciales.

4. DISCUSION

Cuando analizamos el comportamiento de una línea de espera, la selección del modelo a utilizar, así como el estudio de los distintos parámetros que gobiernan el sistema seleccionado son capítulos inevitables. Desde la Metodología Clásica, este último problema se centra en la estimación de dichos parámetros a partir de la información muestral obtenida, utilizando dichas estimaciones para así obtener un conocimiento aproximado del comportamiento de las distintas variables que intervienen en la cola.

Aunque en el ejemplo considerado hemos podido observar que la Metodología Clásica y la Bayesiana dan resultados numéricos parecidos

cuando se utilizan distribuciones iniciales de referencia, creemos que sólo desde la Metodología Bayesiana se puede dar un tratamiento adecuado a la condición $\rho < 1$ de equilibrio del sistema, y consecuentemente al hecho de que un determinado período de ocupación sea o no finito.

Finalmente, en líneas de espera para las que se disponga de poca información muestral, la Metodología Bayesiana nos permite incorporar al análisis la información inicial que tenga cualquier persona conocedora del funcionamiento de esa cola, lo cual es sumamente valioso en dichos casos.

REFERENCIAS

- BERNARDO, J. M. (1979a): «Reference posterior distributions for Bayesian inference», *J. Roy. Statist. Soc. B* 41, 113-147 (with discussion).
- FERRANDIZ, J. (1981): *Una alternativa bayesiana al contraste de hipótesis*, Tesis Doctoral, Universidad de Valencia.
- GRADSHTEYN, I. S., y RYZHIK, I. M. (1965): *Tables of integrals and products*, New York: Academic Press.
- KLEINROCK, L. (1976): *Queueing Systems. Volume I: Theory*, New York, Wiley.