

بهینه‌سازی در مدل‌های صف‌بندی^۱

دکتر محمدرضا صالحی‌راد^۲

چکیده

بهینه‌سازی در مدل‌های صف‌بندی معمولاً بر اساس اندازه‌ای مؤثر بودن، مانند میانگین تعداد در سامانه، میانگین زمان انتظار متقاضیان در سامانه، میانگین دوره‌ی اشتغال (و یا دوره‌ی بیکاری) سرویس دهنده‌ها و غیره انجام می‌گیرد. اما برای به دست آوردن این اندازه‌ها لازم است که متغیرهای موجود در مدل شناسایی شوند. سپس توزیع‌های احتمال این متغیرها را به دست آوریم. این توزیع‌ها دارای پارامترهایی هستند که معمولاً نامعلوم می‌باشند. بنابراین برای بهینه‌سازی بر اساس اندازه‌ای مؤثر بودن نیازمند برآورد این پارامترها هستیم. البته در برخی موارد توزیع‌های مورد نظر معلوم نیستند. در این حالت ابتدا باید این توزیع‌ها را تقریب کرده و پس از برآورد پارامترهای آن‌ها به تحلیل و بهینه‌سازی مدل بپردازیم.

در این تحقیق، دو مدل صف‌بندی $M/M/c$ و $M/G/1$ را در حالت پایا در نظر گرفته و با یک مثال شیوه بهینه‌سازی در مدل اول را بررسی می‌کنیم.

واژه‌های کلیدی: مدل صف $M/M/c$ ، مدل صف $M/G/1$ ، روش‌های MCMC، توزیع گامای آمیخته، اندازه‌های مؤثر بودن، حالت پایا.

طبقه‌بندی MSC: 62F15, 60K2.

۱. مبحثی تعیین‌کننده در حوزه مدل‌های صف‌بندی که در تاریخ هفدهم بهمن‌ماه سال ۹۱ در نشست‌های کرسی‌های ترویجی در محل دانشکده

اقتصاد دانشگاه علامه طباطبائی در حضور منتقدان محترم جناب آقای دکتر رضا پورطاهری و جناب آقای دکتر عبدالرحیم بادامچی‌زاده ارائه شد.

۲. عضو هیأت علمی گروه آمار، ریاضی و کامپیوتر دانشکده اقتصاد دانشگاه علامه طباطبائی – salehirad@atu.ac.ir

۱- مقدمه

نظریه‌ی صف بندی یکی از شاخه‌های احتمال و فرایندهای تصادفی است که هم از لحاظ کاربردی و هم از لحاظ نظری مورد علاقه بسیاری از محققین آمار و علوم دیگر می‌باشد. این مدل‌ها به بررسی سامانه‌هایی می‌پردازد که در آن‌ها خدماتی را به متقاضیان ارائه می‌دهند. یکی از اهداف مهم در این مدل‌ها، چگونگی سرویس دهی به متقاضیان است به طوری که این سرویس به شکل بهینه ارائه گردد.

بهینه سازی در مدل‌های صف بندی معمولاً بر اساس اندازه‌ای مؤثر بودن، انجام می‌گیرد. برای به دست آوردن این اندازه‌ها دو حالت اتفاق می‌افتد. در حالت اول، توزیع‌های زمان‌های بین ورود و زمان‌های سرویس معلوماند. در این حالت، پارامترهای نامعلوم در مدل را برآورد می‌کنیم. برآورد این پارامترها به یکی از دو شیوه‌ی کلاسیک یا بیزی انجام می‌گیرد. در حالت دوم، توزیع‌های فوق نامعلوم (کلی) هستند. در این حالت، ابتدا باید این توزیع‌ها را تقریب کرده، سپس پارامترهای مدل را برآورد کنیم.

در روش‌های کلاسیک برای مدل‌های صف بندی، معمولاً فرض می‌شود که پارامترها مقادیر ثابتی هستند اما، نامعلوم. پس از برآورد این پارامترها به بهینه سازی مدل می‌پردازیم. اما در شیوه‌ی بیزی پارامترها را متغیرهای تصادفی در نظر گرفته وبا انتخاب توزیع‌های پیشین مناسب برای آن‌ها، و همچنین نمونه‌ای که از جامعه به دست می‌آید، تابع چگالی پسین پارامترها را به دست آورده و به کمک آن برآوردهای بیزی را پیدا می‌کنیم. سپس بر اساس این برآوردها مدل را بهینه می‌کنیم.

بعضی مواقع تابع چگالی‌های پسین فرم بسته‌ای ندارند و یا غیر قابل محاسبه می‌باشند. در این صورت برای تقریب تابع چگالی‌های پسین از روش‌های شبیه سازی، مانند زنجیر مارکوف مونت کارلو (MCMC)^۱ که شامل روش‌های نمونه گیری متروپولیس^۲ و نمونه گیری گیبز^۳ و غیره است، استفاده می‌کنیم.

برای مدل‌های صف بندی که توزیع‌های زمان‌های بین ورود یا سرویس و یا هردو کلی هستند، اولین بار هیمن (۱۹۷۵)، روش تقریب انتشار را به کار برد. همچنین جین و راثو (۲۰۰۰)، با استفاده از یک فرایند براونی این توزیع‌ها را تقریب کردند. البته در این

1. Markov Chain Monte Carlo

2. Metropolis

3. Gibbs

زمینه از روش‌های ناپارامتری نیز می‌توان استفاده کرد. استفاده از روش‌های بیزی در سامانه‌های صف‌بندی به اوایل دهه ۱۹۷۰ بر می‌گردد. اولین بار باگچی^۱ و همکاران (۱۹۷۲) یک مدل صف‌بندی مارکوفی با یک سرویس دهنده را در نظر گرفته و برای آن نرخ سرویس بهینه و ظرفیت سامانه را با روش بیزی برآورد کردند. اما در دهه ۱۹۹۰ به دلیل پیچیده بودن روابط و فرمول‌های به دست آمده، استفاده از روش‌های شبیه‌سازی رواج پیدا کرد. مهم‌ترین کارهای انجام شده در این دهه، کارهای وپر^۲ و همکارانش است. ریوز^۳ و همکاران (۱۹۹۸a) یک مدل $M/G/1$ را در نظر گرفتند و تحت روش سنتی و به کمک دو مدل $M/Er/1$ و $M/H_k/1$ ، این مدل را از دیدگاه بیزی با استفاده از روش‌های شبیه‌سازی مورد بررسی قرار دادند. ریوز و همکاران (۱۹۹۸b) همچنین یک مدل $M/G/1$ را در نظر گرفتند و توزیع زمان سرویس را از طریق توزیع گامای آمیخته تقریب زده و آن را با روش سنتی مقایسه کردند. آوسین و همکاران (۲۰۰۴)، در مدل $M/G/1$ توزیع زمان سرویس را با توزیع ارلنگ آمیخته تقریب زده و پارامترهای مدل را با کمک نمونه‌گیری متروپولیس بازگشت‌پذیر برآورد کردند. آوسین و همکاران (۲۰۰۷) در ادامه تحقیقات قبلی خود، یک مدل صف‌بندی $GI/G/1$ که یک مدل کلی‌تر نسبت به مدل $M/G/1$ می‌باشد را مورد بررسی قرار دادند.

در مورد مسئله‌ی بهبودسازی در مدل‌های صف‌بندی، بررسی‌های فراوانی انجام گرفته است. در بعضی از آن‌ها بهبودسازی بر اساس نرخ‌های ورود و سرویس و در برخی دیگر بر اساس اندازه‌های مؤثر بودن در سامانه انجام شده است. برای مثال، سرفوزو (۱۹۹۹)، در یک مدل $M/M/1$ ، با تعریف یک تابع هزینه مناسب، بهبودسازی را بر اساس انتخاب نرخ‌های ورود متقاضیان و سرویس دهی انجام داده است. ما در این مقاله بهبودسازی را با تعریف یک تابع هزینه مناسب، بر اساس تعداد سرویس دهنده‌های بهینه انجام می‌دهیم.

در ادامه این مقاله، ابتدا در بخش ۲ به توصیف چگونگی برآورد در مدل صف‌بندی $M/M/1$ می‌پردازیم. سپس در بخش ۳ مدل $M/G/1$ از دیدگاه بیزی مورد بررسی قرار می‌دهیم. در این بخش، توزیع زمان سرویس برای مدل فوق را به کمک توزیع گامای

آمیخته تقریب زده و از روش‌های بیزی، پارامترهای مدل را برآورد می‌کنیم. سپس با استفاده از نتایج به دست آمده، اندازه‌های مؤثر بودن برای این مدل را محاسبه می‌کنیم. همچنین، برآورد تابع احتمال اندازه سامانه در حالت پایا را برای مدل $M/G/1$ ارایه می‌دهیم. در بخش ۴ یک مدل $M/M/c$ را در نظر گرفته و با تعریف یک تابع هزینه مناسب، تعداد سرویس دهنده‌های بهینه برای این مدل را پیدا می‌کنیم. در بخش ۵، با یک نتیجه‌گیری مقاله را به پایان می‌بریم.

۲-۲ مدل $M/M/1$

فرض کنید یک مدل صف بندی مارکوفی ساده $M/M/1$ داریم که توزیع‌های ورود و خروج از هم مستقل باشند. این مدل حالت خاصی از مدل $M/M/c$ است که در آن تعداد سرویس دهنده‌ها، c ، برابر یک است. فرض کنید نمونه‌ای به حجم n_a از زمان‌های بین دو ورود متقاضیان به صورت $\underline{t} = \{t_1, t_2, \dots, t_{n_a}\}$ داشته باشیم که در آن توزیع این زمان‌ها نمایی با پارامتر λ است. همچنین، فرض کنید نمونه‌ای به حجم n_s از زمان‌های سرویس به صورت $\underline{s} = \{s_1, s_2, \dots, s_{n_s}\}$ داشته باشیم که در آن توزیع این زمان‌ها نیز نمایی با پارامتر θ است. در اینجا فرض می‌کنیم زمان‌های بین ورود به سامانه و زمان‌های سرویس از یکدیگر مستقلند.

حال با توجه به توضیح‌های بالا به بررسی برآوردهای حداکثر درست نمایی و بیزی پارامترهای مدل $M/M/1$ می‌پردازیم.

۲-۱. برآورد حداکثر درست‌نمایی

با توجه به فرض‌های بالا، تابع درست‌نمایی برای این سامانه صف بندی به صورت زیر است

$$L(\lambda, \theta | \underline{t}, \underline{s}) = L(\lambda | \underline{t}) L(\theta | \underline{s}) \quad (1)$$

$$= \lambda^{n_a} e^{-\lambda T} \theta^{n_s} e^{-\theta S}$$

که در آن $T = \sum_{i=1}^{n_a} t_i$ و $S = \sum_{i=1}^{n_s} s_i$ هستند. حال می‌توانیم پارامترهای مدل را با استفاده از روش حداکثر درست‌نمایی برآورد کنیم. برای این منظور، از تابع درست‌نمایی داده شده در (۱)، داریم

$$\hat{\lambda}_{ML} = \frac{n_a}{T} = \frac{1}{\bar{T}}$$

$$\hat{\theta}_{ML} = \frac{n_s}{S} = \frac{1}{\bar{S}}$$

حال می‌توانیم به کمک این برآوردها اندازه‌های مؤثر بودن در سامانه را به دست آوریم. برای مثال، می‌دانیم شدت ترافیک در مدل $M/M/1$ از رابطه $\rho = \frac{\lambda}{\theta}$ محاسبه می‌شود. با قرار دادن برآوردهای حداکثر درستی پاراترهای λ و θ در ρ ، خواهیم داشت

$$\hat{\rho}_{ML} = \frac{\bar{S}}{T}$$

۲-۲. برآورد بیزی

برای برآورد بیزی پاراترهای مدل $M/M/1$ ، فرض می‌کنیم توزیع پیشین روی λ یک توزیع گاما به صورت $\pi(\lambda) \approx \Gamma(\alpha, \beta)$ باشد که در آن α و β معلومند. آنگاه توزیع پسین λ به صورت زیر است

$$\pi(\lambda | t) \approx \Gamma(n_a + \alpha, T + \beta)$$

که در آن $T = \sum_{i=1}^{n_a} t_i$ است. حال تحت تابع زیان درجه دوم خطا، برآورد بیز λ ، عبارت است از

$$\hat{\lambda}_{Bayes} = E(\lambda | t) = \frac{n_a + \alpha}{T + \beta}$$

همچنین، برای برآورد بیزی نرخ سرویس دهی، θ ، فرض می‌کنیم توزیع پیشین روی θ یک توزیع گاما به صورت $\pi(\theta) \approx \Gamma(v, \eta)$ باشد که در آن v و η معلوماند. آنگاه توزیع پسین θ به صورت زیر است

$$\pi(\theta | s) \approx \Gamma(n_s + v, S + \eta)$$

که در آن $S = \sum_{i=1}^{n_s} s_i$ است. حال تحت تابع زیان درجه دوم خطا، برآورد بیز λ ، عبارت است از

$$\hat{\theta}_{Bayes} = E(\theta | s) = \frac{n_s + v}{S + \eta}$$

همچنین، برآورد بیز ρ به صورت زیر به دست می‌آید

$$\hat{\rho}_{Bayes} = \frac{(n_a + \alpha)(S + \eta)}{(n_s + v - 1)(T + \beta)}$$

در بخش بعد مدل $M/G/1$ را بررسی می‌کنیم.

۲. مدل M/G/1

در مدل M/M/1، توزیع پسین پارامترها فرم بسته و مناسبی دارد و برآوردهای بیز پارامترها به راحتی به دست می‌آیند. اما، بعضی مواقع تابع چگالی‌های پسین فرم بسته‌ای ندارند و یا غیرقابل محاسبه می‌باشند. در این صورت، همان طور که قبلاً اشاره شد، برای تقریب تابع چگالی‌های پسین از روش‌های شبیه سازی، مانند MCMC که شامل روش‌های نمونه گیری متروپولیس^۱ و نمونه گیری گیبز^۲ و غیره است، استفاده می‌کنیم. این مسئله در مدل M/G/1 پیش می‌آید. در این مدل، چون توزیع زمان‌های سرویس کلی و نامعلوم می‌باشند لذا، ابتدا توزیع زمان سرویس را تقریب زده، سپس با انتخاب توزیع‌های پیشین مناسب برای پارامترهای مدل، اندازه‌های مؤثر بودن و تابع احتمال اندازه سامانه را برآورد می‌کنیم. برای این منظور، ابتدا در بخش (۳-۱)، روش‌های سنتی برای تقریب توزیع زمان سرویس در مدل صف بندی M/G/1 را از دیدگاه بیزی مورد بررسی قرار می‌دهیم. سپس در بخش (۳-۲)، تقریب توزیع زمان سرویس با گامای آمیخته را ارایه می‌دهیم.

۳-۱. بررسی روش سنتی برای مدل M/G/1 از دیدگاه بیزی

در روش‌های سنتی، تقریب توزیع زمان سرویس در مدل صف بندی M/G/1 بر اساس ضریب تغییرات نمونه (CV)، که از زمان‌های سرویس محاسبه می‌شود، انجام می‌گیرد. در این روش اگر مقدار ضریب تغییرات کوچک‌تر از یک باشد از توزیع ارلنگ استفاده می‌کنیم. اگر مقدار ضریب تغییرات بزرگ‌تر از یک باشد توزیع ابر نمایی^۳ را به کار می‌بریم و اگر مقدار ضریب تغییرات نزدیک یک باشد، توزیع نمایی مناسب است. مدل‌های فوق را به ترتیب با $M/Er/1$ ، $M/H_k/1$ و $M/M/1$ نشان می‌دهیم.

برای حالت سوم یعنی مدل M/M/1 که در آن توزیع ورود و خروج هر دو نمایی هستند، آرمو (۱۹۹۴) و آرمو و همکاران (۱۹۹۴)، این مدل را به طور کامل مورد بررسی قرار داده‌اند. از آنجایی که این مدل حالت خاصی از مدل M/Er/1 است بنابراین در ادامه دو مدل دیگر را مورد بررسی قرار می‌دهیم.

1. Metropolis
2. Gibbs
3. Hyper exponential

۲-۳. توزیع زمان سرویس ارلنگ

در این حالت فرض می‌کنیم زمان‌های سرویس دارای توزیع ارلنگ با پارامترهای μ و k ، به صورت زیر باشد

$$f(t|\mu, k) = \frac{(\mu k)^k}{(k-1)!} t^{k-1} \exp(-k\mu t), \quad t > 0$$

همچنین فرض می‌کنیم که توزیع پیشین توأم پارامترهای μ و k به صورت

$$f(\mu, k) \propto \frac{\theta_s^{k-1} k^{v_s k} \mu^{v_s k-1} e^{-\alpha_s k \mu}}{\Gamma(k)^{v_s}}, \quad \mu \in \mathbb{R}^+, k \in \mathbb{N}$$

باشد که در آن $v_s \in \mathbb{N}$ و $\alpha_s, \theta_s \in \mathbb{R}^+$ ثابت‌های معلومی هستند. همچنین فرض می‌کنیم μ به شرط k دارای توزیع گاما به صورت $\mu|k \sim \Gamma(v_s k, \alpha_s k)$ است. در این صورت توزیع احتمال پارامتر k به صورت

$$P(k) \propto \frac{\Gamma(\alpha_s k)}{\Gamma(k)^{v_s}} \left(\frac{\theta_s}{\alpha_s^{v_s}} \right)^{k-1}$$

خواهد بود که با انتگرال گیری از تابع چگالی تمام μ و k نسبت به μ به راحتی به دست می‌آید.

قابل ذکر است که اگر $k=1$ باشد آنگاه مدل صف بندی، یک مدل صف $M/M/1$ می‌باشد و همچنین k نمی‌تواند مقدار بی نهایت را اختیار کند، یعنی k در فاصله‌ای به صورت $1 \leq k \leq k_{\max}$ مقدار می‌گیرد. همچنین اگر $v_s=1$ و $\alpha_s=\theta_s$ باشند در این صورت یک توزیع پیشین نا آگاهی بخش^۱ برای k داریم. توجه کنید که توزیع‌های پیشین تعیین شده در بالا، توزیع‌های پیشین مزدوج هستند.

با توجه به توضیحات بالا، تابع درستنمایی زمان‌های سرویس به صورت زیر است

$$L(\mu, k|s) \propto \left(\frac{(\mu k)^k}{(k-1)!} \right)^{n_s} T_s^{k-1} \exp(-t_s k \mu)$$

که در آن n_s تعداد مشاهدات زمان سرویس، t_s مجموع T_s حاصل ضرب زمان‌های سرویس می‌باشند. حال در نتیجه زیر، توزیع پسین تمام μ و k را ارایه می‌دهیم.

نتیجه ۱: با توجه به توضیحات بالا، تابع چگالی پسین توأم μ و k برابر است با

$$f(\mu, k | s) \propto \frac{(\theta_s T_s)^{k-1} k^{(v_s + n_s)k} \mu^{(v_s + n_s)k-1} e^{-(\alpha_s + s)k\mu}}{\Gamma(k)^{v_s + n_s}}$$

در ادامه توزیع زمان سرویس را ابرنمایی در نظر گرفته و مدل حاصل را مورد بررسی قرار می‌دهیم.

۳-۳. توزیع زمان سرویس ابرنمایی

فرض می‌کنیم توزیع زمان سرویس ترکیبی از k توزیع نمایی به صورت زیر است

$$f(s|\theta) = q_1 \mu_1 e^{-\mu_1 s} + \dots + q_k \mu_k e^{-\mu_k s}$$

$$= \sum_{i=1}^k q_i \mu_i e^{-\mu_i s}$$

که در آن $\theta = (q_1, \mu_1; \dots; q_k, \mu_k)$ ، $\sum q_i = 1$ و $q_i > 0$ ، برای $i=1, \dots, k$ هستند.

حال توزیع‌های پیشین را به صورت زیر در نظر می‌گیریم

$$\mu_i \sim \Gamma(a_i, b_i); \quad i=1, \dots, k$$

و برای بردار q فرض می‌کنیم دارای توزیع دریکله به صورت زیر است

$$q \sim D(\alpha_1, \dots, \alpha_k)$$

که در آن $\alpha_i > 0$ ، برای $i=1, \dots, k$ می‌باشند.

همچنین با توجه به توزیع زمان سرویس ارایه شده و برای راحتی محاسبات برای هر s_j (که همان مشاهده j ام از زمان سرویس می‌باشد) بردار متغیرهای کمکی $z_j = (z_{1j}, \dots, z_{kj})$ را به صورت زیر تعریف می‌کنیم

$$z_j | q, \mu \sim M_k(1; q_1, \dots, q_k)$$

که در آن علامت M_k ، نشان دهنده توزیع چند جمله‌ای و $z_{ij} \in \{0, 1\}$ ، برای

$i=1, \dots, k$ و $\sum_{i=1}^k z_{ij} = 1$ است. حال با توجه به فرض‌های بالا در دو نتیجه بعد، توزیع

شرطی $s_j | z_j, q, \mu$ و همچنین توزیع‌های پسین پارامترهای μ و q را ارایه می‌دهیم.

نتیجه ۲: با توجه به فرض‌های بالا، برای متغیر s_j داریم

$$s_j | z_j, q, \mu \sim \text{Exp}(\mu_j)$$

نتیجه ۳: تحت شرایط بالا، توزیع‌های پسین به صورت زیر هستند

$$z_j | s_j, q, \mu \sim M_k \left(1; \frac{q_1 \mu_1 e^{-\mu_1 s_j}}{\sum_{i=1}^k q_i \mu_i e^{-\mu_i s_j}}, \dots, \frac{q_k \mu_k e^{-\mu_k s_j}}{\sum_{i=1}^k q_i \mu_i e^{-\mu_i s_j}} \right) \quad (\text{الف})$$

برای $j=1, \dots, n_s$.

$$\mu_j | s, z \sim \Gamma \left(a_j + \sum_{i=1}^{n_s} z_{ji}, b_j + \sum_{i=1}^{n_s} z_{ji} s_i \right) \quad (\text{ب})$$

برای $j=1, \dots, k$.

$$q | s, z \sim D \left(\alpha_1 + \sum_{i=1}^{n_s} z_{1i}, \dots, \alpha_k + \sum_{i=1}^{n_s} z_{ki} \right) \quad (\text{ج})$$

حال با توجه به تابع چگالی‌های پسین به دست آمده و با استفاده از روش‌های شبیه‌سازی زنجیر مارکوف مونت‌کارلو، یک نمونه از توزیع‌های پسین فوق، شبیه‌سازی می‌کنیم. برای این منظور، اگر $(q^{(i)}, \mu^{(i)})$ و $z^{(i)} = (z_1^{(i)}, \dots, z_{n_s}^{(i)})$ به ترتیب نشان دهنده پارامترها و متغیرهای کمکی برای i امین تکرار باشند، آنگاه مراحل انجام الگوریتم نمونه‌گیری گیبز به صورت زیر است:

(۱) برای $i=0$ با مقدار دلخواه $(q^{(0)}, \mu^{(0)}, z^{(0)})$ ، شروع می‌کنیم،

سپس مراحل زیر را تا زمان همگرایی نمونه تکرار می‌کنیم

(۲) $z_j^{(i+1)} \sim z_j^{(i)} | t_j, q^{(i)}, \mu^{(i)}$ ، برای $j=1, \dots, n_s$ ، را تولید می‌کنیم،

(۳) $q^{(i+1)} \sim q | t, z^{(i+1)}$ را تولید می‌کنیم،

(۴) $\mu_j^{(i+1)} \sim \mu_j | t, z^{(i+1)}$ ، برای $j=1, \dots, k$ ، را تولید می‌کنیم،

(۵) قرار می‌دهیم $i = i+1$ و به مرحله (۲) برمی‌گردیم.

این الگوریتم پس از چند مرحله تکرار، مثلاً m بار، به همگرایی می‌رسد. اما بر اساس روش نمونه‌گیری فوق این m مشاهده را به عنوان داده‌های سوخته در نظر گرفته و N تکرار بعدی را به عنوان نمونه اختیار می‌کنیم. نمونه حاصل که به عنوان نمونه‌ای از پارامترهای مدل می‌باشد را به صورت $\{(q^{(1)}, \mu^{(1)}), \dots, (q^{(N)}, \mu^{(N)})\}$ نشان می‌دهیم.

همچنین فرض می‌کنیم که پارامتر λ دارای توزیع پیشین گاما به صورت $\Gamma(v_a, \alpha_a)$ باشد، در این صورت، توزیع پسین λ ، نیز یک توزیع گاما به صورت $\Gamma(v_a + n_a, \alpha_a + \sum_{i=1}^n t_i)$ است. همچنین یک نمونه از توزیع پسین λ به حجم N به صورت $\{\lambda^{(1)}, \dots, \lambda^{(N)}\}$ تولید می‌کنیم. حال با استفاده از نمونه‌های تولید شده از توزیع‌های پسین پارامترهای کل سامانه، می‌توان اندازه‌های مؤثر بودن در سامانه را برآورد نمود که آن را در بخش بعد توضیح می‌دهیم.

۲-۳. تقریب توزیع زمان سرویس با گامای آمیخته

روش‌های سنتی برای تقریب توزیع زمان سرویس، تمام حالت‌های ممکن را در بر نمی‌گیرند؛ بنابراین باید روش کلی‌تری را در نظر بگیریم که علاوه بر شامل بودن سه مدل ارایه شده در بخش قبل، برازش بهتری برای توزیع زمان سرویس ارایه دهد. برای این منظور، قضیه زیر را در نظر می‌گیریم.

قضیه (دیور^۱ و لورنتز^۲ (۱۹۹۳)). برای هر تابع پیوسته $f(x)$ روی بازه $[0, \infty)$ با شرط

$$\lim_{x \rightarrow \infty} f(x) = 0$$

و تابع

$$S_u(x) = e^{-ux} \sum_{k=0}^{\infty} \frac{(ux)^k}{k!} f\left(\frac{k}{u}\right), \quad u > 0$$

داریم

$$\lim_{u \rightarrow \infty} S_u(x) = f(x)$$

این قضیه بیان می‌دارد که تابع $f(x)$ را می‌توان به وسیله ترکیبی از توابع گاما به صورت $\Gamma(k, u)$ با وزن‌های $(1/u)f(k/u)$ تقریب زد.

حال با توجه به قضیه فوق می‌توان توزیع زمان سرویس در مدل $M/G/1$ را به وسیله ترکیبی از توزیع‌های گاما تقریب کرد. برای این منظور، اگر تابع چگالی زمان سرویس را با $f(s|\theta)$ نشان دهیم، داریم

$$\begin{aligned} f(s|k, q, \mu, v) &= q_1 \frac{v_1}{\mu_1 \Gamma(v_1)} \left(\frac{v_1}{\mu_1}\right)^{v_1-1} \exp\left(-\frac{v_1}{\mu_1} s\right) + \dots + q_k \frac{v_k}{\mu_k \Gamma(v_k)} \left(\frac{v_k}{\mu_k}\right)^{v_k-1} \exp\left(-\frac{v_k}{\mu_k} s\right) \\ &= \sum_{i=1}^k q_i \text{Er}\left(s \middle| v_i, \frac{v_i}{\mu_i}\right) \end{aligned}$$

که در آن $\sum_{i=1}^k q_i = 1$ ، $q_i > 0$ و $v_i, \mu_i > 0$ هستند. توجه کنید که در این حالت k جزء پارامترها می‌باشد.

قابل توجه است که این مدل، مدل‌های سنتی ذکر شده در بخش قبل را نیز در بر می‌گیرد. به این صورت که اگر $k=1$ و $v_1=1$ باشد، توزیع نمایی را نتیجه می‌دهد. اگر $k=1$ و $v_1 \in \mathbb{N}$ باشد توزیع ارلنگ را خواهیم داشت و اگر $k>1$ و $v_i=1$ ، برای $i=1, \dots, k$ باشند توزیع ابر نمایی به دست می‌آید.

در این حالت، پارامترهای کل سامانه، λ ، پارامتر ورود و $\theta = (k, q_1, v_1, \mu_1; \dots; q_k, v_k, \mu_k)$ ، پارامترهای زمان سرویس، هستند. حال فرض می‌کنیم یک نمونه به حجم n_a از توزیع زمان‌های بین ورود به صورت $\underline{t} = \{t_1, t_2, \dots, t_{n_a}\}$ و یک نمونه به حجم n_s از زمان‌های سرویس به صورت $\underline{s} = \{s_1, s_2, \dots, s_{n_s}\}$ داشته باشیم. همچنین فرض می‌کنیم زمان‌های ورود به سامانه و زمان‌های سرویس از هم مستقل هستند. در این صورت تابع درستنمایی کل سامانه به صورت زیر است:

$$L(\lambda, \theta | \underline{t}, \underline{s}) \propto L(\lambda | \underline{t}) L(\theta | \underline{s})$$

که در آن

$$L(\lambda | \underline{t}) \propto \lambda^{n_a} \exp\left(-\lambda \sum_{i=1}^{n_a} t_i\right)$$

و

$$L(\theta | \underline{s}) = \prod_{j=1}^{n_a} \left(\sum_{i=1}^k q_i \text{Er}\left(s \left| v_i, \frac{v_i}{\mu_i} \right. \right) \right)$$

با توجه به این که توزیع‌های زمان‌های بین ورود و زمان‌های سرویس از هم مستقلاند، در نتیجه توزیع پسین پارامترهای θ و λ نیز از هم مستقل خواهند بود. برای پارامترهای زمان سرویس، θ ، به دلیل مشکل بودن محاسبات نیاز به استفاده از روش‌های شبیه سازی MCMC می‌باشد. در ادامه با ارایه توزیع‌های پیشین مناسب برای پارامترهای θ ، توزیع‌های پسین را محاسبه می‌کنیم.

۳-۲-۱. تعیین توزیع پیشین برای θ و الگوریتم MCMC برای نمونه گیری از

$$f(\theta | \underline{s})$$

ابتدا یک توزیع پیشین گسسته برای پارامتر k ، مثلاً توزیع یکنواخت از ۱ تا $k_{\max}$ را اختیار می‌کنیم. سپس توزیع‌های پیشین دیگر پارامترها را به شرط k تعریف می‌کنیم. برای راحتی محاسبات برای هر s_j ، $j=1, \dots, n_s$ ، بردار متغیرهای کمکی $\underline{z}_j = (z_{1j}, \dots, z_{kj})$ را به صورت زیر تعریف می‌کنیم

$$\underline{z}_j | q, \underline{\mu} \sim M_k(1; q_1, \dots, q_k)$$

که در این صورت توزیع زمان سرویس به شرط این متغیرها برابر است با

$$s_j | \underline{z}_j, \underline{\mu}, v \sim \text{Er}\left(s \left| v_i, \frac{v_i}{\mu_i} \right. \right)$$

حال با توجه به مطالب بالا، می‌توانیم توزیع پیشین $(q, \mu, \nu) | z, k$ را به صورت زیر بنویسیم

$$f(q, \mu, \nu | z, k) \propto f(z | q, k) f(q | k) f(\mu | k) f(\nu | k)$$

همچنین توزیع‌های پیشین پارامترهای سرویس به شرط k را به صورت زیر تعریف می‌کنیم

$$q | k \approx D(\alpha_1, \dots, \alpha_k)$$

و برای بردار پارامترهای ν ، فرض می‌کنیم که ν_j ها از هم مستقل و هر یک دارای توزیع نمایی به صورت

$$\nu_j | k \sim E(\theta) \quad \forall \quad j=1, \dots, k$$

باشند. همچنین توزیع گامای وارون را برای هر μ_j به صورت زیر در نظر می‌گیریم

$$\mu_j | k \sim \Pi(\alpha, \beta) \quad \forall \quad j=1, \dots, k$$

و فرض می‌کنیم که μ_j ها از هم مستقلاند. حال در نتیجه زیر توزیع‌های پسین پارامترها را به دست می‌آوریم.

نتیجه ۴: با توجه به توزیع‌های پیشین تعیین شده در بالا، توزیع‌های پسین به صورت زیر هستند

$$P(z_j | s, k, q, \mu, \nu) \propto q_j \frac{(\nu_j / \mu_j)^{\nu_j}}{\Gamma(\nu_j)} s_i^{\nu_j-1} \exp\left(-\frac{\nu_j}{\mu_j} s_i\right) \quad \text{الف)}$$

برای $j=1, \dots, n_s$.

$$q | s, z, k \sim D(\alpha_1 + n_1, \dots, \alpha_k + n_k) \quad \text{ب)}$$

$$\mu_j | s, z, k \sim \Pi(\alpha + n_j \nu_j, \beta + n_j \nu_j); \quad j=1, \dots, k \quad \text{ج)}$$

که در آن $n_j = \sum_{i=1}^n z_{ji}$ ، برای $j=1, \dots, k$ و $\sum_{j=1}^k n_j = n_s$ است.

$$f(\nu_j | s, \mu, k) \propto \frac{\nu_j^{n_j \nu_j}}{\Gamma(\nu_j)^{n_j}} \exp\left(-\nu_j \left(\theta + \frac{W_j}{\mu_j} + n_j \log \mu_j - \log P_j\right)\right) \quad \text{د)}$$

که در آن $P_j = \prod_{i: z_i=j} s_i$ است.

از آنجایی که مراحل اثبات‌ها طولانی می‌باشد، از آوردن آن‌ها صرف نظر می‌کنیم. برای این منظور، به محمدی و صالحی‌راد (۲۰۱۲) مراجعه نمایید.

حال با توجه به توزیع‌های پسین به دست آمده، مراحل انجام الگاریتم MCMC با نمونه‌گیری گیبز به صورت زیر است

(۱) برای $i=0$ ، با مقدار دلخواه $(q^{(0)}, \mu^{(0)}, v^{(0)})$ شروع می‌کنیم،

سپس مراحل زیر را تا زمان همگرایی نمونه تکرار می‌کنیم

(۲) $z_j^{(i+1)} \sim z_j | s, k, q^{(i)}, \mu^{(i)}, v^{(i)}$ ، برای $j=1, \dots, k$ ، را تولید می‌کنیم،

(۳) $q^{(i+1)} \sim q | s, k, z_j^{(i+1)}, \mu^{(i)}, v^{(i)}$ ، را تولید می‌کنیم،

(۴) $\mu_j^{(i+1)} \sim \mu_j | s, k, z_j^{(i+1)}, q^{(i+1)}, v^{(i)}$ ، برای $j=1, \dots, k$ ، را تولید می‌کنیم،

(۵) $v_j^{(i+1)} \sim v_j | s, k, z_j^{(i+1)}, q^{(i+1)}, \mu_j^{(i+1)}$ ، برای $j=1, \dots, k$ ، را تولید می‌کنیم،

(۶) قرار می‌دهیم $i=i+1$ و به مرحله (۲) بر می‌گردیم.

در اینجا، مانند الگوریتم بخش قبل، پس از رسیدن به همگرایی و نادیده گرفتن مشاهدات سوخته، N تکرار بعد را به عنوان نمونه اختیار کرده و پارامترهای مدل را برآورد می‌کنیم.

در ادامه، با توجه به نمونه‌های شبیه‌سازی شده، برآورد اندازه‌های مؤثر بودن را به دست می‌آوریم.

۲-۲-۳. برآورد اندازه‌های مؤثر بودن

با توجه به نمونه شبیه‌سازی شده $\{(\mu^{(1)}, v^{(1)}, q^{(1)}), \dots, (\mu^{(N)}, v^{(N)}, q^{(N)})\}$ که به کمک الگوریتم نمونه‌گیری گیبز به دست آمده است و همچنین نمونه $\{\lambda^{(1)}, \dots, \lambda^{(N)}\}$ که از پارامتر زمان بین ورود متقاضیان (λ) به دست می‌آید، می‌توانیم اندازه‌های مؤثر بودن در سامانه را برآورد کنیم. برای مثال، در اکثر مدل‌های صف‌بندی حالت پایای سامانه، از اهمیت بالایی برخوردار است. برای این منظور، می‌دانیم که سامانه زمانی در حالت پایا است که شدت ترافیک در سامانه، ρ ، کمتر از یک باشد. برای این منظور، ما می‌توانیم احتمال این که شدت ترافیک کمتر از یک باشد را به صورت زیر تقریب کنیم

$$P(\rho < 1 | t, s) \approx \frac{1}{N} \# \{ \rho^{(n)} < 1 \}$$

که در آن

$$\rho^{(n)} = \lambda^{(n)} \sum_{i=1}^k q_i^{(n)} \mu_i^{(n)} \quad \forall \quad n=1, \dots, N$$

است. حال برآورد بیزی شدت ترافیک را در نتیجه زیر ارائه می‌دهیم.

نتیجه ۵: با توجه به نمونه شبیه‌سازی شده یک برآورد بیزی برای شدت ترافیک به صورت زیر است

$$E(\rho|t, \underline{s}) \approx E(\lambda|t) \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^k q_i^{(n)} \mu_i^{(n)}$$

که در آن

$$E(\lambda|t) = (v_a + n_a) \left(\alpha_a + \sum_{i=1}^{n_a} t_i \right)^{-1}$$

می‌باشد.

در ادامه با یک نتیجه‌ی دیگر، تقریب تابع چگالی زمان سرویس را به دست می‌آوریم.
نتیجه ۶: با توجه به نمونه شبیه سازی شده تابع چگالی زمان سرویس را می‌توان به صورت زیر برآورد کرد

$$f(s|\underline{s}) \approx \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{k^{(i)}} w_j^{(i)} \Gamma \left(s \middle| v_j^{(i)}, \frac{v_j^{(i)}}{\mu_j^{(i)}} \right)$$

به همین ترتیب توزیع حاشیه‌ای پارامتر k به صورت زیر است

$$P(k|\underline{s}) \approx \frac{1}{N} \# \{n : k^{(n)} = k\}$$

نتیجه فوق روشی برای انتخاب مدل توزیع زمان سرویس می‌باشد. برای مثال، احتمال این که توزیع پسین فقط شامل یک توزیع ارلنگ باشد، برابر $P(k=1|\underline{s})$ است. اگر این احتمال به اندازه کافی بزرگ باشد، می‌توانیم توزیع زمان سرویس در مدل صف بندی را ارلنگ اختیار کنیم، یعنی مدل صف بندی به صورت $M/Et/1$ خواهد بود. سپس اندازه‌های مؤثر بودن را بر حسب این مدل محاسبه کنیم. به همین ترتیب ما می‌توانیم احتمال پسین این که توزیع زمان سرویس یک ابر نمایی باشد را به صورت زیر محاسبه کنیم

$$P(v=1|\underline{s}) \approx \frac{1}{N} \# \{n : v_i^{(n)} = 1 \quad \forall \quad i=1, \dots, k^{(n)}\}$$

اگر این احتمال به اندازه کافی بزرگ باشد، می‌توانیم توزیع زمان سرویس در مدل صف بندی را ابرنمایی اختیار کنیم، یعنی مدل صف بندی به صورت $M/H_k/1$ خواهد بود. سپس اندازه‌های مؤثر بودن را بر حسب این مدل محاسبه کنیم. به همین ترتیب احتمال پسین این که زمان سرویس دارای توزیع نمایی باشد را به صورت زیر برآورد می‌کنیم

$$P(v=1, k=1|\underline{s}) \approx \frac{1}{N} \# \{n : k^{(n)} = 1 \& v_i^{(n)} = 1 \quad \forall \quad i=1, \dots, k^{(n)}\}$$

که این همان برآورد در مدل صف بندی $M/M/1$ است.

فرض می‌کنیم سامانه صف بندی مورد بررسی در حالت پایا باشد و تحت این شرط می‌خواهیم توزیع احتمال تعداد متقاضیان در سامانه (اندازه سامانه)، را برآورد کنیم. برای این منظور، این تابع احتمال را با $\Pi(i)$ نشان می‌دهیم، یعنی احتمال این که i متقاضی در سامانه باشد. آنگاه، داریم

$$\begin{cases} \Pi(0)=1-\rho \\ \Pi(j)=\Pi(0)\delta_i+\sum_{j=1}^{i+1}\Pi(j)\delta_{i-j+1} \end{cases}$$

که در آن

$$\delta_i = \sum_{i=1}^k q_i \frac{\Gamma(v_j+i)}{i! \Gamma(v_j)} \left(\frac{\lambda}{v_j/\mu_j + \lambda} \right)^i \left(\frac{v_j/\mu_j}{v_j/\mu_j + \lambda} \right)^{v_j}$$

برای جزئیات بیشتر به محمدی و صالحی‌راد (۲۰۱۲) مراجعه کنید.

همان‌طور که در بخش قبل بیان شد، با توجه به نمونه شبیه‌سازی شده از پارامترهای سامانه با روش نمونه‌گیری گیبز، می‌توانیم مقدار $\rho^{(i)}$ را به ازای هر i به صورت

$$\rho^{(i)} = \lambda^{(i)} \sum_{j=1}^k q_j^{(i)} \mu_j^{(i)} \quad \forall \quad i=1, \dots, N$$

محاسبه کنیم. در این صورت برآورد توزیع احتمال تعداد متقاضیان در سامانه در حالت پایا به صورت

$$\Pi(j) \approx \frac{1}{R} \sum_{i: \rho^{(i)} < 1} \Pi(j | \lambda^{(i)}, q^{(i)}, \mu^{(i)}, k)$$

خواهد بود که در آن $R = \#\{\rho^{(i)} < 1\}$ ، یعنی تعداد نمونه‌هایی که در آن‌ها $\rho^{(i)} < 1$ است.

۴- بهینه‌سازی در مدل M/M/c

در این بخش، یک مدل صف بندی M/M/c را در نظر گرفته و با تعریف یک تابع هزینه مناسب، بهترین مقدار c ، یعنی تعداد سرویس دهنده‌های بهینه موازی را تعیین می‌کنیم به طوری که تابع هزینه مینیمم گردد. برای این منظور مثال زیر را در نظر می‌گیریم. فرض کنید یک شرکت خطوط هوایی دارای مرکز تلفنی با سه خط تلفن است. هر یک از این خط‌ها در طول دوره‌ی اشتغال توسط یک کارمند اداره می‌شود. در طول ۲۴ ساعت سه ساعت پراشتغال داریم که در طول این ساعت‌ها بسیاری از متقاضیان نمی‌توانند ارتباط برقرار کنند، یعنی همه خط‌ها اشغال هستند. به تجربه مشخص شده

است که حدود ۶۰ درصد افرادی که نتوانند تماس تلفنی برقرار کنند از شرکت‌های هوایی دیگر استفاده می‌کنند. فرض کنید تعداد ارتباط‌های تلفنی در طول دوره‌های پراشتغال دارای توزیع پواسون با میانگین ۲۰ تماس در ساعت باشند و هر کارمند به طور متوسط ۶ دقیقه با متقاضی صحبت کند و طول این زمان دارای توزیع نمایی باشد. الف) اگر متوسط مبلغی که هر متقاضی در هر پرواز می‌پردازد ۷۰۰ تومان باشد، در این صورت متوسط زیان روزانه به دلیل محدودیت امکانات سرویس دهی چقدر است؟ ب) اگر حقوق و مزایای شغلی هر کارمند ۸۰ تومان در ساعت باشد و ناچار باشد هر روز یک نوبت کاری، ۸ ساعت کار کند تعداد اپتیمم کارمندانی که باید به کار گرفته شوند را به دست آورید؟

فرضیات مسئله

سه ساعت کار پراشتغال، طی هشت ساعت نوبت کار روزانه اتفاق می‌افتد. در زمان‌های دیگر، یک کارمند می‌تواند ترافیک را اداره کند. شرکت هیچ گاه مرکز تلفن را تعطیل نمی‌کند و یک کارمند را در نوبت‌های کاری تعطیل به کار می‌گیرد. هزینه اضافه کردن خط‌های جدید به مرکز را ناچیز فرض می‌کنیم.

حل مسئله. با توجه به فرضیات مسئله، یک مدل صف بندی $M/M/c$ داریم. هدف در این مسئله تعیین تعداد سرویس دهنده‌های بهینه، یعنی c ، می‌باشد. برای این منظور، ابتدا هزینه‌های سامانه را محاسبه می‌کنیم که تابعی از c است. سپس مقدار c ای این هزینه‌ها را مینیمم می‌کند را به دست می‌آوریم. برای این منظور، داریم:

B_C : متوسط هزینه کل سامانه در صورت داشتن C سرویس دهنده
 P_C : احتمال زیان سامانه (احتمال زیان ارلانگ) یا احتمال پر بودن سامانه یا احتمال مشغول بودن تمام خط‌های تلفن که به صورت زیر است:

$$P_C = \frac{\frac{r^c}{c!}}{\sum_{i=0}^c \frac{r^i}{i!}} \quad (2)$$

از طرفی با توجه به اطلاعات مسئله، داریم:

$$\lambda = 20 \text{ تماس در ساعت}$$

$$\mu = \frac{60}{6} = 10 \text{ مکالمه در ساعت}$$

$$r = \frac{\lambda}{\mu} = \frac{20}{10} = 2$$

همچنین طی سه ساعت دوره شلوغی، به طور متوسط $3 \times 20 = 60$ نفر تماس می‌گیرند، بنابراین به طور متوسط $60P_c$ متقاضی طی این سه ساعت موفق به برقراری تماس نمی‌شوند. همچنین 60% این متقاضیان به شرکت‌های دیگر می‌روند، یعنی تعداد $36P_c$. در نتیجه به ازای از دست دادن این متقاضیان، مبلغ $700 \times 36P_c = 25200 P_c$ تومان به سامانه زیان وارد می‌شود (درآمد سامانه کاهش می‌یابد).

از طرفی به ازای هشت ساعت کار هر کارمند، مبلغ $8 \times 80 \times c = 640 c$ تومان هزینه دستمزد کارمندان می‌باشد. بنابراین متوسط هزینه کل سامانه در یک روز کاری برابر است با:

$$B_c = 25200 P_c + 640 c$$

$$\begin{aligned} &= 25200 \frac{\frac{r^c}{c!}}{\sum_{i=0}^c \frac{r^i}{i!}} + 640 c \\ &= \frac{25200 \times 2^c}{c! \sum_{i=0}^c \frac{2^i}{i!}} + 640 c \end{aligned}$$

حال باید این تابع هزینه را نسبت به c ، تعداد سرویس دهنده‌ها مینیمم کنیم. با توجه به شکل این تابع، یک روش برای یافتن تعداد سرویس دهنده‌های بهینه، آزمون و خطا می‌باشد که در جدول (۱-۱) آن را آورده‌ایم.

مقادیر B_c	مقادیر c
۱۷۴۴۰	۱
۱۱۳۶۰	۲
۷۲۱۲	۳
۴۹۶۱	۴
۴۰۴۰	۵
۴۱۳۴	۶
۴۴۸۱	۷

جدول (۱-۱): متوسط هزینه‌های سامانه به

ازای انتخاب C سرویس دهنده

همان طور که در جدول (۱-۱) می‌بینیم، به ازای $C=5$ سرویس دهنده، تابع هزینه داده شده در (۲)، مینیمم می‌شود. یعنی، با شرایط داده شده در مسئله، تعداد بهینه خط‌های تلفن (کارمندان) مورد نیاز این شرکت هوایی برابر پنج است.

۵- نتیجه گیری

در این مقاله، دو مدل صف بندی $M/M/c$ و $M/G/1$ را بررسی کردیم. ابتدا، مسئله برآوردیابی در مدل $M/M/1$ را ارایه دادیم که شامل روش‌های حداکثر درستیابی ویزی بود. سپس مسئله برآوردیابی در مدل $M/G/1$ که در آن توزیع سرویس کلی است را بررسی نمودیم. برای این مدل، از تقریب ارلنگ آمیخته برای برآورد توزیع زمان سرویس استفاده کردیم. اما با توجه به قضیه ارایه شده در بخش سوم می‌توانیم از تابع چگالی‌های آمیخته نرمال بریده شده در بازه $[0, \infty)$ نیز استفاده کنیم.

هدف اصلی این مقاله بررسی چگونگی بهینه سازی در مدل‌های صف بندی بوده است. در حالتی که پارامترهای مدل معلوم باشند، بهینه سازی مدل صف ساده‌تر انجام می‌گیرد. اما، در حالتی که پارامترهای مدل نامعلوم باشند، بعد از برآورد این پارامترها، به بهینه سازی مدل می‌پردازیم. در بخش پایانی، با فرض معلوم بودن پارامترهای مدل، بهینه سازی یک مدل صف $M/M/c$ را در قالب یک مثال بررسی کردیم.

لازم به ذکر است که تئوری ارایه شده در این مقاله را می‌توان به مدل‌های دیگر صف‌بندی، برای مثال مدل‌های $G/M/1$ ، $G/M/c$ و همچنین مدل‌های $G/M/c$ و $G/G/c$ تعمیم داد.

- Armero, C. (1994), "Bayesian inference in Markovian queues" *Queueing Systems* 15 (1994) 419–426.
- Armero, C., Bayarri, M.J. (1994), "Bayesian prediction in M/M/1 queues", *Queueing Systems*, 15, 401–417.
- Ausin, M.C., Wiper, M.P. (2004) "Bayesian estimation for the M/G/1 queue using a phase type approximation" *Journal of Statistical Planning and Inference*, 118, 83-101.
- Ausin, M.C., Wiper, M.P., Lillo, R.E. (2007) "Bayesian prediction of the transient behavior and busy period in short- and long-tailed GI/G/1 queueing systems" *Computational Statistics & Data Analysis*.
- Bagchi, T.P., Cunnincham, A. A. (1972), "Bayesian approach to the design of queueing systems" *INFOR.*, 10, 36-46.
- DeVore, R., Lorentz, G. (1993), "Constructive Approximation" Springer-Verlag, New York.
- Heyman, D. P. (1975), "A Diffusion Model Approximation for the GI/G/1 Queue in Heavy Traffic", *The Bell System Technical Journal*, Vol. 54, No. 9, 1637-1646.
- Jain, Sudha, Rai, T. S. S. (2000), "Statistical Inference for the GI/G/1 Queue Using Diffusion Approximation", *Information and Management Sciences*, Vol. 11, No. 2, 1-12.
- Mohammadi, A., Salehi Rad, M. R. (2012), "Bayesian Inference and Prediction in an M/G/1 queue with optional second service", *Comm. in Stat.-Simul. and Compu.*, 41, 419-435.
- Rios Insua, D., Wiper, M., Ruggeri, F. (1998), "Bayesian model selection for M/G/1 queues" *Queueing Systems*, 30, 289-308.
- Rios, D., Wiper, M.P., Ruggeri, F. (1998a), "Bayesian analysis of M/Er/1 and M/Hk/1 queues" *Queueing Systems*, 30, 289–308.
- Serfozo, Richard F. (1999), "M/M/1 Queues with Switching Costs and Hysteretic Optimal Control", *Oper. Res.*, Vol. 47, No. 2, 310-312.
- Wiper, M.P., Rios, D., Ruggeri, F. (2001), "Mixtures of gamma distributions with applications" *J. Comput. Graphical Statist.* 10, 440–454.

