Queueing Syst (2011) 68:285–293
DOI 10.1007/s11134-011-9245-y

Estimating Loynes' exponent

Ken R. Duffy · Sean P. Meyn

Received: 10 May 2011 / Revised: 10 May 2011 / Published online: 2 August 2011

Abstract Loynes' distribution, which characterizes the one dimensional marginal of the stationary solution to Lindley's recursion, possesses an ultimately exponential tail for a large class of increment processes. If one can observe increments but does not know their probabilistic properties, what are the statistical limits of estimating the tail exponent of Loynes' distribution? We conjecture that in broad generality a consistent sequence of non-parametric estimators can be constructed that satisfies a large deviation principle. We present rigorous support for this conjecture under restrictive assumptions and simulation evidence indicating why we believe it to be true in greater generality.

Keywords Single server queue · Loyne's exponent · Estimating large deviations

Mathematics Subject Classification (2000) 60K25 · 60F10

1 Introduction

If $\{X(n)\}$ is a stationary, ergodic process with $E(X(1)) < 0$, then Loynes [22] proved that there is a stationary process satisfying Lindley's recursion [21],

$$W(n+1) = \left[W(n) + X(n+1)\right]^{+} \quad \text{for all } n \in \mathbb{Z},$$

K.R. Duffy
Hamilton Institute, National University of Ireland Maynooth, Maynooth, Ireland
e-mail: ken.duffy@nuim.ie

S.P. Meyn (✉)
Department of Electrical and Computer Engineering, and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
e-mail: meyn@illinois.edu

and that all other solutions couple to it in almost surely finite time. A one dimensional marginal of the stationary solution is equal to Loynes' distribution, the distribution of the random variable $W = \sup_{n\geq 0} \sum_{i=1}^{n} X(i)$, where the empty sum $\sum_{i=1}^{0} X(i)$ is defined to be 0.

We are interested in estimating the tail behavior of Loynes' distribution. Consider the partial sums process $\{S(n)/n\}$, where $S(n) := X(1) + \cdots + X(n)$. Its associated scaled Cumulant Generating Function (sCGF) λ and associated Loynes' exponent θ^* are defined as follows,

$$\lambda(\theta) := \lim_{n\to\infty} \frac{1}{n} \log E\big[e^{\theta S(n)}\big], \qquad \theta^* := \sup\big(\theta : \lambda(\theta) \leq 0\big). \tag{1}$$

In an advance on the generality of earlier results, Glynn and Whitt [16, Theorem 1] proved the following, which justifies the terminology Loynes' *exponent*.

Theorem 1 [16] *Assume that*

(i) $\{X(n)\}$ *is strictly stationary*;
(ii) $\lambda(\theta)$ *is finite in a neighborhood of* θ^*, *differentiable at* θ^* *with* $\lambda(\theta^*) = 0$ *and* $\lambda'(\theta^*) > 0$;
(iii) $E(\exp(\theta^* S(n)) < \infty$ *for all* $n \geq 1$.
Then Loynes' distribution has an ultimately exponential tail

$$\lim_{n\to\infty} \frac{1}{n} \log P(W > n) = -\theta^*. \tag{2}$$

Note that Lelarge [20, Proposition 1] has recently shown that condition (ii) can be relaxed. The uniform bound, condition (iii), ensures that large W is caused by the cumulative behavior of a collection of increments $\{X(n)\}$ rather than a single exceedingly large increment.

Theorem 1 is of practical interest as it says that, for large n and a sizeable collection of increment processes $\{X(n)\}$, $P(W > n) \sim e^{-n\theta^*}$. If θ^* is known, this provides an estimate of the likelihood of long waiting times or large queue-lengths. This originally generated interest in estimating θ^* on-the-fly from observations to predict queueing behavior in ATM networks. For example, see early work of Courcoubetis et al. [5] and Duffield et al. [8].

2 Estimating Loynes' exponent, a conjecture

If you can observe a sequence of consecutive increments, $X(1), \ldots, X(n)$, and wish to estimate Loynes' exponent, θ^*, what are the statistical properties that your estimator can have?

Conjecture 1 *In broad generality* (*i.e.*, *conditions similar to those in Theorem* 2) *without knowing anything further about the process* $\{X(n)\}$, *one can build a sequence* *of estimates* $\{\theta^*(n)\}$ *that satisfy a Large Deviation Principle (LDP)* [6]. *That is, for all Borel sets*

$$-\inf_{x\in B^\circ} J(x) \le \liminf_{n\to\infty} \frac{1}{n}\log P\big(\theta^*(n)\in B\big) \le \limsup_{n\to\infty}\frac{1}{n}\log P\big(\theta^*(n)\in B\big) \le -\inf_{x\in\bar{B}} J(x), \quad (3)$$

where B° *denotes the interior of* B *and* $\bar{B}$ *denotes the closure of* B, *and* $J:[0,\infty]\mapsto[0,\infty]$ *is a good rate function (lower semi-continuous and with compact level sets). Moreover, one can construct consistent estimates* ($J(x)=0$ *if and only if* $x=\theta^*$).

Equation (3) can be compared with [24, Proposition 11.3.4] or [10, Theorem 3]. The following is a corollary of the latter.

Proposition 1 [10] *Suppose that the sequence* $\{X(n)\}$ *is i.i.d. with* $P(X(1)>0)>0$ *and assume that* $\lambda(\theta)=\log E[e^{\theta X(1)}]$ *is finite in a neighborhood of the origin. Then with*

$$S_w(n) := W(1)+\cdots+W(n),$$

the process $\{S_w(n)/n^2\}$ *satisfies the LDP with a good rate function* K. *That is,*

$$P\left\{\sum_{i=1}^{n} W(i) > n^2 x\right\} \sim \exp\big(-nK(x)\big)$$

with $K(0)=0$, $K(x)>0$ *for all* $x>0$, *and* $K(x)<\infty$ *for some* $x>0$.

Consequently, we are conjecturing that estimating the tail exponent of Loynes' distribution is easier than estimating its mean.

Estimation schemes A range of approaches to estimating θ^* can be deduced from the literature. Following from the seminal work [8], our focus in this article is on estimates of θ^* deduced from the MLE of the sCGF, $\lambda(\cdot)$, constructed from observations of the increments process $\{X(n)\}$. This frequentist approach is described in detail in Sect. 3 where we give theorems for it and a variant of it based on Markovian ideas. An alternative, Bayesian approach has also been proposed [14, 15, 26] where one assumes that $\{X(n)\}$ is i.i.d. or forms a Markov chain. Assuming a prior for the increment distribution or transition matrix, from observations the method identifies the most likely distribution or matrix and, hence, Loynes' exponent.

Other estimators based on observations of $\{W(n)\}$ have also been proposed, which we shall not consider here. For example, the extremal estimator $\log(n)/\max(1, W(1), \ldots, W(n))$, which converges to θ^* in a suitable sense, is studied in [4, 29] and, for the $GI/G/1$ queue, methods been developed for estimating the whole of Loynes' distribution [27] from which θ^* can be deduced (cf. [17]).

For interesting, distinct discussions on inferring input traffic and performance criteria of a queueing system from partial observations with short range dependent increments, see [3, 19]; and for possibly long range dependent Gaussian inputs, see [23].

3 Rigorous evidence in support of the conjecture

Our limited rigorous evidence for the validity of Conjecture 1 is based on two frequentist sCGF estimation schemes. We show that the conjecture holds if the non-overlapping partial sums of the increments $\{X(n)\}$ are i.i.d. and bounded for a fixed block size. This result is deduced from [9]. Moving away from independence, we also present a new result: if $\{X(n)\}$ forms a finite state irreducible Markov chain, then the conjecture holds. For the conjecture to be established in generality, new ideas are needed to extend to unbounded and non-independent increments that take more than a finite number of values, but—arguably—the unboundedness is more technically challenging. We hope this will become apparent in the exposition that follows.

The first result is based on the estimation scheme proposed in [8]. For an integer $B < \infty$, construct the blocked process

$$Y(i) := \sum_{j=(i-1)B+1}^{iB} X(j). \tag{4}$$

Select B sufficiently large that you believe the blocked process $\{Y(n)\}$ is close to being i.i.d. If the process $\{Y(n)\}$ was i.i.d., then $\lambda(\theta)$ in (1) reduces to $\lambda(\theta) = B^{-1}\log E(\exp(\theta Y(1)))$. Given observations $X(1), \dots, X(n)$, this suggests using the MLE for $\lambda(\theta)$[1] and θ^*:

$$\hat{\lambda}(n,\theta) := \frac{1}{B}\log\left(\frac{1}{\lfloor n/B\rfloor}\sum_{i=1}^{\lfloor n/B\rfloor} e^{\theta Y(i)}\right) \quad \text{and finally} \quad \theta^*(n) := \sup\bigl(\theta : \hat{\lambda}(n,\theta) \le 0\bigr).$$

A central limit theorem for $\{\theta^*(n)\}$ is proved in [8]. As a corollary to a result regarding a related estimation problem, [9, Theorem 2] proves that the sequence of estimates $\{\theta^*(n)\}$ satisfy the LDP under less restrictive conditions than those of the following theorem, but does not establish its consistency.

Theorem 2 [9] *If, for some B, $\{Y(i)\}$ is i.i.d. and $Y(i)$, with measure μ, takes values in a closed, bounded subset Σ of the real line that does not include an open ball around the origin, then Conjecture 1 holds with rate function*

$$J(x) = \inf_{\nu}\bigl\{H(\nu|\mu) : x = \sup\bigl(\theta : \lambda_\nu(\theta) \le 0\bigr)\bigr\},$$

where $\lambda_\nu(\theta) := \log \nu(\exp(\theta x))$ and $H(\nu|\mu)$ is relative entropy.

[1] This is known to be a biased estimator [13].

A sketch of the proof of Theorem 2 is as follows. By Sanov's Theorem, the empirical laws $\{L(n)\}$, defined by $L(n) := n^{-1}\sum_{i=1}^{n} 1_{Y(i)}$ for $n \geq 1$, satisfy the LDP in $\mathcal{M}_1(\Sigma)$, the space of probability measures on Σ, equipped with the topology of weak convergence. If $Y(1)$ has measure μ, then the good rate function for the LDP is the relative entropy $H(\nu|\mu)$. As Σ is bounded, for each $\theta \in \mathbb{R}$ the function $x \mapsto \exp(\theta x)$ is continuous and bounded. Thus if $\mu_n \Longrightarrow \mu$ in $\mathcal{M}_1(\Sigma)$, then $\log \mu_n(\exp(\theta x)) \to \log \mu(\exp(\theta x))$ in $\mathbb{R}$. As point-wise convergence of convex functions implies uniform convergence on bounded subsets, by the contraction principle $\{\hat{\lambda}(n,\cdot)\}$ satisfies the LDP in the space of convex functions equipped with the topology of uniform convergence on bounded subsets. To obtain the LDP for the tail exponent estimates $\{\theta^*(n)\}$, one considers the continuity of $\hat{\lambda}(n,\cdot) \mapsto \sup(\theta : \hat{\lambda}(n,\theta) \leq 0)$ and applies Puhalskii's extension of the contraction principle [28, Theorem 2.1]. As $H(\nu|\mu) = 0$ if and only if $\nu = \mu$, consistency follows from the variational form of the rate function for $\{\theta^*(n)\}$ that is given by the contraction principle.

The need to exclude $Y(n)$ taking values in an open ball around the origin is due to an artifact of the estimation scheme. If $Y(1) = \cdots = Y(n) = 0$, then $\hat{\lambda}(n,\theta) = 0$ for all θ and $\theta^*(n) = \infty$. However, in the topologically nearby situation where $Y(1) = Y(2) = \cdots = Y(n) = \epsilon > 0$, $\theta^*(n) = 0$. That is, the estimation scheme possesses a discontinuity and the assumption is imposed to avoid it.

Our second result is based on a frequentist version of the estimator used in [26]. Consider a finite state Markov chain $\{X(n)\}$ with an irreducible transition matrix $\Pi = (\pi_{i,j}) \in [0,1]^{M\times M}$, taking values $\{f(1),\ldots,f(M)\}$. For each $\theta \in \mathbb{R}$, define the matrix $\Pi_\theta = \Pi D_\theta$, where D_θ is the matrix with diagonal entries $\exp(\theta f(1)), \exp(\theta f(2)), \ldots, \exp(\theta f(M))$ and all off-diagonal entries equal to zero. The sCGF of the partial sums process $\{S(n)/n\}$ can be identified as $\lambda(\theta) = \log \rho(\Pi_\theta)$, where ρ is the spectral radius [6, Theorem 3.1.2]. With $0/0 := 0$, this suggests that one constructs the MLE for Π, $\hat{\Pi}(n)$, defined by

$$\hat{\pi}(n)_{i,j} := \left(\sum_{k=1}^{n} 1_{\{(X(k-1),X(k))=(i,j)\}}\right) \Big/ \left(\sum_{k=1}^{n} 1_{\{X(k-1)=i\}}\right),$$

and then estimates $\lambda(\theta)$ and θ^* by $\hat{\lambda}(n,\theta) = \log \rho(\hat{\Pi}(n)_\theta)$ and $\theta^*(n) = \sup(\theta : \hat{\lambda}(n,\theta) \leq 0)$.

Theorem 3 *If $\{X(n)\}$ is a finite state Markov chain with an irreducible transition matrix Π and $f(i) \neq 0$ for all $i \in \{1,\ldots,M\}$, then Conjecture 1 holds with rate function*

$$J(x) = \inf_{\nu}\bigl\{H(A|\Pi) : x = \sup\bigl(\theta : \log \rho(A_\theta) \leq 0\bigr)\bigr\},$$

where for a stochastic matrix $A = (a_{i,j})$ with stationary distribution $\psi = (\psi_i)$

$$H(A|\Pi) = \sum_{i=1}^{M} \psi_i \sum_{j=1}^{M} a_{i,j} \log \frac{a_{i,j}}{\pi_{i,j}}. \tag{5}$$

Theorem 3 follows the arguments of Theorem 2 once we establish that $\{\hat{\lambda}(n,\cdot)\}$ satisfies the LDP. By [6, Theorem 3.1.13], the empirical laws of the transitions $\{L_2(n)\}$, $L_2(n) := n^{-1}\sum_{i=1}^n 1_{(X(n-1),X(n))}$, satisfy the LDP in $\{(i,j) : \pi_{i,j} > 0\}$. With $\phi = (\phi_1,\ldots,\phi_M)$ being the stationary distribution of Π, the rate function defined by [6, (3.1.14)] is zero only at the values $\phi_i \pi_{i,j}$. If $\pi_{i,j} = 0$, then $\hat{\pi}(n)_{i,j} = 0$. For all (i,j) such that $\pi_{i,j} > 0$, we have that $\hat{\pi}(n)_{i,j}$ can be expressed as a ratio of integrals against $L_2(n)$:

$$\hat{\pi}(n)_{i,j} = \big(L_2(n)(1_{(x,y)=(i,j)})\big) / \big(L_2(n)(1_{(x,y)=(i,\cdot)})\big).$$

Thus the estimate $\hat{\Pi}(n)$ is a continuous construction from $L_2(n)$ so that contraction principle can be applied and the estimates $\{\hat{\Pi}(n)\}$ satisfy the LDP with a rate function $H(\cdot|\Pi) : [0,1]^{M\times M} \mapsto [0,\infty]$ defined in (5) that satisfies $H(A|\Pi) = 0$ if and only if $A = \Pi$. If the sequence of matrices A_n converge entry-wise to A, then by the continuous dependence of eigenvalues on entries [18], $\log \rho(A_n D_\theta)$ converges to $\log \rho(A D_\theta)$ point-wise for all θ and therefore uniformly on compact subsets of θ. Thus, from the contraction principle, $\{\hat{\lambda}(n,\cdot)\}$ satisfies the LDP and the rest of the proof follows as for Theorem 2.

Following the logic in [9, Theorem 1], as an aside we note that based on either of these two estimation schemes one can construct estimates of the rate function for the partial sums process $\{S(n)/n\}$. Defining $\hat{I}(n,x) := \sup_\theta(\theta x - \hat{\lambda}(n,\theta))$, if $\{\hat{\lambda}(n,\cdot)\}$ satisfies the LDP, then it can be shown that $\{\hat{I}(n,\cdot)\}$ satisfies the LDP in the space of $\mathbb{R} \cup \{\infty\}$ valued convex functions equipped with the Attouch–Wets topology [1, 2]. That is, there is a LDP for estimating large deviation rate functions and, under the conditions of Theorems 2 and 3, the estimates are consistent.

Example Assume that the increments process $\{X(n)\}$ forms a two-state Markov chain on the state space $\{-1,+1\}$ with transition matrix

$$\Pi = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \quad \text{where } 0 < \alpha < \beta < 1.$$

Then

$$\phi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right) \quad \text{and} \quad \theta^* = \log\left(\frac{1-\alpha}{1-\beta}\right).$$

The sequence $\{\hat{\Pi}(n)\}$ satisfies the LDP with a good rate function $H(A|\Pi)$ that is finite only at matrices of the form

$$A(a,b) = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}, \quad \text{where } a, b \in (0,1),$$

in which case (5) gives

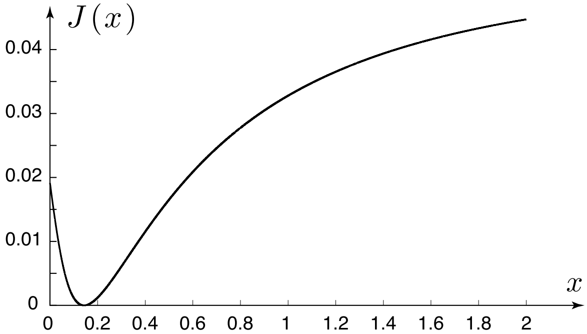

Fig. 1 Rate function for $\{\theta^*(n)\}$ based on frequentist Markov chain estimator. Markov increments $\{X(n)\}$ on $\{-1,+1\}$ with $\alpha = 1/16$, $\beta = 3/16$ and $\theta^* = \log(15/13)$

$$H\big(A(a,b)|\Pi\big) = \frac{b}{a+b}\left((1-a)\log\left(\frac{1-a}{1-\alpha}\right) + a\log\left(\frac{a}{\alpha}\right)\right) + \frac{a}{a+b}\left(b\log\left(\frac{b}{\beta}\right) + (1-b)\log\left(\frac{1-b}{1-\beta}\right)\right).$$

Consequently, the rate function for $\{\theta^*(n)\}$ is given by the one dimensional optimization

$$J(x) = \inf\left(H\big(A(a,b)|\Pi\big) : \log\left(\frac{1-a}{1-b}\right) = x\right) = \inf_{a\in(0,1)} H\big(A\big(a, 1-(1-a)e^{-x}\big)|\Pi\big).$$

While $J(x)$ cannot be determined in closed form, it can be readily calculated numerically. Figure 1 provides an example for given parameters; it is convex below Loynes' exponent, but its non-convex nature for large x is apparent.

Simulation evidence to support the conjecture For the conjecture to be substantiated, the conditions under which Theorems 2 and 3 hold need to be significantly extended to cope with more general dependence structure of the increments and, perhaps more challengingly, to remove the boundedness assumptions. A prototypical example of the later is where Lindley's recursion describes the waiting times at the $D/M/1$ queue. That is, the increments $\{X(n)\}$ are i.i.d. with $P(X(1) > x + 1/\beta) = \exp(-\alpha x)$ for $x \geq -\beta^{-1}$ and $\alpha > \beta$. This example satisfies the conditions of Theorem 1 with $\theta^*(< \alpha)$ being the positive solution of the transcendental equation

$$\log\big(\alpha/\big(\alpha - \theta^*\big)\big) - \theta^*/\beta = 0, \tag{6}$$

which can be readily solved numerically.

This example is delicate because the tail of increments decay exponentially, albeit with a larger exponent than θ^*. Even though this example breaks the hypotheses of Theorem 2, we can implement the estimation scheme in [8] and see how it performs. Fixing $B = 1$, for a single realization $X(1), \ldots, X(50{,}000)$ the left hand side of Fig. 2 plots $\theta^*(n)$ as a function of n as well as Loynes' exponent, θ^*, identified

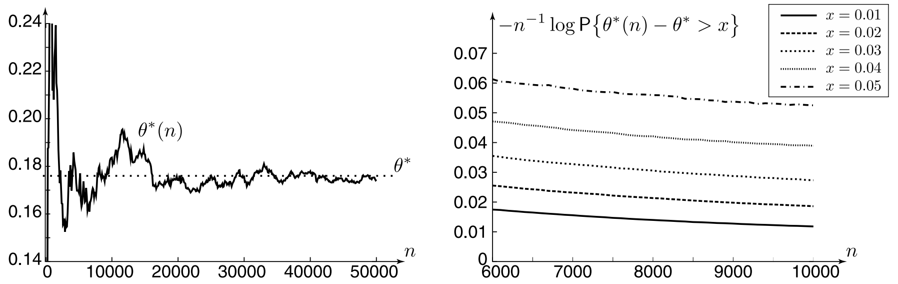

Fig. 2 $D/M/1$ queue with $\alpha = 1$ and $\beta = 10/11$, giving $\theta^* \approx 0.176$. Shown *on the left hand side* are estimates of θ^* based on (4) with $B = 1$. Plots shown *on the right hand side* are estimates of the rate function for this estimator evaluated at $\theta^* + x$ for several values of x

by solving (6). This plot indicates that the estimates are converging to the correct value. The right-hand plot is an attempt to consider the existence of an LDP. For 2×10^5 independent simulations and a range of values of x, it plots n^{-1} times the logarithm of the number of samples such that $\hat{\theta}^*(n) - \theta^* > x$, which we expect to converge to the rate function for the tail exponent estimates. This figure is suggestive of the existence of an LDP despite the departure from the boundedness conditions in Theorems 2 and 3.

Further questions Conjecture 1 is challenging, but Theorem 1 has been extended considerably so that more difficult questions can be asked. In a development of the sCGF approach in [16], [7] considers the case where the partial sums $\{S(n)/n\}$ satisfy the LDP at a non-linear speed. By reconsidering the general scaling problem in terms of rate functions rather than sCGFs, [12, Theorem 2.2] extends the results further while also correcting a lacuna (the omission of an assumption in the vein of Theorem 1(iii)).

Theorem 4 [12] *Assume that* $\{S(n)/n^A\}$ *satisfies the LDP at speed* n^V *with rate function* I *(so that, roughly speaking,* $P(S(n) > xn^A) \sim \exp(-I(x)n^V)$*) and define* $\theta^* := \inf_{x>0} x^V I(1/x^A)$. *If, in addition,* $n^{-V} \log P(S(n) > xn^A) \leq -\theta^*$ *for all* n *and all* x *sufficiently large, then*

$$\lim_{n\to\infty} \frac{1}{n^{V/A}} \log P(W > n) = -\theta^*.$$

For example, this theorem holds if $\{X(n)\}$ is i.i.d. with a Weibull distribution [25] or if $\{S(n)\}$ corresponds to sampled fractional Brownian motion or a sampled two state process with Weibull sojourn times [11].

Among our questions are: can one simultaneously estimate V/A while estimating θ^*? What impact does the non-linear scaling have on the properties of estimators?

Acknowledgements Financial support from the National Science Foundation (NSF CCF 07-29031), AFOSR (FA9550-09-1-0190) and Science Foundation Ireland (07/IN.1/1901) is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, AFOSR or SFI.

References

1. Attouch, H., Wets, R.J.-B.: A convergence theory for saddle functions. Trans. Am. Math. Soc. **280**(1), 1–41 (1983)
2. Attouch, H., Wets, R.J.-B.: Isometries for the Legendre–Fenchel transform. Trans. Am. Math. Soc. **296**(1), 33–60 (1986)
3. Baccelli, F., Kauffmann, B., Veitch, D.: Inverse problems in queueing theory and Internet probing. Queueing Syst. **63**(1–4), 59–107 (2009)
4. Berger, A.W., Whitt, W.: Maximum values in queueing processes. Probab. Eng. Inf. Sci. **9**(3), 375–409 (1995)
5. Courcoubetis, C., Kesidis, G., Ridder, A., Walrand, J., Weber, R.: Admission control and routing in ATM networks using inferences from measured buffer occupancy. IEEE Trans. Commun. **43**(2–4), 1778–1784 (1994)
6. Dembo, A., Zeitouni, O.: Large Deviation Techniques and Applications. Springer, Berlin (1998)
7. Duffield, N.G., O'Connell, N.: Large deviations and overflow probabilities for the general single server queue, with applications. Math. Proc. Camb. Philos. Soc. **118**(2), 363–374 (1995)
8. Duffield, N.G., Lewis, J.T., O'Connell, N., Russell, R., Toomey, F.: Entropy of ATM traffic streams: a tool for estimating QoS parameters. IEEE J. Sel. Areas Commun. **13**(6), 981–990 (1995)
9. Duffy, K., Metcalfe, A.P.: The large deviations of estimating rate functions. J. Appl. Probab. **42**(1), 267–274 (2005)
10. Duffy, K.R., Meyn, S.P.: Most likely paths to error when estimating the mean of a reflected random walk. Perform. Eval. **67**(12), 1290–1303 (2010)
11. Duffy, K.R., Sapozhnikov, A.: The large deviation principle for the on–off Weibull sojourn process. J. Appl. Probab. **45**(1), 107–117 (2008)
12. Duffy, K., Lewis, J.T., Sullivan, W.G.: Logarithmic asymptotics for the supremum of a stochastic process. Ann. Appl. Probab. **13**(2), 430–445 (2003)
13. Ganesh, A.J.: Bias correction in effective bandwidth estimation. Perform. Eval. **27**(8), 319–330 (1996)
14. Ganesh, A.J., O'Connell, N.: A large-deviation principle for Dirichlet posteriors. Bernoulli **6**(6), 1021–1034 (2000)
15. Ganesh, A., Green, P., O'Connell, N., Pitts, S.: Bayesian network management. Queueing Syst. Theory Appl. **28**(1–3), 267–282 (1998)
16. Glynn, P., Whitt, W.: Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. J. Appl. Probab. **31A**, 413–430 (1994)
17. Hall, P., Teugels, J.L., Vanmarcke, A.: The abscissa of convergence of the Laplace transform. J. Appl. Probab. **29**(2), 353–362 (1992)
18. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
19. Kauffmann, B., Baccelli, F., Veitch, D.: Towards multihop available bandwidth estimation. ACM SIGCOMM Comput. Commun. Rev. **37**(2), 83–84 (2009)
20. Lelarge, M.: Tail asymptotics for discrete event systems. Discrete Event Dyn. Syst. **18**(4), 563–584 (2008)
21. Lindley, D.V.: The theory of queues with a single server. Proc. Camb. Philos. Soc. **48**, 277–289 (1952)
22. Loynes, R.M.: The stability of a queue with non-independent inter-arrival and service times. Proc. Camb. Philol. Soc. **58**, 497–520 (1962)
23. Mandjes, M., van de Meent, R.: Resource dimensioning through buffer sampling. IEEE/ACM Trans. Netw. **17**(5), 1631–1644 (2009)
24. Meyn, S.P.: Control Techniques for Complex Networks. Cambridge University Press, Cambridge (2008)
25. Nagaev, S.V.: Large deviations of sums of independent random variables. Ann. Probab. **7**(5), 745–789 (1979)
26. Paschalidis, I.Ch., Vassilaras, S.: On the estimation of buffer overflow probabilities from measurements. IEEE Trans. Inf. Theory **47**(1), 178–191 (2001)
27. Pitts, S.M.: Nonparametric estimation of the stationary waiting time distribution function for the $GI/G/1$ queue. Ann. Stat. **22**(3), 1428–1446 (1994)
28. Puhalskii, A.: Large deviation analysis of the single server queue. Queueing Syst. **21**(1–2), 5–66 (1995)
29. Zeevi, A., Glynn, P.W.: Estimating tail decay for stationary sequences via extreme values. Adv. Appl. Probab. **36**(1), 198–226 (2004)