Calculus of Variations and Sample Path Large Devations for the Classical Occupancy Problem

Carl Nuzman, Phil Whiting
Bell Laboratories, Lucent Technologies
P.O. Box 636, Murray Hill,
NJ 07974, USA
Email: {cnuzman, pawhiting}@lucent.com
Tel: +1 908 582 {3365, 6912} FaX: +1 908 582 3340
Corresponding author: Phil Whiting

Abstract

The solution to a calculus of variations problem arising in connection with (sample path) large deviations asymptotics for the classical occupancy problem is presented. The approach allows tail probability estimates to be constructed for applications in computer science and communications. Some example applications are described.

1 Introduction

The classical occupancy problem centres on the distribution of r balls in n urns, each ball independently assigned to a given urn with probability $1/n$. There is a vast literature on this subject, see for example [4], [7],[6] and [1]. Examples of applications in computer science and communications include [2], [9] and [3]. [2] utilises results for the distribution of unoccupied urns only. [9] describes an application in computer science concerned with memory access and uses an urn model in which there is overflow from urns with fixed capacity. In [3] the application is to determine the requirement for wavelength converters in an optical packet switch the urn model works with the distribution of occupied urns but the number of balls thrown is randomised.

Our results are connected with the determination of the probability of rare events which is the domain of large deviations asymptotics. In our set up the scale parameter n corresponds to the number of urns and $\lceil n\beta \rceil$ balls are thrown, where the balls per urn thrown β is fixed. Let $\gamma_i(x)$ be the fraction of urns with i balls in them after x balls per urn have been thrown. We will concentrate on rare events which can be determined via $\gamma_i(\beta), i = 0, \cdots, I$, the fractions of urns with I balls or less, when all balls have been thrown. In sample path large deviations [11] the probability of a given rare event E may be estimated asymptotically by identifying the limit $-1/n \log \mathbb{P}\{E\} = \mathcal{J}$. $\mathcal{J}$ is determined as the solution to a calculus of variations problem. Solving this calculus of variations problem for the classical occupancy problem is the focus of this paper and we determine both the minimal cost and the (global) extremal curves.

Establishment of the large deviations principle itself is detailed in a companion paper [8]. This analysis relies heavily on theory developed in chapter 5 of [11] but is modified to be for a discrete

time setting and additional work is also required to deal with the boundary behaviour which arises in the classical occupancy problem.

The rest of the paper is as follows. In section 2 we provide additional background and detail our main results. Section 3 presents results in connection with the applications mentioned above. In section 4 we give an analytical outline of the solution to the calculus of variations problem. Finally there are some brief conclusions.

2 Main Results

The purpose of this section is to provide the reader with some background as to our results and then to give a formal statement of them.

In sample path large deviations the probability exponent $\mathcal{J}$ corresponding to a rare event may be computed as the solution to a calculus of variations problem. In loose terms, for large n,

$$\mathbb{P}\{Rare\ event\} \quad \approx \quad e^{-n\mathcal{J}}$$

where n, is the number of urns and the total balls thrown per urn β is fixed. In many cases the rare event is determined by the state of the urns after all the balls have been thrown. The extremal curves appearing in the solution determine the "most likely" way in which the urns fill up to produce such an event. The integral to be minimised is given via the *rate function* [11], l. In our case, l may be defined in terms of the vector fraction of urns with i balls or less in them, $\psi_i(x), i = 0, \cdots, I$ as a function of x the balls per urn currently thrown, $x \in [0, \beta]$. Define $\psi_{-1}(x) = 0$. The $\psi_i(x)$ are continuous monotone decreasing functions, with derivatives $-\dot{\psi}_i(x) \equiv \theta_i(x), i = 0, \cdots, I$

$$J_\Gamma = \int_0^\beta l(\Gamma(x), \Theta(x))dx \quad . \tag{1}$$

Here $\Gamma(x)$ is the vector fraction of urns with i balls in them so that $\gamma_i(x) = \psi_i(x) - \psi_{(i-1)}(x), i = 0, \cdots, I$ with $\gamma_{(I+1)}(x) = 1 - \psi_I(x)$. $\Theta(x)$ is the vector of derivatives $\theta_i(x)$ with $\theta_{I+1}(x) \equiv 1 - \theta_0(x) - \cdots - \theta_I(x)$. $\Theta(x)$ may be thought of as the "instantaneous" rate at which balls are entering urns with i balls in them when there are x balls per urn. For fixed x both Γ, Θ are probability vectors and

$$l(\Gamma(x), \Theta(x)) = D\left(\Theta(x) \mid\mid \Gamma(x)\right) \tag{2}$$

is the divergence or Kullback-Liebler number. This is the "instantaneous" cost of the rate Θ when the expected rate is Γ.

Optimization is over a class of continuous monotone decreasing curves, which we term *cumulative occupancy functions* which must satisfy certain consistency constraints but are otherwise arbitrary. It turns out that there is little loss of generality in focusing on the all empty initial state and so our results are for this case. It is to be emphasised that the *global* minimum solution must be determined in sample path large deviations and indeed the results which we present are *globally* minimal. It takes us too far afield to demonstrate this and so we will simply show how the curves arise as solutions to the Euler-Lagrange necessary conditions.

Before stating our main result we make some definitions.

Definition 2.1 *A sequence of $I+1$ continuous monotone decreasing functions $\psi_i : [0,\beta] \to [0,1], i = 0,\cdots,I$, satisfying*

$$0 \le \psi_0(x) \le \psi_1(x) \le \cdots \le \psi_I(x) \le 1, \tag{3}$$

and additionally,

$$\sum_i (\psi_i(x_0) - \psi_i(x_1)) \le x_1 - x_0, \quad 0 \le x_0 < x_1 \le \beta, \tag{4}$$

are said to be a sequence of cumulative occupancy functions.

A similar set of inequalities apply in the infinite case in which the inequality in (4) is replaced with equality. Cumulative occupancy functions determine the limiting sample paths of the fractions of urns with i balls in them under the large deviations scaling.

Definition 2.2 *A (continuous, monotone decreasing) function ψ_0 defined on the interval $[0,\beta]$ is said to be completely monotone provided $(-1)^i \psi_0^{(i)}(x) > 0, x \in [0,\beta]$ and $\psi_0(0) = 1$.*

In what follows we will also have $\psi_i(0) = 1, i = 0,\cdots,I$, corresponding to the empty initial state. The terminal condition will be one or more linear constraints on the fraction of urns with occupancy i, that is on $\gamma_i(\beta), i = 0,\cdots,I$. This is the case with all the applications mentioned. If all are specified we write $\gamma_i(\beta) = \xi_i, \quad i = 0,\cdots,I$ and set $\Xi = (\xi_i)_{i=0}^I$.

Our main results now follow.

Theorem 2.1 *The global minimising occupancy functions are determined via the fraction of unoccupied urns, $\psi_0(x)$ where*

$$\psi_0(x) = Ce^{-\rho x} + \sum_{i=0}^{I} a_i x^i, \quad x \in [0,\beta] \tag{5}$$

with $C, \rho > 0$ and the a_i are constants. $\psi_0(x)$ is completely monotone on the interval $[0,\beta]$ and

$$\psi_i(x) = \sum_{k=0}^{i} \frac{(-x)^k}{k!} \psi_0^{(k)}(x), i = 1,2,\cdots, \tag{6}$$

In the important special case where $\gamma_i(\beta), i = 0,\cdots,I$ are specified, C and ρ are determined via two conditions. The first condition is a normalisation constraint,

$$\sum_{i=0}^{I} \xi_i + C \sum_{i=I+1}^{\infty} e^{-\rho\beta} \frac{(\beta\rho)^i}{i!} = 1. \tag{7}$$

The second condition is on the total proportion of balls used which is

$$\sum_{i=1}^{I} i\xi_i + C \sum_{i=I+1}^{\infty} i e^{-\rho\beta} \frac{(\beta\rho)^i}{i!} = \beta \tag{8}$$

Eliminating C from these two equations yields a fixed point for ρ,

$$\frac{\beta - \sum_{i=1}^{I} i\xi_i}{1 - \sum_{i=0}^{I} \xi_i} = \frac{\sum_{i=I+1}^{\infty} i e^{-\rho\beta} \frac{(\beta\rho)^i}{i!}}{\sum_{i=I+1}^{\infty} e^{-\rho\beta} \frac{(\beta\rho)^i}{i!}} \tag{9}$$

which is the conditional mean number of balls per urn for urns which have more than I balls. In addition,

$$\psi_0(x) = \sum_{i=0}^{I} \xi_i \left(1 - \frac{x}{\beta}\right)^i - \sum_{i=0}^{I} \frac{(-\rho\beta x)^i}{\beta^i i!} C e^{-\rho\beta} \sum_{j=i}^{I} \frac{(\rho\beta)^{I-j}}{(I-j)!} + C e^{-\rho x}. \tag{10}$$

The solutions for C, ρ are unique.

The second result determines the optimal cost, $\mathcal{J}$. Let $\mathcal{Q}$ denote the probability distributions on the non-negative integers.

Theorem 2.2 $\mathcal{J} = min_{\Gamma} D\left(\Gamma \mid\mid \mathcal{P}(\beta)\right)$ (11)

where $\Gamma \in \mathcal{Q}$ satisfy the linear terminal constraints and $\mathcal{P}(\beta)$ is the Poisson distribution with mean β.

The result is related to Sanov's theorem, which however applies to independent random variables. Γ has a Poisson tail distribution with parameter $\beta\rho$ in general. ρ reflects the change in rate with which balls enter urns with more than I balls, $\rho < 1$ representing a reduced rate and $\rho > 1$ an excess rate.

The above results may be further generalised to arbitrary initial conditions and to two or more types of ball being used for example. In section 4 we sketch the proof of these results. We now illustrate these results with two applications.

3 Worked Examples

3.1 An End-Point constraint on ψ_0 only, [3],[2]

In this example we require that $\psi_0(\beta) = \xi_0$, the proportion of unoccupied cells. In this case we may obtain C either from the boundary conditions, $\psi_0(0) = -\dot{\psi}_0(0) = 1$ or from (8), so that $C = 1/\rho$ and,

$$\psi_0(x) = \frac{1}{\rho} e^{-x\rho} + 1 - \frac{1}{\rho} \quad . \tag{12}$$

is the most likely path to reach ξ_0. $\rho > 1$ yields an exponential curve lying above the expected curve e^{-x} and vice versa. Using (9) we find that ρ is determined by the nonnegative solution to,

$$\beta\rho = -\log\left(1 - \rho(1 - \xi_0)\right) \quad . \tag{13}$$

The most likely path for the other urn occupancies is

$$\begin{aligned}
\gamma_1(x) &= x e^{-x\rho} \\
\gamma_2(x) &= \rho \frac{x^2}{2} e^{-x\rho} \\
\vdots &= \vdots \\
\gamma_i(x) &= \frac{1}{\rho} e^{-x\rho} \frac{(x\rho)^j = i}{i!}
\end{aligned}$$

which is the Poisson tail with parameter modified by ρ. After an application of theorem 2.2 and C is eliminated via (7) we find,

$$\mathcal{J} = (\beta + \xi_0 - 1)\log\rho + \xi_0 \log\xi_0 - \frac{1-(1-\xi_0)\rho}{\rho}\log\left(1-(1-\xi_0)\rho\right) \tag{14}$$

3.2 The Overflow Problem [3]

This problem allows a generalisation of the results in [3] to the case where there are multiple fibres between destinations. Under the overflow problem the urns are regarded as having a finite capacity $I > 0$. Balls which are cast to a full urn are regarded as having fallen to the "floor". In [3] the case when an above average numbers of balls are lost is relevant. We proceed to recast this problem as one in which the capacity constraint is replaced by one on the fraction of urns with occupancy I. Let w be the number of overflowing balls, and m_i, m_{I+} the number of urns with $i, i \leq I$ balls, respectively more than I balls thrown at it, so that,

$$\begin{aligned} n &= m_0 + \cdots + m_I + m_{I+} \\ w &= r - m_1 - 2m_2 - \cdots - Im_I - Im_{I+} \end{aligned} \tag{15}$$

Eliminating m_I, m_{I+},

$$w = r + m_0 + (m_0 + m_1) + \cdots + (m_0 + \cdots + m_{I-1}) - In \tag{16}$$

Dividing by n and setting $\alpha \equiv w/n$, the fraction of balls per urn which overflow,

$$\alpha = \beta - I + \sum_{i=0}^{I}(I-i)\gamma_i \ . \tag{17}$$

We note that this constraint,

$$\sum_{i=0}^{I}(I-i)\gamma_i = \zeta = \alpha + I - \beta \ .$$

is of the form given in theorem 2.2. Minimising the divergence is therefore under the constraints,

$$\sum_{i=0}^{I}(I-i)\gamma_i = \zeta, \quad \sum_{i=0}^{\infty}\gamma_i = 1 \quad \sum_{i=0}^{\infty} i\gamma_i = \beta. \tag{18}$$

The method of Lagrange multipliers yields a solution,

$$\gamma_i = Ce^{-\rho\beta}\frac{(\beta\rho)^i}{i!}, \quad i \geq I \quad \gamma_i = Ce^{-\rho\beta}\frac{\beta^i\rho^i\nu^I}{\nu^i i!}, \quad i \leq I \ , \tag{19}$$

and we may obtain two fixed point equations for ρ and ν via the constraints.

$$\begin{aligned} R_I + Q_I &= \nu^{-1}R_I + \rho Q_I \\ \zeta(R_I + Q_I) &= (I - \frac{\rho\beta}{\nu})R_I + Ie^{-\rho\beta}\frac{(\beta\rho)^I}{I!} \ . \end{aligned} \tag{20}$$

where $Q_I(\rho) \equiv e^{-\rho\beta} \sum_{i=I}^{\infty} \frac{(\beta\rho)^i}{i!}$ and $R_I(\rho,\nu) \equiv e^{-\rho\beta}\nu^I \sum_{i=0}^{I-1} \frac{(\rho\beta)^i}{\nu^i i!}$. The divergence is

$$J(\zeta,\beta,I) \equiv D\left(\Gamma \mid\mid \mathcal{P}_\beta\right) = \beta + \log\zeta + \beta\log\rho + (\zeta - I)\log\nu - \log(S_I) \tag{21}$$

on eliminating C, where $S_I(\rho,\nu) \equiv \sum_{i=0}^{I}(I-i).\frac{(\rho\beta)^i}{\nu^i i!}$. For the overflow problem we have,

$$\mathcal{J} \quad = \quad J(\alpha + I - \beta, \beta, I), \quad [\beta - I]_+ \leq \alpha < \beta \tag{22}$$

ρ, ν are obtained from (21) with $\alpha + I - \beta$ substituted for ζ.

4 Analysis Overview

We begin by outlining the main steps in demonstrating theorem 2.1. After setting up the finite Euler-Lagrange equations these steps are:-

i) Write the formal equivalent for an infinite system and cast into canonical form
ii) Show that the canonical equation may be solved in terms of the derivatives of ψ_0
iii) Obtain the $\psi_i, i = 1, 2, \cdots$ by inspection
iv) Obtain the form of ψ_0 by inspection
v) Demonstrate that for terminal conditions specifying a given consistent end point Ξ, such a solution exists.
vi) Derive the finite solution from the infinite case
vii) Verify that the solutions define a sequence of cumulative occupancy functions.

Some comments are in order. An end point Ξ is consistent provided it can be realised using β balls per urn i.e.

$$\beta \geq \sum_{i=0}^{I} i\xi_i + (I+1)\left(1 - \sum_{i=0}^{I}\xi_i\right) \tag{23}$$

The solutions thus define a family of extremals which attain every possible endpoint. That the $\psi_i, i = 0, \cdots, I$ determine a sequence of cumulative occupancy functions is a consequence of ψ_0 being completely monotone as we will see. This latter holds for choices of the a_i so as to pass through a consistent endpoint.

As far as theorem 2.2 is concerned, integration over the extremal establishes the cost to be of the form $D\left(\Gamma \mid\mid \mathcal{P}(\beta)\right)$. That this cost should be minimal over $\Gamma \in \mathcal{Q}$ to a given end point Ξ may be shown via a Lagrangean argument in which the C, ρ are used to construct Lagrange multipliers. The minimising property establishes the uniqueness of C, ρ for a given end point Ξ. These results hold more generally since the minimum will be determined by some end point satisfying the terminal conditions.

We begin our outline by stating the Euler-Lagrange equations (necessary conditions) for an extremal family of curves $\psi_i, i = 0, \ldots, I$ see [10],

$$\begin{aligned} \frac{\partial l}{\partial \psi_i} &= \frac{d}{dx}\left\{\frac{\partial l}{\partial \dot{\psi}_i}\right\}, \\ -\frac{\theta_i}{(\psi_i - \psi_{i-1})} + \frac{\theta_{i+1}}{(\psi_{i+1} - \psi_i)} &= \frac{d}{dx}\left\{-\log\frac{\theta_i}{(\psi_i - \psi_{i-1})} + \log\frac{\theta_{I+1}}{(1 - \psi_I)}\right\} \end{aligned} \tag{24}$$

where we take $\psi_{I+1} = 1$. The ψ_i must satisfy the initial (empty) and terminal conditions.

The formal equivalent of the Euler-Lagrange equations for an infinite sequence of functions ψ_i, correspond to the integral functional,

$$\mathcal{J}_\infty(\beta) = \int_0^\beta \tilde{l}(\Psi;\Theta)\,dx \tag{25}$$

with $\tilde{l}(\Psi,\Theta) \equiv \sum_{j=0}^{\infty} \theta_j \log \frac{\theta_j}{(\psi_j - \psi_{j-1})}$ but constraints on the values of $\psi_i(\beta), i = 0, 1, \cdots, I$ only. Formal use of variations leads to the infinite sequence of Euler-Lagrange equations

$$-\frac{\theta_i}{(\psi_i - \psi_{i-1})} + \frac{\theta_{i+1}}{(\psi_{i+1} - \psi_i)} \quad = \quad \frac{d}{dx}\left\{ -\log \frac{\theta_i}{(\psi_i - \psi_{i-1})} \right\} \quad .$$

These can be reformulated as a double sequence of first order ordinary differential equations, [10], by introducing the auxiliary variables

$$p_i = -\log \frac{-\dot{\psi}_i}{(\psi_i - \psi_{i-1})}, \quad i = 0, 1, 2, 3, \cdots \tag{26}$$

so that the Euler-Lagrange equations become,

$$\dot{p_i} = e^{-p_{i+1}} - e^{-p_i}, \quad i = 0, 1, 2, 3, \cdots . \tag{27}$$

The following determines the $p_i, i = 0, 1, \cdots$ in terms of ψ_0 and its derivatives.

Lemma 4.1 *Suppose that $\psi_0, \psi_1, \cdots$ are a sequence of functions which satisfy the equations (26) and (27) on the interval $[0, \beta]$, Suppose also that $\psi_0(x)$ is completely monotone on $[0, \beta]$. Then the infinite sequence of functions $\{p_i\}$ satisfy*

$$p_i = -\log\left(-\frac{\psi_0^{(i+1)}}{\psi_0^{(i)}} \right) \tag{28}$$

and are a solution to (27).

The results allow the $\psi_i, i = 1, 2, \cdots$ to be expressed in terms of the derivatives of ψ_0.

Lemma 4.2 *Suppose that ψ_0 is infinitely differentiable and completely monotone on $[0, \beta]$. Then the infinite sequence of functions $\{\psi_i\}$ defined by*

$$\psi_i(x) = \sum_{j=0}^{i} \frac{(-x)^j}{j!} \psi_0^{(j)}(x) \tag{29}$$

solve the Euler-Lagrange equations (26) and (27) as well as the initial conditions.

The lemma may be verified by differentiation and substitution into (26).

If the function $\psi_0(x)$ is *completely monotone* over $[0, \infty)$ with $\psi_0(0) = 1$ then ψ_0 must be the Laplace transform of a probability distribution, see [5]. Since we essentially "don't care" about what happens to urns with higher occupancy than I we expect that this will be reflected only as a change of rate of hitting such urns. This suggests that we require

$$\frac{-\dot{\psi}_i(x)}{\psi_i(x) - \psi_{(i-1)}(x)} = \rho > 0, \quad i > I. \tag{30}$$

and, moreover, that we should take

$$\psi_0(x) = Ce^{-\rho x} + \sum_{i=0}^{I} a_i x^i. \tag{31}$$

(5) ensures that (30) is satisfied. The a_i can be chosen so as to meet the terminal and initial conditions. Indeed (10) may be verified by using $\gamma_i(\beta) = \frac{(-\beta)^i}{i!}\psi_0^{(i)}(\beta)$. Finally a continuity argument shows that there is at least one solution for ρ via (9), C is then determined via (7). This completes the outline for steps (i) - (v). We now pause to evaluate the cost of a path.

Define $\mathcal{J} \equiv \int_0^\beta \sum_{k=0}^\infty -\dot{\psi}_k(x) \log \frac{-\dot{\psi}_k(x)}{\dot{\psi}_k(x)-\dot{\psi}_{(k-1)}(x)} dx$. Using (26) together with (28) and after some rearrangement we find,

$$\mathcal{J} \;=\; \sum_{k=0}^{\infty} \int_0^\beta (\dot{\psi}_k(x) - \dot{\psi}_{(k-1)}(x)) \log \left|\psi_0^{(k)}(x)\right| dx \quad \text{since } \psi_{-1} = 0 \ .$$

Integrating by parts and after further algebra we find,

$$\mathcal{J} \;=\; \sum_{k=0}^{\infty} \left[\gamma_i \log |\psi_0^{(i)}|\right]^\beta + \beta.$$

Where we have used $\gamma_0(0) = 1$ and $\gamma_i(0) = 0$ otherwise. Using $\gamma_i(x) = (-x)^i/i! \cdot \psi_0^{(i)}(x)$, we find,

$$\mathcal{J} = D\left(\Gamma \;||\; \mathcal{P}(\beta)\right), \tag{32}$$

where Γ denotes the terminal distribution.

The curves $\psi_0, \cdots, \psi_I$ from the infinite system, solve the finite system of Euler-Lagrange equations and the path cost coincides with the infinite case. To see this note $1-\theta_0-\cdots-\theta_I = \dot{\psi}_{I+1}+\dot{\psi}_{I+2}+\cdots$ from which it follows that

$$\begin{aligned} \frac{1-\theta_0-\cdots-\theta_I}{1-\psi_I} &= \frac{-\left(\dot{\psi}_{I+1}+\dot{\psi}_{I+2}+\cdots\right)}{(\psi_{I+1}-\psi_I)+(\psi_{I+2}-\psi_{I+1})+\cdots} \\ &= \rho \end{aligned} \tag{33}$$

by choice of ψ_0. It therefore follows that the finite system of Euler-Lagrange equations collapse to the first I equations of the infinite system. That is

$$\frac{\theta_{i+1}}{\psi_{i+1}-\psi_i} - \frac{\theta_i}{\psi_i-\psi_{i-1}} = \frac{d}{dx}\left\{\log\frac{1-\theta_0-\cdots-\theta_I}{1-\psi_I} - \log\frac{\theta_i}{\psi_i-\psi_{i-1}}\right\}, i = 0, \cdots, I-1 \ , \tag{34}$$

because the first term on the right is a constant and so disappears. Similar considerations apply to the remaining Ith equation. As far as the cost is concerned the terms beyond I sum to

$$(1-\theta_0-\cdots-\theta_I)\log\frac{(1-\theta_0-\cdots-\theta_I)}{(1-\psi_I)} = \sum_{i=I+1}^{\infty} -\dot{\psi}_i \log\frac{-\dot{\psi}_i}{(\psi_i-\psi_{i-1})} \quad . \tag{35}$$

This is step (vi).

For step (vii) it is readily shown that ψ_0 as given in (10) is completely monotone so that (3) follows. Further

$$\sum_{i=0}^{I} \theta_i(x) = -\sum_{i=0}^{I} \dot{\psi}_i(x) = -\sum_{i=0}^{I} \frac{(-x)^i}{i!}\psi_0^{(i+1)}(x) \leq -\sum_{i=0}^{\infty} \frac{(-x)^i}{i!}\psi_0^{(i+1)}(x) = \psi_0^{(1)}(0) = 1 \qquad (36)$$

so that (4) is satisfied. Also $\psi_i(0) = 1, i = 0, 1, \cdots$ by construction.

For theorem 2.2, consider the minimisation problem

$$\min_{\Gamma} D\left(\Gamma \mid\mid \mathcal{P}(\beta)\right)$$

over probability distributions Γ such that $\sum_{i=0}^{\infty} i\gamma_i = \beta$ and $\gamma_i = \xi_i > 0,\ i = 0, \cdots, I$. We form the infinite Lagrangean,

$$\mathcal{L} = \sum_{i=I+1}^{\infty} \gamma_i \log \frac{\gamma_i}{p_i(\beta)} + y\left(1 - \sum_{j=0} \xi_j - \sum_{k=I+1}^{\infty} \gamma_i\right) + \lambda\left(\beta - \sum_{j=0}^{I} j\xi_j - \sum_{i=I+1}^{\infty} i\gamma_i\right) \qquad (37)$$

defined on $\xi_i \geq 0, i = I+1, I+2, \cdots$. Here $p_i(\beta) = e^{-\beta}(\beta^i)/i!$ are the Poisson probabilities with mean β. Set $y, \lambda > 0$ to be $e^{(y-1)-\beta} = Ce^{-\beta\rho}$ and $e^{\lambda} = \rho$. Since the Lagrangean is a sum of functions over γ_i, we may minimise term by term to obtain,

$$\begin{aligned} \log \gamma_i + 1 - y - i\lambda - \log p_i(\beta) &= 0 \qquad (38) \\ \gamma_i &= Ce^{-\rho\beta}\frac{(\beta\rho)^i}{i!} \end{aligned}$$

which determines the minimum since each term is convex in γ_i. However this solution satisfies the equality constraints and hence by the Lagrange sufficiency theorem [12] we have minimised the objective. We have thus shown,

Lemma 4.3 *The tail distribution $\gamma_i = Ce^{-\rho\beta} \cdot (\beta\rho)^i/i!$ where C, ρ are obtained from (8),(9) determine a solution to the problem,*

$$\min_{\Gamma} D\left(\Gamma \mid\mid \mathcal{P}(\beta)\right)$$

over probability distributions Γ such that $\sum_{i=0}^{\infty} i\gamma_i = \beta$ and $\gamma_i = \xi_i > 0,\ 0, \cdots, I$.

Corollary 4.1 *C, ρ are uniquely determined by (8),(9) and the solution to the minimisation problem is unique.*

Proof
$D\left(\Gamma \mid\mid \mathcal{P}(\beta)\right)$ is a strictly convex function of $\Gamma \in \mathcal{Q}$ there cannot be two distinct probability distributions achieving the minimum. □

The following is also immediate,

Corollary 4.2 *The cost function corresponding to the extremal determined by $\psi_0(x)$ as given in (10) equals $\min_{\Gamma} D\left(\Gamma \mid\mid \mathcal{P}(\beta)\right)$ where the minimum is taken over distributions satisfying the constraints.*

5 Conclusions

The calculation of tail probaility (large deviations) estimates in connection with the classical occupancy problem arise in applications connected with computer memory and communication network modelling. Use of exact results to analyse these problems is often hampered by numerical intractability. The results here provide exact formulas for the large deviations exponent and involve an implicit equation which is much more readily solved. Moreover the extremals are relevant to importance sampling of the classical occupancy problem. The solution may be generalised to other initial conditions. Models with distinct probabilities for entering urns as well as those with balls of different colours may also be analysed.

References

[1] A. Charalambides. Occupancy and Sequential Occupancy Distributions. In *Advances in Combinatorial Methods and Applications to Probability and Statistics, N. Balakrishnan Ed.*, pages 259–273, 1997.

[2] A. Kamath et al. Tail bounds for occupancy and the satisfiability threshold conjecture. *Random Structures and Algorithms*, 7(1):59 – 80, 1995.

[3] V. Eramo M. Listanti et al. Optical switch dimensioning and the classical occupancy problem. Submitted to 39th Allerton Conference, 2001.

[4] W. Feller. *An Introduction to Probability Theory and Its Applications Vol. 1.* John Wiley, New York, 1968.

[5] W. Feller. *An Introduction to Probability Theory and Its Applications Vol. 2.* John Wiley, New York, 1971.

[6] Lars Holst. On the Coupon Collectors and other Urn Problems. *International Statistical Review*, 54(1):15–27, 1986.

[7] N.L.Johnson and S. Kotz. *Urn Models and their Applications.* John Wiley, 1977.

[8] Carl Nuzman and Phil Whiting. Sample path large deviations for the classical occupancy problem. Technical Memorandum Lucent Technologies (in preparation).

[9] M.V. Ramakrishna and P. Mukhopadhyay. Analysis of Bounded Disorder File Organisation. In *Proc. of the 7th ACM Symposium of the Principles of DataBase Systems, Austin, Texas*, pages 117–125, 1988.

[10] Hans Sagan. *Introduction to the Calculus of Variations.* Dover, 1969.

[11] Adam Shwartz and Alan Weiss. *Large Deviations for Performance Analysis.* Chapman and Hall, 1995.

[12] Peter Whittle. *Optimization under Constraints.* John Wiley, 1971.