STATIONARY ANALYSIS OF AN \((R, Q)\) INVENTORY MODEL WITH NORMAL AND EMERGENCY ORDERS

ONNO BOXMA, *Eindhoven University of Technology
DAVID PERRY, **Holon Institute of Technology
WOLFGANG STADJE, ***University of Osnabrück

Abstract

We consider an \((R, Q)\) inventory model with two types of orders, normal orders and emergency orders, which are issued at different inventory levels. These orders are delivered after exponentially distributed lead times. In between deliveries, the inventory level decreases in a state-dependent way, according to a release rate function \(\alpha()\). This function represents the fluid demand rate; it could be controlled by a system manager via price adaptations. We determine the mean number of downcrossings \(\theta(x)\) of any level \(x\) in one regenerative cycle, and use it to obtain the steady-state density \(f(x)\) of the inventory level. We also derive the rates of occurrence of normal deliveries and of emergency deliveries, and the steady-state probability of having zero inventory.

Keywords: \((R, Q)\) inventory; stochastic lead time; steady-state analysis; level crossings

2020 Mathematics Subject Classification: Primary 90B05
Secondary 60K30

1. Introduction

\((R, Q)\) policies, also known as reorder-point/order-quantity policies, belong to the most common policies in inventory control systems. In \((R, Q)\)-type models considered in this paper the inventory level is gradually decreasing, due to a deterministic fluid demand. When the inventory level drops below a certain control level \(R\), an order of size \(Q\) is placed. That order arrives after a certain lead time, and at this time epoch the inventory increases by \(Q\). In contrast to classical EOQ models, the lead times are not negligible but stochastic variables.

The problem faced by the controller is when to place an order and how much to order. The decision variable \(R\) answers the when question and the decision variable \(Q\) answers the how much question. In the broad literature of economics and finance, so-called trigger–target models are frequently considered [3, 4, 5, 6, 24]; in these models the lead times are negligible, there is a trigger value, and when hit the inventory level is reset to a fixed target value. When the lead times are random, there cannot be a fixed target value and a heuristic \((R, Q)\) policy seems the most natural policy (and accepted in the literature).

Received 30 September 2020; revision received 24 March 2022.
* Postal address: Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email address: o.j.boxma@tue.nl
** Postal address: Holon Institute of Technology, PO Box 305, Holon 5810201, Israel. Email address: dperry@stat.haifa.ac.il
*** Postal address: Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany. Email address: wstadje@uos.de

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.
The main contribution of our study is the analysis of an inventory model, called \((\tilde{R}, \tilde{Q})\), with two types of orders: normal orders and emergency orders. Here \(R = (r_n, r_e)\) and \(Q = (q_n, q_e)\). There are two trigger levels: \(r_n\) which triggers a normal order with normal lead time \(\sim \exp(\lambda_n)\), and \(r_e\) which triggers an emergency order with emergency lead time \(\sim \exp(\lambda_e)\). We also have two target sizes \((q_n, q_e)\). We assume that \(r_e < r_n\), since an emergency order is placed when the normal order is still pending while the inventory level is close to 0 (we use the terminology that during its lead time an order is pending). We assume, for the sake of simplicity, that \(q_e = q_n\), although often in practice \(q_e > q_n\) because the set-up cost of the emergency order might be higher than that of the normal order. Our analysis goes through for \(q_e > q_n\) (or \(q_n > q_e\)) as well, but at the expense of an (even) more intricate distinguishing of cases. These complications will become clear in the proof of Theorem 1. Below we simplify the notation and set \(r_n = a\), \(r_e = b > a\) and \(q_n = q_e = q\). Accordingly, the inventory considered here is characterized by the triple \((a, b, q)\) with \(q > b > a\). This latter condition ensures that any delivery takes the inventory level back to a size deemed satisfactory without immediate further orders.

A key feature of our study is that the inventory level process \(V = \{V(t) : t \geq 0\}\) decreases continuously according to a quite general state-dependent release rate function \(\alpha(x)\), when the inventory level is \(x\), between positive replenishments of size \(q\) which are delivered at the end of the respective lead times. The motivation for introducing a state-dependent release rate \(\alpha(\cdot)\) stems from pricing. When the content level is too high the controller will be interested in getting rid of excess inventory as quickly as possible, to reduce the holding costs. To achieve this goal she lowers the price continuously in order to raise the demand rate (it seems natural that price and demand rate are related deterministically). Similarly, when the inventory level is low she will raise the price and thereby reduce the demand rate. This way the risk of losses due to shortage and unsatisfied demand is also reduced.

The main result of the paper is an exact (non-transform) closed-form expression for the stationary density of the inventory level, which turns out to be highly complicated. The approach to derive this density also appears to be of methodological interest. We circumvent the problems that the inventory level process is not Markovian, and that deliveries do not arrive according to a Poisson process, by relating the stationary inventory level density \(f(x)\) to the mean number of downcrossings \(\theta(x)\) of level \(x\) in one regenerative cycle, and by first determining \(\theta(x)\), which subsequently also gives us \(f(x)\). We also derive the rates of occurrence of normal deliveries and of emergency deliveries, and the steady-state probability of having zero inventory.

Different supply modes are used in many real-world inventory systems; having normal and emergency orders is one example of this feature. For example, shipping of orders can be carried out by sea or by land but also by air, probably at a higher cost, or there may be more expensive sources of replenishments which are only used in case of very low inventory levels. Complex emergency situations of this latter kind have recently occurred in the worldwide hunt for medical equipment in the face of the coronavirus pandemic. Our model is a modest theoretical contribution to this problem area.

**Literature review.** Continuous-review inventory models with multiple sourcing (multiple supply modes; both normal and emergency shipping) are studied extensively in the operations research literature, almost exclusively with the focus on optimization. Models with two supply modes have been frequently suggested (see e.g. [9], [17], [20], [25], [26], [31], [32], and [33]). In contrast to this paper, they assume stochastic demand, usually of jump type, and two deterministic lead times for the different order types; their objective is multi-period optimization. In general, these models are characterized by objective functions whose components are the
An (R, Q) inventory model

3
cost (or profit) functionals and include measures of uncertainty, but the cost functionals are
not too difficult to obtain. In our model the computation of the cost measures is challenging,
requiring an intricate probabilistic analysis (based on the level crossing technique), which is
the core of this paper. To the best of the authors’ knowledge, a steady-state analysis as in this
paper is nowhere to be found. Only in [25], by Moinzadeh and Nahmias, has the policy studied
here been considered, albeit in a different model and using approximations. They consider a
model with a continuous review policy, two supply modes, deterministic lead times, and ran-
dom demand (whereas we have stochastic lead times and deterministic demand); their goal is
to find the order quantities and reorder points for each supply mode in order to minimize total
inventory costs.

Among the other studies that assume the use of supply modes with deterministic lead times,
Moinzadeh and Schmidt [26] present a model based on a one-for-one (S – 1, S) inventory pol-
icy with two supply options, where the decision on whether to place an order with the regular
or emergency supplier depends on the age of the pending order. Jain et al. [19] consider an
(R, Q) model in which two transportation modes exist, each with a deterministic lead time and
fixed plus variable shipping cost structure. After the manufacturing lead time, the buyer must
decide how much of the order to ship via each transportation mode under stochastic demand.
Several papers on multiple sourcing consider a fast emergency supplier, which can be used
to avoid shortages by placing an emergency order (based on some triggering condition). For
example, Tagaras and Vlachos [30] assume an emergency order is placed when the probability
of stockout is high; Chiang and Gutierrez [12] propose an ‘indifference level’ for an inventory
that triggers a second order; in Johansen and Thorstenson [20], the reorder point for the emer-
gency supplier is a function of the time until the regular order arrives. Axsäter [1, 2] and Huang
et al. [18] propose and study several heuristic decision rules for triggering emergency orders.

Some references on models with multiple sourcing, and with stochastic lead times, are
[11], [15], [22], [23], [27], and [29]. Chiang and Chiang [11] propose an inventory model with
one supplier and multiple deliveries. They assume the use of a continuous review policy with
normally distributed demand, constant lead time, and a predetermined service level. In their
model an order of size Q is placed at the beginning of the replenishment cycle, which is split
between deliveries with interarrival times. A similar approach is found in Chiang [10], who
proposes the use of one supply mode and multiple deliveries for periodic review. Pan et al. [27],
based on order statistics, present expressions to estimate the parameters for the distributions of
the effective lead time and time between arrivals when two supply modes with identical real-
time distributions are used. As an example, they consider three lead time distributions: normal,
uniform, and exponential. They observe a reduction in the mean and variance of the effective
lead time, compared with the use of a single supply mode.

Most of the relevant models that assume stochastic lead times also assume (R, Q) inventory
policies, and seek the optimal order quantity, reorder point, and the proportion of order split-
ting, as in Lau and Lau [22] and Lau and Zhao [23], the latter accounting for stochastic demand.
These papers show that most of the savings from order splitting result from a reduction in cycle
stock, which often exceeds the cost savings due to associated safety stock reduction. Ramasesh
et al. [29] show that, under the assumption of identical lead-time distributions for the differ-
ent supply modes, an optimal solution is found when the order is split equally and that the
savings come from reduced holding and backordering costs. A different approach is presented
by Ganeshan, Tyworth, and Guo [15], who propose a discounting option for the less reliable
supplier, and present exchange curves to define when it is beneficial to split the order, as well
as the percentage discount necessary to make the model attractive.
Finally we refer to Poormoaid et al. [28] for an extensive recent review of the literature on continuous review and/or periodic review inventory models with emergency shipments; to Zipkin [34] for a detailed introduction to \((R, Q)\) models; and to Federgruen and Wang [14] for conditions for monotonicity of the optimal policy parameter in a wide range of \((R, Q)\) models.

The remainder of the paper is organized as follows. The model is described in more detail in Section 2. In Section 3 we discuss the regenerative structure of the inventory level process. Its formidable steady-state distribution is subsequently determined in Section 4. We derive the rate of emergency deliveries in Section 5 and that of normal deliveries in Section 6. Section 7 is devoted to a variant of the model with a somewhat different procedure regarding normal orders. Finally, Section 8 contains some suggestions for further research. In particular, we briefly indicate how our steady-state results could be used for optimization purposes – an important topic but beyond the scope of the present paper.

2. Model description

In this section we describe the model in more detail. The inventory level process \(V = \{V(t) : t \geq 0\}\) under consideration is a jump-fluid process with state-dependent deterministic release rate: for every state \(x > 0\) the release rate (due to the fluid demand) between jumps is \(\alpha(x)\). The jumps are all upwards and of size \(q\); they represent the arrival (delivery) of orders. There are two types of orders: normal orders and emergency orders. The lead times for both orders are exponentially distributed. Below we successively discuss the release rate process, the policy for placing normal orders and emergency orders, and the lead times in more detail.

Release rate assumption. We assume that

\[
A(x) := \int_0^x \frac{1}{\alpha(w)} \, dw < \infty,
\]

for every \(x > 0\). The latter assumption simply says that level 0 can be reached in a finite amount of time, since for every \(0 \leq x_1 < x_2\), \(A(x_2) - A(x_1)\) is the time it takes to go from level \(x_2\) all the way down to level \(x_1\) if no jumps (deliveries) occur in between. We refer to Harrison and Resnick [16] for an early study of inventory and storage processes with a state-dependent release rate, and to Boxma et al. [7] for a steady-state analysis of storage processes with state-dependent input and output rates.

The finiteness assumption (2.1) simplifies the analysis of the model, but is not necessary. It is enough to assume that \(\int_a^x 1/(\alpha(w)) \, dw < \infty\). This means that level \(a\) can be reached from any level \(x > a\).

Remark 2.1. In practice, the price often cannot be changed continuously over time. Thus there are certain predetermined levels such that each of their downcrossings leads to a change in the price. An example is the case when there are switchover levels \(\gamma_1, \gamma_2, \ldots, \gamma_n\), such that

\[
\alpha(x) = \begin{cases} 
\alpha_1, & 0 < x \leq \gamma_1, \\
\alpha_2, & \gamma_1 \leq x < \gamma_2, \\
\ddots & \\
\alpha_n, & \gamma_{n-1} \leq x < \gamma_n.
\end{cases}
\]

That is, the controller switches between \(n\) different demand rates for which it is natural to assume that \(\alpha_1 < \alpha_2 < \cdots < \alpha_n\).
The policy for placing normal and emergency orders. The jumps of $V$ represent the arrival (delivery) of normal and emergency orders. If just before downcrossing level $b$ no order is pending, the controller places a normal order, of size $q$, immediately after the downcrossing of level $b$; but if at least one order is pending (a normal order or an emergency order) she will not place a new one. If just before downcrossing level $a$ (recall that $a < b$) no emergency order is pending, the controller places an emergency order, also of size $q$, immediately after the downcrossing. This done even if a normal order is still pending. This policy is natural, since level $a$ can be interpreted as a warning level: the inventory level approaches 0 and the controller would like to avoid the risky situation of an empty system, which is interpreted as a period of unsatisfied demand (during which sales may be lost). If just before downcrossing level $a$ no normal order is pending, then also a normal order is placed. This policy ensures that just after hitting level $a$ exactly two orders are pending, one normal and one emergency. Since the lead times of these orders are independent and both exponentially distributed, $V$ is a regenerative process with, as regeneration epochs, the successive epochs at which $V$ hits level $a$ from above. In Lemma 3.1 this will be formally stated and proved. A level below $a$ is considered to be precariously low, so that having two orders underway seems justified.

Remark 2.2. In Section 7 we study the following variant of the policy described above. If just before downcrossing level $b$ no normal order is pending, the controller places a normal order, even if an emergency order is pending. The implication is that, in this variant, there is always already a normal order pending when level $a$ is reached.

Lead times assumptions. The lead times for normal orders, and also for emergency orders, are assumed to be independent and identically distributed; normal lead times and emergency lead times are also independent of each other. We suppose that normal lead times are exponentially distributed with rate $\lambda_n$, while emergency lead times are exponentially distributed with rate $\lambda_e$. It would be natural to assume that $\lambda_e > \lambda_n$, as an emergency lead time should typically be stochastically shorter than a normal lead time; however, in our analysis such a condition is not necessary. Thus there are two independent suppliers and from each of them one may make just one order at a time. The lead times cannot be expedited.

Parameter assumptions. In this study we restrict attention to parameters $a, b, q$ satisfying $0 < a < b < q$ and $b < a + q$ (see Figure 1). This ordering seems natural, but one can think of other parameter relations, such as $0 < a < q < b$ or $a + q < b$. In principle the derivations for such other cases are similar. For example, if $a + q < b$, it can happen that after the arrival of the last pending order the inventory level is still below $b$ but our policy does not indicate a new order. Therefore one might be inclined to consider other parameter combinations. If one wants to do optimization, then of course all possible orderings have to be taken into account.

3. The regenerative structure of the inventory level process

In this section we discuss the regenerative structure of the inventory level process, with the periods between two successive downcrossings of level $a$ as regeneration cycle.

The process $V$ is not a Markov process (relative to its natural filtration), even though the lead times are exponential. To see this, consider $V$ for example at some time instant $t$ at which it hits level $b$. Then either one order is pending (possibly an emergency order) or no order is pending, in which case a new normal order is issued. Therefore the process continues with one pending order whose residual lead time can be $\exp(\lambda_n)$ or $\exp(\lambda_e)$, both cases occurring with positive probability. However, from the evolution of $V$ before $t$ it might be known that exactly
two jumps have happened before the first hitting of $b$ at time $t$; in this case it is clear that the two jumps are due to the arrivals of the normal and the emergency order which were pending at the start of the current cycle, and now the order which is pending immediately after $t$ must be a normal one. Hence the conditional probability law of the process after $t$ is different from the one only conditional on $V(t) = b$. In other words, given $V(t)$, the future development from $t$ on depends on the history prior to $t$. Thus $V$ is not a Markov process. Note however that the joint process of $V$ and the numbers of pending normal and emergency orders is a Markov process, whereas $V$ alone is also not Markovian relative to the filtration generated by the three processes together. Furthermore, the jump process of $V$ is not a Poisson process, since the arrival times of future deliveries depend on the history of $V$. Notwithstanding, as we shall prove below, $V$ is a regenerative process, governed by three known parameters – the decision variables $(a; b; q)$ – and the known release rate function $\alpha(x)$.

**Lemma 3.1.** $V$ is a regenerative process and the time between two downcrossings of level $a$ is a regenerative cycle that starts with two pending orders.

**Proof.** We focus on the question of how $V$ enters the set of states $[0, a)$. Obviously $V$ enters this set at a downcrossing time of level $a$. As we have seen in Section 2, at that downcrossing time an emergency order was still pending or else a new emergency order is placed, and the same holds for a normal order. Hence exactly two orders, one normal and one emergency, are pending when the inventory process decreases from $a$ onwards. The assumption that all lead times are independent and exponentially distributed now implies that $V$ is a regenerative process, with regeneration epochs beginning and ending at the time instants at which level $a$ is downcrossed. □

We are almost ready to derive the steady-state distribution of $V$ and close this section with an observation that plays an important role in that derivation, and – first – with a lemma that provides some useful insight. In this second lemma we relate the number of pending orders with certain subsets of the sample space of $V$. 

---

**Figure 1.** The inventory level $V(t)$, with state-dependent release rate, normal and emergency orders and constant order sizes $q$; $C_i$ is the length of a regenerative cycle.
Lemma 3.2. The state space of $V$ is $[0, a + 2q)$ and the following implications hold:

(a) $V \in [0, a) \implies \{\text{two orders pending}\}$.
(b) $V \in [a, b) \implies \{\text{exactly one order pending}\}$.
(c) $V \in [b, a + q) \implies \{\text{at most one order pending}\}$.
(d) $V \in [a + q, a + 2q) \implies \{\text{no orders pending}\}$.

Proof. Since there are at most two orders pending, and the process decreases in between order deliveries, while $b < a + q$, we have $V < a + 2q$.

(a) This assertion was already derived in the proof of Lemma 3.1.

(b) Consider the moment of the first delivery in the cycle. Immediately after this epoch, $V \in [q, a + q)$. There are two possibilities.

   (i) The first delivery is of the emergency type. Then no more emergency orders will be placed until the end of the cycle. By Lemma 3.1, after $V$ downcrosses level $b$ only the normal order is pending.

   (ii) The first delivery is of the normal type. Then the emergency order is still pending. If it continues to be pending until the end of the cycle, no more normal orders will be placed and level $a$ will be reached with an emergency order pending. If the delivery of the emergency order occurs before the end of the cycle, immediately after that delivery $V \geq q > b$. Once $V$ returns to level $b$ a normal order is placed and by Lemma 3.1, during the time period it takes to go from level $b$ down to level $a$, a normal order is always pending.

(c) As in (b), immediately after the first jump in the cycle we have $V \in [q, a + q)$. If during the way down towards level $b$ no jump (delivery) occurs, $V$ will reach level $b$ with one pending order. However, if during the way down towards level $b$ the other jump (delivery) occurs, $V$ will reach level $b$ with no pending orders.

(d) $V$ enters this region only if the second delivery occurs in relatively close proximity to the first one. Another jump will then only be possible after $V$ reaches level $b$. □

Remark 3.1. Suppose that a normal order is pending right after a downcrossing of level $b$. This implies that no emergency order is pending (see statement (b) of Lemma 3.2). During a time period between a downcrossing of level $b$ and the first subsequent downcrossing of level $a$, there is now a geometrically distributed number of normal deliveries, with parameter $e^{-\lambda_0[A(b)−A(a)]]}$ (which is the probability of zero normal deliveries during this time interval). To see this, recall that order lead times are memoryless. If the delivery occurs before $V$ downcrosses level $a$, $V$ jumps above level $b$. When, later, $V$ downcrosses level $b$ again the controller will face a probabilistic replication of the situation at the previous downcrossing of level $b$. This means that there are no interruptions during the time period of going from level $b$ down to level $a$, and the probability of the latter event is $e^{-\lambda_0[A(b)−A(a)]}$.

4. Steady-state analysis

In this section we determine the steady-state density $f(\cdot)$ of the inventory level process $V$. Notice that this density always exists, and has finite support $(0, a + 2q)$. As observed in
the previous section, \( V \) is not a Markov process. Moreover, the jump process representing the delivery of orders is not a Poisson process since given the ‘present’ value of \( V \), future deliveries depend on past deliveries. Thus PASTA (Poisson arrivals see time averages) cannot be applied in order to find the steady-state law by the traditional LCT (level crossing technique: see [8] and [13]) concept of equating the rate of upcrossings of any fixed level \( x \) by \( V \) with the corresponding rate of downcrossings. We obtain \( f(\cdot) \) via a (as far as we know) novel application of the LCT. Introduce, for \( 0 < x < a + 2q \),

\[
\theta(x) := E[\text{number of downcrossings of level } x \text{ during a regenerative cycle}].
\]

Henceforth \( C \) denotes the length of an arbitrary regenerative cycle, the time between two successive downcrossings of level \( a \). We relate \( f(x) \) to \( \theta(x) \) in the following lemma. Subsequently we determine \( \theta(x) \) in Theorem 4.1.

**Lemma 4.1.** We have the following relation between the stationary inventory level density \( f(x) \) and the mean number of downcrossings in a cycle, \( \theta(x) \): for all \( x \in [0, a + 2q) \),

\[
\theta(x) = \frac{\alpha(x)f(x)}{\alpha(a)f(a)}, \tag{4.1}
\]

so that

\[
f(x) = \frac{\alpha(a)f(a)}{\alpha(x)} \theta(x). \tag{4.2}
\]

**Proof.** Since \( \theta(x) \) is the mean number of downcrossings per regenerative cycle, and \( \mathbb{E}[C] \) the mean cycle length, the downcrossing rate of level \( x \) equals \( \theta(x) / \mathbb{E}[C] \). On the other hand, the downcrossing rate of level \( x \) also equals \( \alpha(x)f(x) \). To prove (4.1), and hence also (4.2), it remains to show that

\[
\mathbb{E}[C] = \frac{1}{\alpha(a)f(a)}.
\]

This relation follows by observing that level \( a \) is downcrossed exactly once per cycle, and that the downcrossing rate of level \( a \) equals \( \alpha(a)f(a) \).

We now derive some properties of \( f(\cdot) \) (under the assumption that \( A(x) < \infty \) for all \( x > 0 \)). First, let \( \pi \) be the steady-state probability of having zero inventory. By the normalizing condition we have

\[
\int_0^{a+2q} f(x) \, dx = 1 - \pi.
\]

When \( A(x) = \infty \), as for example in the shot noise case \( \alpha(w) = w \), we have \( \pi = 0 \) and the normalizing condition is

\[
\int_0^{a+2q} f(x) \, dx = 1.
\]

In addition, we have some continuity properties, because the inventory level process decreases in a continuous way. In particular,

\[
f(x-) = f(x+) \quad \text{for } x = a, b, a + q, b + q, 2q. \tag{4.3}
\]

However, \( f(\cdot) \) is not continuous at \( q \), because there is a positive probability of reaching \( q \) from 0:

\[
\alpha(q)[f(q-) - f(q+)] = (\lambda_n + \lambda_e)\pi.
\]
Finally, LCT does work at level 0, where the upward jumps do occur according to a Poisson process:

\[ \alpha(0)f(0) = (\lambda_n + \lambda_e)\pi. \]  

(4.4)

We are now in a position to introduce the steady-state equations which relate the rate of downcrossings among different levels, and thus to obtain expressions for \( \theta(x) \) which, via (4.2), translate into expressions for the content level density \( f(x) \). In order to ease the notation we use the abbreviations

\[ f_n := \frac{\lambda_n}{\lambda_n + \lambda_e}, \quad f_e := \frac{\lambda_e}{\lambda_n + \lambda_e}, \]

which are the probabilities that a normal order comes first or that an emergency order comes first, and

\[ J_\sigma(x; w) := e^{-\sigma[A(x) - A(w)]}. \]

Observe that \( J_\sigma(x; w) \) is the probability that an \( \exp(\sigma) \)-distributed random variable exceeds the time to decrease from \( x \) to \( w \), if no jumps occur. We shall take \( \sigma = \lambda_n, \lambda_e, \lambda_n + \lambda_e \) below. In particular, if both a normal order and an emergency order are pending when the process is at level \( x \), then \( J_{\lambda_n+\lambda_e}(x; w) \) is the probability that level \( w \) is reached from level \( x \) before an order is delivered. In the latter case, we even write \( J(x; w) := J_{\lambda_n+\lambda_e}(x; w) \). It will be convenient, in the next theorem, to observe that

\[ dJ(a; w) = \frac{\lambda_n + \lambda_e}{\alpha(w)} \exp\left(-\frac{1}{\alpha(x)} \int_w^a dx\right) dw \]

represents the probability that the first upward jump of the inventory level process, after \( a \) has been downcrossed, occurs when the level is between \( w \) and \( w + dw \), while \( J(a; 0) \), the atom at zero, is the probability that zero has been reached from \( a \) before an upward jump has occurred.

**Theorem 4.1.** The following expressions hold for the expected number of downcrossings \( \theta(x) \) of level \( x \) during a regenerative cycle.

(a) \( 0 \leq x < a \):

\[ \theta(x) = J(a; x). \]  

(4.5)

(b) \( a \leq x < b \):

\[ \begin{align*}
\theta(x) &= f_e \frac{1}{J_{\lambda_n}(x; a)} \\
&\quad + f_n \int_0^a \left[ (1 - J_{\lambda_e}(q + w; x)) \frac{1}{J_{\lambda_n}(x; a)} \\
&\quad + (J_{\lambda_e}(q + w; x) - J_{\lambda_e}(q + w; a)) \left( 1 + \frac{1}{J_{\lambda_n}(x; a)} \right) \\
&\quad + J_{\lambda_e}(q + w; a) \right] dJ(a; w). \end{align*} \]  

(4.6)
(c) $b \leq x < q$:

$$
\theta(x) = f_e \int_0^a \left[ (1 - J_{\lambda_n}(q + w; x)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; b)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ J_{\lambda_n}(q + w; b) \frac{1}{J_{\lambda_n}(b; a)} \right] \, dJ(a; w) 
+ f_n \int_0^a \left[ (1 - J_{\lambda_n}(q + w; x)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; b)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ J_{\lambda_n}(q + w; b) \frac{1 - J_{\lambda_n}(b; a)}{J_{\lambda_n}(b; a)} \right] \, dJ(a; w) 
+ f_n \int_{x-q}^a \left[ (1 - J_{\lambda_n}(q + w; x)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; b)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ (J_{\lambda_n}(q + w; b) - J_{\lambda_n}(q + w; a)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ J_{\lambda_n}(q + w; a) \right] \, dJ(a; w) 
+ f_n \int_0^{x-q} \left[ (1 - J_{\lambda_n}(q + w; b)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; b) - J_{\lambda_n}(q + w; a)) \frac{1}{J_{\lambda_n}(b; a)} \right] \, dJ(a; w). 
$$

(4.7)

(d) $q \leq x < a + q$:

$$
\theta(x) = f_e \int_{x-q}^a \left[ (1 - J_{\lambda_n}(q + w; x)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; b)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ J_{\lambda_n}(q + w; b) \frac{1}{J_{\lambda_n}(b; a)} \right] \, dJ(a; w) 
+ f_n \int_{x-q}^a \left[ (1 - J_{\lambda_n}(q + w; x)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; b)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ (J_{\lambda_n}(q + w; b) - J_{\lambda_n}(q + w; a)) \left( 1 + \frac{1}{J_{\lambda_n}(b; a)} \right) 
+ J_{\lambda_n}(q + w; a) \right] \, dJ(a; w) 
+ f_n \int_{x-q}^a \left[ (1 - J_{\lambda_n}(q + w; b)) \frac{1}{J_{\lambda_n}(b; a)} 
+ (J_{\lambda_n}(q + w; b) - J_{\lambda_n}(q + w; a)) \frac{1}{J_{\lambda_n}(b; a)} \right] \, dJ(a; w). 
$$

(4.8)
(e) \( a + q \leq x < b + q \):

\[
\theta(x) = f_e \int_0^a \left( 1 - J_{\lambda_e}(q + w; b) \right) \left( 1 + \beta_n(x) + J_{\lambda_n}(q + w; b) \beta_n(x) \right) \, dJ(a; w)
+ f_n \int_0^a \left( 1 - J_{\lambda_e}(q + w; b) \right) \left( 1 + \beta_n(x) + J_{\lambda_n}(q + w; b) \beta_e(x) \right) \, dJ(a; w),
\]

where \( \beta_n(x) \) and \( \beta_e(x) \) are given by

\[
\beta_n(x) = \frac{1 - J_{\lambda_n}(b; x - q)}{J_{\lambda_n}(b; a)},
\]

\[
\beta_e(x) = 1 - J_{\lambda_e}(b; x - q) + (1 - J_{\lambda_n}(b; x - q)) \frac{1 - J_{\lambda_e}(b; a)}{J_{\lambda_n}(b; a)}.
\]

(f) \( b + q \leq x < 2q \):

\[
\theta(x) = f_e \int_0^a \left( 1 - J_{\lambda_n}(q + w; x - q) \right) \, dJ(a; w)
+ f_n \int_0^a \left( 1 - J_{\lambda_e}(q + w; x - q) \right) \, dJ(a; w).
\]

(g) \( 2q \leq x < a + 2q \):

\[
\theta(x) = f_e \int_0^a \left( 1 - J_{\lambda_n}(q + w; x - q) \right) 1_{x - q \geq x - q} \, dJ(a; w)
+ f_n \int_0^a \left( 1 - J_{\lambda_e}(q + w; x - q) \right) 1_{x - q \geq x - q} \, dJ(a; w).
\]

**Proof.** (a) \( 0 \leq x < a \). In this region the number of downcrossings of level \( x \) is at most one. Hence the mean number of downcrossings \( \theta(x) \) of level \( x \in [0, a) \) equals the probability that \( x \) is downcrossed, which is \( J(a; x) = J_{\lambda_n + \lambda_e}(a; x) \).

(b) \( a \leq x < b \). By Lemma 3.2, in this region exactly one order is pending; it can either be the normal order or the emergency order, but not both of them. Accordingly, to determine \( \theta(x) \) we distinguish the events where the emergency delivery occurs first and the normal delivery occurs first. We shall now explain the terms of (4.6).

- **The term in line 1 of (4.6).** This term corresponds to the case that the emergency delivery comes first. The probability of that event is \( f_e = \lambda_e / (\lambda_n + \lambda_e) \). After the arrival of the emergency order, only the normal order is pending, and there will be no more emergency order in the rest of the cycle. Immediately after the arrival of the emergency order, \( \mathbf{V} > b \), and level \( x \in [a, b) \) will be downcrossed at least once until the end of the cycle. That number of downcrossings \( \mathbf{D}_x \) is geometrically distributed,

\[
\mathbb{P}(\mathbf{D}_x = m) = J_{\lambda_n}(x; a)(1 - J_{\lambda_n}(x; a))^{m-1}, \quad m = 1, 2, \ldots,
\]

with mean \( \mathbb{E}\mathbf{D}_x = 1/J_{\lambda_n}(x; a) \). Indeed, each time when \( x \) has been downcrossed, the probability that level \( a \) will be reached (and hence the cycle will be ended) before another normal delivery arrives (in which case \( b \) will be upcrossed and later downcrossed again, and subsequently \( x \) will be downcrossed again) equals \( J_{\lambda_n}(x; a) \).
The term in lines 2–4 of (4.6). The probability that the normal delivery arrives when the state is in \((w, w + dw)\) (for \(0 < w < a\)) and before the emergency delivery is \(f_nJ(a; w)\). Immediately after the arrival of the normal delivery the state becomes \(q + w\) (by definition \(q + w > b\)), and from this moment only the emergency order is pending. If that emergency delivery arrives before level \(x\) is downcrossed (which has probability \(1 - J_{\lambda_n}(q + w; x)\)), the process jumps above \(b\), and at the next downcrossing of \(b\) a normal order is pending. Hence one again has \(1/J_{\lambda_n}(a; x)\) downcrossings of level \(x\). If the emergency delivery arrives after \(x\) is downcrossed but before \(a\) is downcrossed (probability \(J_{\lambda_n}(q + w; x)(1 - J_{\lambda_n}(x; a)) = J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; a)\)), then \(x\) has already been downcrossed once, and the process jumps up by \(q\), subsequently \(b\) is downcrossed again and a normal order is pending, and we have on average \(1/J_{\lambda_n}(x; a)\) more downcrossings of level \(x\). If the emergency delivery does not arrive before \(a\) is reached (probability \(J_{\lambda_n}(q; a)\)), then we have just one downcrossing of \(x\) in the cycle.

At this stage we would like to emphasize that, with probability \(J(a; 0)\), zero is reached before an order has arrived. Indeed, the atom at zero in the integral in (4.6) gives the contribution

\[
\int_{w=0}^{a} \frac{1}{J_{\lambda_n}(b; a)} \, dz \, J(a; w) = \int_{w=0}^{a} \frac{1}{J_{\lambda_n}(b; a)} \, dz \, J_{\lambda_n}(q; a)
\]

These terms correspond to the case in which no delivery arrives before level 0 is reached, after which a normal delivery arrives first (probability \(f_nJ(a; 0)\)). In each of cases (c)–(f) below, where an integral \(\int_{w=0}^{a} z(w) \, dJ(a; w)\) appears, there is a similar contribution \(z(0)J(a; 0)\) from the atom at 0 when \(w\) is in the interior; similarly for \(\int_{w=0}^{a} z(w) \, dJ(a; w)\) in case (d) below.

(c) \(b \leq x < q\). We shall explain all terms of (4.7) successively.

The terms in lines 1–3 of (4.7). This concerns a case in which the emergency delivery arrives first, when the state is in \((w, w + dw)\) (or in 0, the contribution from the atom). The process now jumps to \(q + w > x\). From this moment only the normal order is pending. We now consider the various possibilities for the arrival of a normal delivery.

(i) It arrives with probability \(1 - J_{\lambda_n}(q + w; x)\) before \(x\) is crossed. That leads to a jump in the process, some time later a crossing of \(x\), and no further orders until subsequently \(b\) is reached. The mean number of crossings of \(x\) is now

\[
1 + \frac{1}{J_{\lambda_n}(b; a)} = \frac{1}{J_{\lambda_n}(b; a)}.
\]

(ii) It arrives with probability \(J_{\lambda_n}(q + w; x) - J_{\lambda_n}(q + w; b)\) when \(x\) has already been crossed a first time, but before \(b\) is reached. After \(b\) is reached, there are on average

\[
\frac{1}{J_{\lambda_n}(b; a)}
\]

more crossings of \(x\).

(iii) It arrives with probability \(J_{\lambda_n}(q + w; b)\) not before \(b\) is reached. After \(b\) is reached, there are on average

\[
\frac{1 - J_{\lambda_n}(b; a)}{J_{\lambda_n}(b; a)}
\]

more arrivals of normal deliveries and subsequent crossings of \(x\).
The terms in lines 4–6 of (4.7). This concerns a case in which the normal delivery arrives first, when the state is in \((w, w + dw)\) (or, again, in 0, the contribution from the atom). Lines 4–5 are completely analogous to lines 1–2. Line 6 differs more fundamentally from line 3. We get with probability \(J \lambda e(q + w; b) - J \lambda e(q + w; a)\) an emergency delivery between \(b\) and \(a\), giving rise to, on average, \(1/J_{\lambda a}(b; a)\) more crossings of \(x\). With probability \(J \lambda e(q + w; a)\) there is no emergency delivery before \(a\) is reached and the cycle is ended; in this case there was just one crossing of \(x\).

(d) \(q \leq x < a + q\). We successively explain all terms of the 11 lines of (4.8).

Lines 1–3 of (4.8). These lines concern the case in which an emergency delivery occurs first, when the state is in \((w, w + dw)\), for \(x - q < w < a\) (so now there is no contribution from an atom at 0). Distinguish between the three possibilities in which:

(i) the normal delivery occurs before \(x\) is crossed,
(ii) the normal delivery occurs between the crossings of \(x\) and \(b\), and
(iii) the normal delivery has not yet arrived when \(b\) is crossed.

These three possibilities give rise to, on average,
\[
\frac{1}{J_{\lambda a}(b; a)}, \quad 1 + \frac{1}{J_{\lambda a}(b; a)} \quad \text{and} \quad \frac{1}{J_{\lambda a}(b; a)}
\]
crossings of \(x\).

Lines 4–5 of (4.8). These lines concern the case in which the jump is from \(w \in [0, x - q)\) and hence ends up in \(q + w < x\). Now there are two possibilities:

(i) the normal delivery arrives before \(b\) is reached,
(ii) the normal delivery has not yet arrived before \(b\) is reached.

These two possibilities give rise to, on average
\[
\frac{1}{J_{\lambda a}(b; a)} \quad \text{and} \quad \frac{1 - J_{\lambda a}(b; a)}{J_{\lambda a}(b; a)}
\]
crossings of \(x\).

Lines 6–9 of (4.8). The normal delivery arrives first, when the state is in \((w, w + dw)\), where \(x - q \leq w < a\) (lines 6–9). Now distinguish between the four possibilities in which:

(i) the emergency delivery arrives before \(x\) is crossed,
(ii) it arrives between the crossings of \(x\) and \(b\),
(iii) it arrives between the crossings of \(b\) and \(a\),
(iv) it does not arrive before the end of the cycle.

The first three possibilities now give rise to, on average,
\[
\frac{1}{J_{\lambda a}(b; a)}, \quad 1 + \frac{1}{J_{\lambda a}(b; a)} \quad \text{and} \quad 1 + \frac{1}{J_{\lambda a}(b; a)}
\]
crossings of \(x\) (in fact we could have combined possibilities (ii) and (iii)). Possibility (iv) gives exactly one downcrossing of \(x\) in the whole cycle.
It is easily seen that $\beta(b)$ process jumps to level $q$ right-hand side of (4.12) concerns the case in which the emergency delivery arrives first, the
\begin{equation}
J_{\lambda n}(b;x) - J_{\lambda n}(a;b)\beta_n(x),
\end{equation}

\begin{equation}
J_{\lambda n}(b;x) - J_{\lambda n}(a;b)\beta_n(x),
\end{equation}

It is easily seen that $\beta_n(x)$ and $\beta_e(x)$ satisfy the following two equations:
\begin{equation}
\beta_n(x) = (1 - J_{\lambda n}(b;x - q))(1 + \beta_n(x)) + (J_{\lambda n}(b;x - q) - J_{\lambda n}(b;a))\beta_n(x),
\end{equation}

\begin{equation}
\beta_e(x) = (1 - J_{\lambda n}(b;x - q))(1 + \beta_n(x)) + (J_{\lambda n}(b;x - q) - J_{\lambda n}(b;a))\beta_n(x).
\end{equation}

Hence, with $a \leq x - q < b$, equations (4.10) and (4.11) follow.

\begin{itemize}
  \item **Line 1 of (4.9).** After the arrival of the emergency delivery and the corresponding jump to $q + w$, only the normal order is pending. If it arrives before $b$ is reached, the process jumps up by $q$ and we have one crossing of $x$, and subsequently the process will reach $b$ and we have on average another $\beta_n(x)$ crossings of $x$. If it does not arrive before $b$ is reached, we get on average $\beta_n(x)$ crossings of $x$.
  \item **Line 2 of (4.9).** Similar to line 1, but now the normal delivery is the first to arrive when the process goes down from $a$ to 0, and only the emergency order is pending. If it arrives before $b$ is reached, the process jumps up by $q$ and we have one crossing of $x$, and subsequently the process will reach $b$ and we have on average another $\beta_n(x)$ crossings of $x$. If it does not arrive before $b$ is reached, we have on average $\beta_e(x)$ crossings of $x$.
\end{itemize}

(f) $b + q \leq x < 2q$. In this range there is at most one downcrossing of $x$ during a cycle. Hence $\theta(x)$ is also the probability that level $x$ is downcrossed during the cycle. The first term in the right-hand side of (4.12) concerns the case in which the emergency delivery arrives first, the process jumps to level $q + w$, and the normal delivery arrives before level $x - q$ is reached, resulting in a jump to a level above $x$. From here on no delivery can arrive until level $b$ is reached, and thereafter the process can no longer get higher than $b + q$ in the cycle. The second term does the same for the case in which the normal delivery arrives first.

(g) $2q \leq x < a + 2q$. In this range there is again at most one downcrossing of $x$ during a regenerative cycle. Hence $\theta(x)$ is again the probability that level $x$ is downcrossed during the cycle. The two terms in the right-hand side of (4.13) are the same as the first two terms in the right-hand side of (4.12). It should be noticed that here, unlike cases (b)–(f), there is no contribution from the atom at 0: if 0 is reached in the beginning of the cycle, then the level can no longer reach $2q$ during this cycle. □

**Remark 4.1.** A few terms in (4.7) can be combined, yielding a slight simplification. Observe that for $w \in [b, q)$, the term inside the first square brackets for $\theta(x)$ in (4.7) can be compressed into
\begin{equation}
\frac{1}{J_{\lambda n}(b;a)} + (J_{\lambda n}(q + w;x) - J_{\lambda n}(q + w;b)).
\end{equation}
A similar remark applies in some other formulas. We can now easily integrate that $1/(J_{\lambda_n}(b; a))$ term (taking the atom at 0 into account), which yields

$$f_r = \frac{1}{J_{\lambda_n}(b; a)}.$$

It is reasonably straightforward to verify that (4.3) holds, i.e., the continuity of $f(x)$, and hence of $\theta(x)$, in $x = a, b, a + q, b + q, 2q$.

Theorem 4.1 uniquely determines $\theta(x)$. However, it follows from (4.2) that the content level density $f(x)$ is thus only determined up to the yet unknown constant $f(a)$. The normalizing condition provides one additional equation, i.e., $\int_0^{a+2q} f(x) \, dx = 1 - \pi$, but the stationary probability $\pi$ of having zero inventory is also unknown as yet. LCT at level zero, as given in (4.4), gives us the extra equation that we need:

$$(\lambda_n + \lambda_e)\pi = \alpha(0)f(0),$$

in which an expression for $f(0)$ (or rather $f(0)/f(a)$) is obtained by substituting $x = 0$ in (4.5) and (4.2). It should be observed that $\pi$ is also the fraction of time in which there is zero inventory, during which no demand is satisfied.

**Remark 4.2.** In the region $x \in (0, a)$ we can also use the traditional methodology of LCT to obtain the same solution as given in (4.5). To see this, recall that two orders are pending below level $a$. Since the two lead times are independent and exponentially distributed, we conclude that, below level $a$, the arrival process of jumps is a Poisson process with rate $\lambda_n + \lambda_e$. LCT now gives the balance equation

$$\alpha(x)f(x) = (\lambda_n + \lambda_e)F(x),$$

where the right-hand side, with $F(x)$ the steady-state distribution of the inventory level, gives the rate of upcrossings of level $x \in (0, a)$. Rewriting this as

$$\frac{f(x)}{F(x)} = \frac{\lambda_n + \lambda_e}{\alpha(x)},$$

and integrating gives

$$F(x) = K \exp \left( (\lambda_n + \lambda_e) \int_0^x \frac{1}{\alpha(w)} \, dw \right),$$

with $K$ some constant. Differentiation yields

$$f(x) = \frac{\alpha(a)f(a)}{\alpha(x)}J(a; x),$$

in agreement with (4.5); the last equality was obtained by substituting $x = a$.

This approach does not work for $x > a$ since there the arrival process of the jumps is not a Poisson process.

**Remark 4.3.** For some special choices of the state-dependent release rate function $\alpha(\cdot)$, the expressions in Theorem 4.1 simplify considerably. We mention two special cases.

**Example 1:** $\alpha(x) \equiv \alpha$. In this case $A(x) \equiv x/\alpha$ and hence

$$J_{\sigma}(x; y) = e^{-(\sigma/\alpha)(x-y)}.$$
In this case all integrals appearing in Theorem 4.1 can be easily evaluated.

Example 2: \( \alpha(x) = rx \). In this case \( A(x) \) as defined in (2.1) would diverge. However, we can still determine

\[
J_\sigma(x; y) = \left( \frac{y}{x} \right)^{\sigma/r}, \quad x > y > 0,
\]

and it should be observed that \( J_\sigma(x; 0) = 0 \) (indeed, the origin can never be reached). Various terms in (4.5)–(4.13) now disappear, and others become relatively simple. These two special cases are quite natural. In the first example one might try to choose \( \alpha \) such that a particular profit function is optimized. In the second example the choice of \( r \) has to do with pricing: its value reflects the demand caused by a certain price setting.

5. Rate of emergency deliveries

In this section we determine the rate \( \rho_e \) of emergency deliveries. Let \( I \) be the number of emergency deliveries during an arbitrary regenerative cycle. We have \( \rho_e = \mathbb{E}[I]/\mathbb{E}[C] \). In the next theorem \( \rho_e \) is obtained by determining \( \mathbb{E}[I] \).

**Theorem 5.1.** The rate of emergency deliveries is given by

\[
\rho_e = \alpha(a)f(a) \left\{ f_e + f_n \int_0^a [1 - J_{\lambda_e}(q + w; a)] \, dJ(a; w) \right\}. \tag{5.1}
\]

**Proof.** First of all we observe that the numbers of emergency deliveries \( I_1, I_2, \ldots \) in successive cycles are independent, Bernoulli-distributed random variables, and hence \( \mathbb{E}[I] = \mathbb{P}(I = 1) \). Secondly, since \( \mathbb{E}[C] = 1/(\alpha(a)f(a)) \), we have \( \rho_e = \alpha(a)f(a)\mathbb{P}(I = 1) \). It remains to show that \( \mathbb{P}(I = 1) \) is given by the term between curly brackets in (5.1).

The cycle starts at level \( a \) where the two orders are pending. From this moment two disjoint events may happen, corresponding to the two terms between curly brackets in the right-hand side of (5.1). (i) The emergency delivery arrives first. The probability of this event is \( f_e \). (ii) The normal delivery arrives first and finds \( V \) at level \( w \) \((w = 0 \text{ gives the contribution } J(a; 0)[1 - J_{\lambda_e}(q; a)] \)). Then \( V \) jumps to level \( q + w \). An emergency order arrives before the inventory level decreases from \( q + w \) to \( a \); its probability equals \( 1 - J_{\lambda_e}(q + w; a) \).

6. Rate of normal deliveries

Let \( \rho_n \) be the rate of normal deliveries and \( \mathbb{E}[N] \) the mean number of normal deliveries per cycle. Then

\[
\rho_n = \frac{\mathbb{E}[N]}{\mathbb{E}[C]} = \alpha(a)f(a)\mathbb{E}[N].
\]

Below we determine \( \mathbb{E}[N] \).

We shall repeatedly use the fact that (see (4.14)) the mean number of normal deliveries from the moment \( b \) is reached while a normal order is pending is given by

\[
\mathbb{E}[D_b] - 1 = \frac{1 - J_{\lambda_n}(b; a)}{J_{\lambda_n}(b; a)}.
\]
Theorem 6.1. The rate of normal deliveries is given by

\[
\rho_n = \alpha(a) f(a) \left\{ \int_0^a \left[ (1 - J_{\lambda_n}(q + w; b)) \frac{1}{J_{\lambda_n}(b; a)} + J_{\lambda_n}(q + w; b) \frac{1 - J_{\lambda_n}(b; a)}{J_{\lambda_n}(b; a)} \right] dJ(a; w) \\
+ f_n \int_0^a \left[ 1 + (1 - J_{\lambda_n}(q + w; a)) \frac{1 - J_{\lambda_n}(b; a)}{J_{\lambda_n}(b; a)} \right] dJ(a; w) \right\}.
\] (6.1)

Proof. We need to show that the mean number \( \mathbb{E}[N] \) of normal deliveries per cycle is given by the expression between curly brackets in (6.1). First consider the first term between curly brackets. It corresponds to the case that the emergency delivery comes first. Now distinguish between the two cases in which (i) a normal delivery occurs before \( b \) is first downcrossed (probability \( 1 - J_{\lambda_n}(q + w; b) \)) and (ii) no normal delivery occurs before \( b \) is first downcrossed. In case (i) we get on average \( \mathbb{E}[D_b] \) normal deliveries, and in case (ii) one less.

Next consider the second term between curly brackets. This term corresponds to the case that the normal delivery comes first, and the inventory is at level \( w \), jumping to \( q + w \). That normal delivery contributes one unit to \( \mathbb{E}[N] \). With probability \( 1 - J_{\lambda_n}(q + w; a) \) an emergency delivery arrives before \( a \) is reached. It results in an upcrossing of level \( b \), which is on average followed by \( (1 - J_{\lambda_n}(b; a)) / J_{\lambda_n}(b; a) \) normal deliveries until the end of the cycle. \( \square \)

7. A model variant

In this section we briefly consider the following model variant. The only change with respect to the model of Section 2 is that now, when level \( b \) is downcrossed, a normal order is always placed when it is not already pending, even if an emergency order is already pending. As a result, it becomes less likely that level \( a \) is reached without any delivery having arrived.

Lemma 3.1 remains true in this variant: the period between two downcrossings of level \( a \) is still a regenerative cycle that starts with two pending orders. Statement (b) of Lemma 2 has to be adapted: for \( \lambda \in [a, b) \), at least one order is pending, instead of exactly one (and the normal order is always pending in this interval). Lemma 4.1, which relates the inventory level density \( f(x) \) and the mean number of downcrossings \( \theta(x) \), remains unchanged. Theorem 4.1 needs to be adapted, but only in the region \( x \in [b, b + q) \), which corresponds to the four cases (b)–(e) of the theorem. The reason is that the only difference from the model treated above occurs once level \( b \) is downcrossed. There is no change in the regions \( [b + q, 2q) \) and \( [2q, a + 2q) \) (cases (f) and (g)) because levels higher than \( b + q \) cannot be reached in this model variant after level \( b \) is downcrossed. There is no change in the region \( [0, a) \) (case (a)) because the regenerative cycle starts the same in the model variant as in the original model, with both a normal order and an emergency order pending. In \( [a, b + q) \) the analysis becomes slightly more complicated than in Theorem 4.1. However, all terms corresponding to cases in which the emergency order comes first in the regenerative cycle are unaltered, because a normal order now becomes pending once \( b \) is downcrossed, in both the variant and in the original model.

Below we show how formula (4.6) needs to be adapted; this concerns case (b), i.e. the region \( [a, b) \). Similar adaptations have to be made for the region \( [b, b + q) \).

Case (b) in the model variant: \( a \leq x < b \). The mean number of downcrossings of level \( x \) in the model variant equals

\[
\theta(x) = f_e \left\{ \frac{1}{J_{\lambda_n}(x; a)} \right\} + f_n \int_0^a \left[ (1 - J_{\lambda_n}(q + w; b)) \frac{1}{J_{\lambda_n}(b; a)} + J_{\lambda_n}(q + w; b) g(x) \right] dJ(a; w).
\] (7.1)
where \( g(x) \) is the mean number of downcrossings of level \( x \in [a, b] \) in a cycle, from the moment level \( b \) is downcrossed in a cycle, given that an emergency order is pending.

The rationale behind (7.1) is as follows.

- The first term in the right-hand side is the same as the first term in (4.6): this is the case in which the emergency delivery occurs first.
- The second term reflects the case in which the normal delivery occurs at level \( w \) and before the emergency delivery. The process jumps to \( q + w \). One possibility now is that the emergency delivery comes before \( b \) is downcrossed. From then on, there are on average \( 1 / J_{\lambda_n}(x; a) \) downcrossings of \( x \) in the remainder of the cycle. The second possibility is that the emergency delivery has not occurred before \( b \) is reached. From then on, one has on average \( g(x) \) downcrossings of level \( x \).

Here \( g(x) \) satisfies the following relation:

\[
g(x) = J_{\lambda_n + \lambda_e}(b; a) \left( 1 - J(b; x) \right) \frac{1}{J_{\lambda_n}(x; a)} + \frac{f_e(J(b; x) - J(b; a))}{1 + \frac{1}{J_{\lambda_n}(x; a)}} \left( 1 - J_{\lambda_n}(q + w; b) \right) \frac{1}{J_{\lambda_n}(x; a)} + J_{\lambda_e}(q + w; b) g(x) \left( 1 + \frac{1}{J_{\lambda_n}(x; a)} \right) \frac{dJ(b; w)}{J_{\lambda_n}(x; a)} + J_{\lambda_e}(q + w; b)(1 + g(x)) \right) dJ(b; w).
\]

\[(7.2)\]

\( g(x) \) can be trivially solved from this relation; observe that one can add the two terms with \( g(x) \) in the right-hand side of (7.2) into the product of \( g(x) \) and an integral from \( a \) to \( b \) that does not involve \( x \).

The rationale behind (7.2) is as follows. At the beginning of the remaining cycle, both a normal order and an emergency order are pending. There are now five cases, corresponding to the five lines in (7.2).

- **Case 1:** both orders remain pending until \( a \) is reached. That gives exactly one downcrossing of \( x \).
- **Case 2:** an emergency delivery occurs first, and before \( x \) is downcrossed. The process jumps up by \( q \) and eventually reaches \( b \) again, but there will be no more emergency order. We have on average \( 1 / J_{\lambda_n}(x; a) \) downcrossings of \( x \).
- **Case 3:** an emergency delivery occurs first, but between levels \( x \) and \( a \). Now \( x \) has already been downcrossed once, but thereafter the process jumps up by \( q \), and eventually reaches \( b \) again; now see Case 2 above.
- **Case 4:** a normal delivery occurs first, and before \( x \) is downcrossed. The process jumps up by \( q \). Now there are two sub-cases: either an emergency delivery occurs before \( b \) is reached, or it does not. In the former case one gets \( 1 / J_{\lambda_n}(x; a) \) downcrossings of \( x \), and in the latter case the process reaches \( b \) again and we are back in the old situation: one has on average \( g(x) \) more downcrossings of \( x \).
• Case 5: a normal delivery occurs first, but between levels $x$ and $a$. Now $x$ has already been downcrossed once, but thereafter the process jumps up by $q$. Now there are again two sub-cases, just as for Case 4.

Similar adaptations have to be made in cases (c)–(e) of Theorem 4.1.

8. Conclusions and suggestions for further research

We have presented a steady-state analysis of the inventory level of an $(R, Q)$ model with normal orders and emergency orders, which are delivered after exponentially distributed lead times. We have allowed a quite general state-dependent release rate for the inventory level.

Our results could be used for optimization purposes. Remember that $\rho_e$ and $\rho_n$ are the average number of emergency and normal orders per unit time, respectively, and $\pi$ is the fraction of time having zero inventory and $\mathbb{E}[V]$ is the steady-state mean inventory level. Clearly, the four quantities $\rho_e$, $\rho_n$, $\pi$, and $\mathbb{E}[V]$ are functions of the decision variables $(a, b, q)$. We have shown how to derive explicit expressions for them. A linear combination of these characteristics, say

$$ R(a; b; q) = k_e\rho_e + k_n\rho_n + k_u\pi + h\mathbb{E}[V], $$

would be a meaningful cost functional for evaluating the performance of the system. Here $k_e$ is the set-up cost per emergency order, $k_n$ is the set-up cost per normal order, $k_u$ is the penalty for a time unit of zero inventory, and $h$ is the holding cost per unit of inventory. As the underlying stationary distribution of the inventory level turns out to be of an extremely complicated form, calculating functionals such as $R(a; b; q)$ will certainly be a numerically challenging endeavor; in view of Remark 4.3, the case $\alpha(x) \equiv \alpha$ is probably the first candidate for such an optimization study.

Of course, next to costs there are also profits. The gross profit of the system can be measured by the average value of the sales. For this a price function $p(x)$ has to be specified, where $p(x)$ is the price charged for a unit when the inventory level is $x$. The inventory level process was shown to be regenerative, and it therefore makes sense to consider $V$ and $C$, the stationary inventory level and the first cycle length, respectively. Then, using the ergodic theorem for regenerative processes, the average sales value is given by

$$ S(a; b; q) = \mathbb{E}[p(V)] = \frac{1}{\mathbb{E}[C]} \mathbb{E} \left[ \int_0^C p(V(t)) \, dt \right] = \int_0^{a+2q} p(x)f(x) \, dx. $$

Given $p(x)$, the objective function to be maximized is $S(a; b; q) - R(a; b; q)$. The selection of a suitable price function will then be a second step for future research based on numerical work.

Next to optimization, one could consider some more variants of the main model under consideration. For example, one could take $q_e > q_n$ (larger emergency orders), or even allow random order sizes, or let the order size be state-dependent, too.

Acknowledgement

The authors are grateful to Mahmut Parlar for his valuable contributions.

Funding information

The research of Onno Boxma is partly funded by the NWO Gravitation Programme NETWORKS (grant 024.002.003). The research of David Perry was partly funded by ISF (Israel Science Foundation), grant 3274/19.
Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References


