Confidence-based optimization for the Newsvendor problem

Roberto Rossi[1] Steven D Prestwich[2] S Armagan Tarim[3] Brahim Hnich[4]

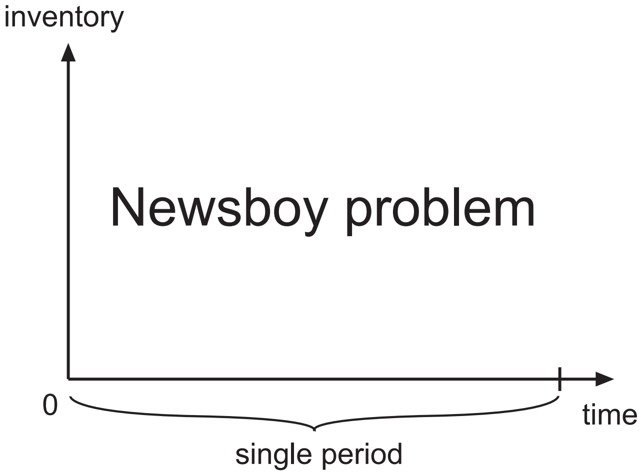
[1]University of Edinburgh, United Kingdom
[2]University College Cork, Ireland
[3]Hacettepe University, Turkey
[4]Izmir University of Economics, Turkey

Jun 2014, Università degli Studi di Bologna, Italy

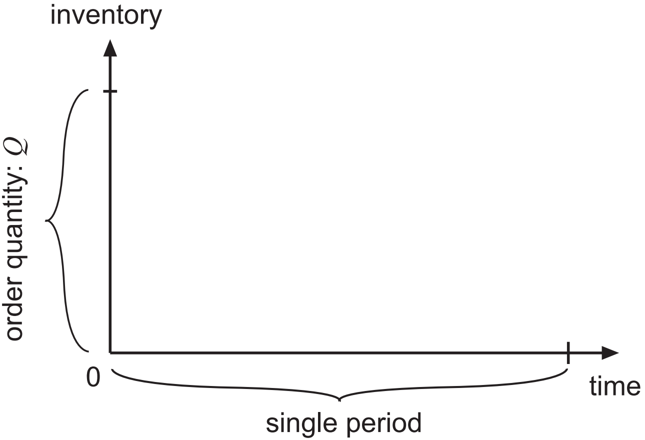
http://dx.doi.org/10.1016/j.ejor.2014.06.007

UNIVERSITY OF EDINBURGH
Business School

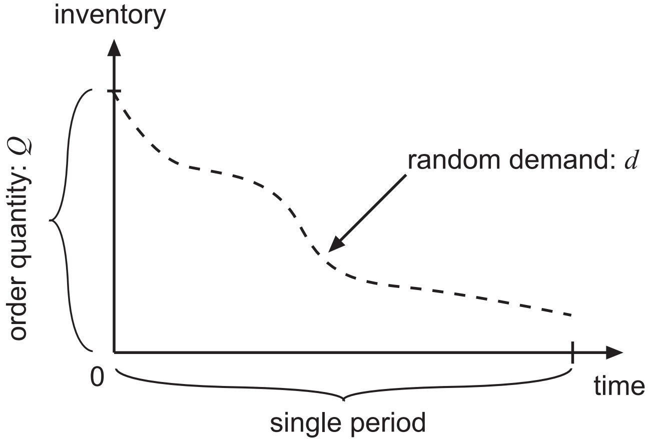
The Newsboy problem



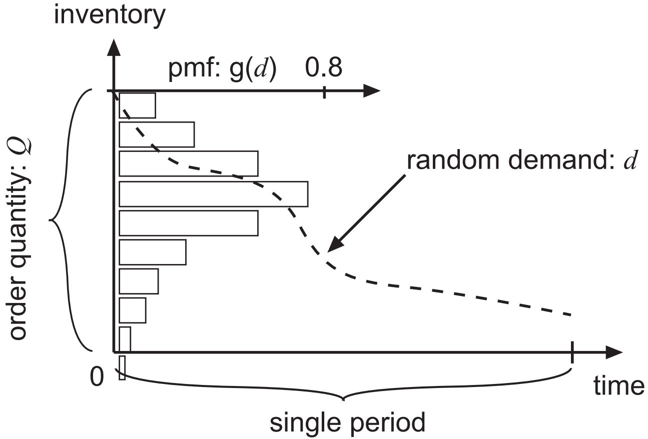
Order quantity



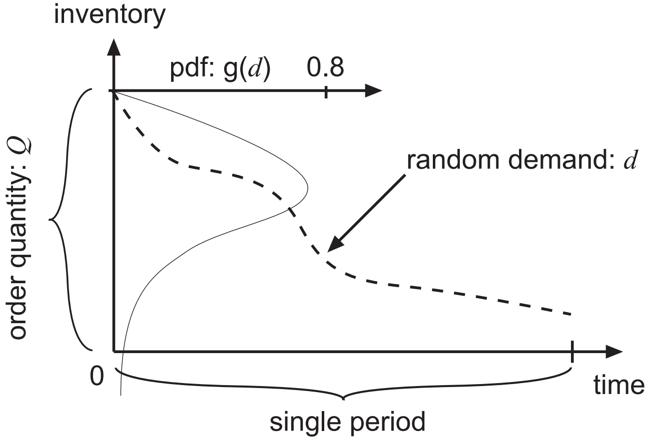
Demand structure



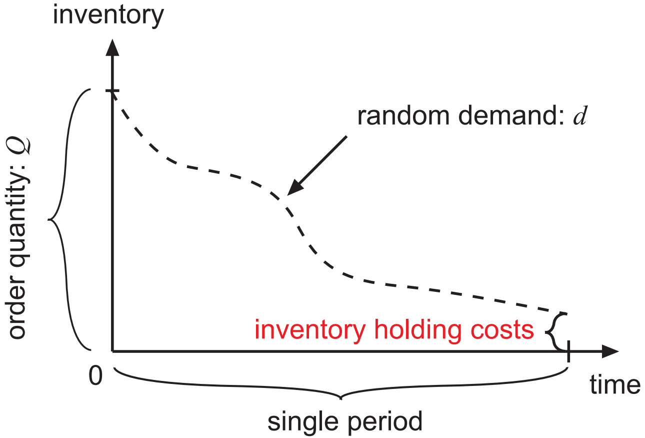
Demand structure



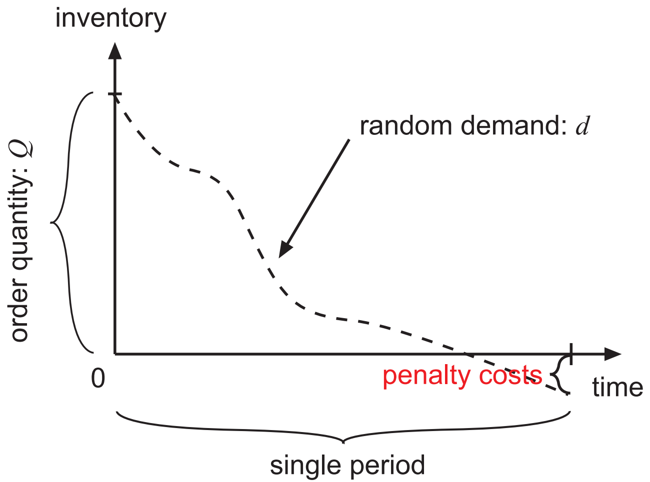
Demand structure



Cost structure



Cost structure



Mathematical formulation

Consider

- d: a **one-period** random demand that follows a **probability distribution** $f(d)$
- h: unit **holding cost**
- p: unit **penalty cost**

Let

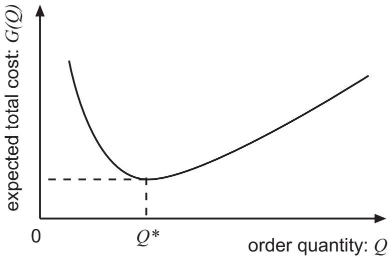
$$g(x) = hx^+ + px^-,$$

where $x^+ = \max(x, 0)$ and $x^- = -\min(x, 0)$.

The **expected total cost** is $G(Q) = E[g(Q - d)]$, where $E[\cdot]$ denotes the expected value.

Solution method

If d is continuous, $G(Q)$ is **convex.**



The optimal order quantity is

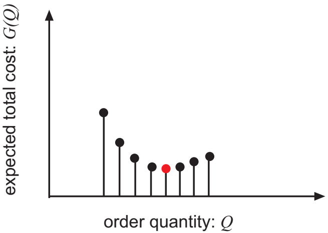
$$Q^* = \inf\{Q \geq 0 : \Pr\{d \leq Q\} = \frac{p}{p+h}\}.$$

Solution method

If d is discrete (e.g. Poisson),

$$\Delta G(Q) = G(Q+1) - G(Q) = h - (h+p)\Pr\{d > j\}$$

is **non-decreasing** in Q.



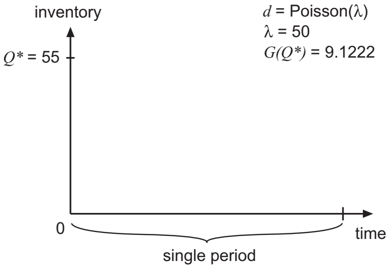
$$Q^* = \min\{Q \in \mathbb{N}_0 : \Delta G(Q) \geq 0\}.$$

Solution method: example

Demand follows a Poisson distribution $Poisson(\lambda)$, with demand rate $\lambda = 50$.

Holding cost $h = 1$, penalty cost $p = 3$.

The optimal order quantity Q^* is equal to 55 and provides a cost equal to 9.1222.



Assumptions on demand distribution

What happens if we consider different assumptions on demand distribution?

Khouja (2000), among other extensions, surveyed those dealing with different states of information about demand.

Demand					
Moments	Known	X	X		
	Unknown			X	X
Distribution	Known	X			X
	Unknown		X	X	
Observations				X	X

Assumptions on demand distribution

Known moments & distribution

What happens if we consider different assumptions on demand distribution?

Khouja (2000), among other extensions, surveyed those dealing with different states of information about demand.

Demand					
Moments	Known	X	X		
	Unknown			X	X
Distribution	Known	X			X
	Unknown		X	X	
Observations				X	X
		↑			

Assumptions on demand distribution

Known moments & unknown distribution

What happens if we consider different assumptions on demand distribution?

Khouja (2000), among other extensions, surveyed those dealing with different states of information about demand.

Demand					
Moments	Known	X	X		
	Unknown			X	X
Distribution	Known	X			X
	Unknown		X	X	
Observations				X	X
			↑		

Assumptions on demand distribution

Known moments & unknown distribution

All works below assume that demand distribution is not known, i.e. *distribution free* setting.

Authors	**Methodology**
Scarf et al. (1958)	"maximin approach," i.e. maximise the worst-case profit
Gallego & Moon (1993)	four extensions to Scarf et al. (1958)
Moon & Choi (1995)	extends Scarf et al. (1958) to account for balking: customers balk when inventory level is low
Perakis & Roels (2008)	"minimax regret," i.e. minimises its maximum cost discrepancy from the optimal decision.

see Notzon (1970); Gallego et al. (2001); Bertsimas & Thiele (2006); Bienstock & Özbay (2008); Ahmed et al. (2007); See & Sim (2010) for multi-period inventory models.

Assumptions on demand distribution

Unknown moments & unknown distribution

What happens if we consider different assumptions on demand distribution?

Khouja (2000), among other extensions, surveyed those dealing with different states of information about demand.

Demand					
Moments	Known	X	X		
	Unknown			X	X
Distribution	Known	X			X
	Unknown		X	X	
Observations				X	X
				↑	

Assumptions on demand distribution

Unknown moments & unknown distribution

All works below operate without any access to and assumptions on the true demand distributions, i.e. *non-parametric* setting.

Authors	**Methodology**
Hayes, 1971	order statistics
Lordahl & Bookbinder, 1994	order statistics
Bookbinder & Lordahl, 1989	bootstrapping
Fricker & Goodhart, 2000	bootstrapping
Levi et al. (2007)	determine bounds for the number of samples needed to guarantee an arbitrary approximation of the optimal policy
Huh et al. (2009)	adaptive inventory policy that deal with censored observations

Assumptions on demand distribution

Unknown moments & known distribution

What happens if we consider different assumptions on demand distribution?

Khouja (2000), among other extensions, surveyed those dealing with different states of information about demand.

Demand					
Moments	Known	X	X		
	Unknown			X	X
Distribution	Known	X			X
	Unknown		X	X	
Observations				X	X
					↑

Assumptions on demand distribution

Unknown moments & known distribution

According to Berk et al. (2007) there are two general approaches for dealing with this setting: the **Bayesian** and the **frequentist**.

According to Kevork (2010) another distinction can be made between approaches assuming that demand is **fully observed** and approaches assuming that demand may be **censored**.

Assumptions on demand distribution

Unknown moments & known distribution

Bayesian approaches in the literature:

Fully observed demand	**Censored demand**
Scarf (1959, 1960)	Lariviere & Porteus (1999)
Iglehart (1964)	Ding & Puterman (1998)
Azoury (1985)	Berk et al. (2007)
Lovejoy (1990)	Chen (2010)
Bradford & Sugrue (1990)	Lu et al. (2008)
Hill (1997)	Mersereau (2012)
Eppen & Iyer (1997)	
Hill (1999)	
Lee (2008)	
Bensoussan et al. (2009)	

Assumptions on demand distribution

Unknown moments & known distribution

Frequentist approaches in the literature:

Authors	**Methodology**
Nahmias (1994)	stock level is given
Agrawal & Smith (1996)	stock level is given
Liyanage & Shanthikumar (2005)	"operational statistics:" optimal order quantity directly estimated from the data
Kevork (2010)	exploits the sampling distribution of the demand parameters to study the variability of the estimates for the optimal order quantity and associated expected total profit.
Akcay et al. (2011)	ETOC: expected one-period cost associated with operating under an estimated inventory policy
Klabjan et al. (2013)	integrate distribution fitting and robust optimisation

Assumptions on demand distribution

Unknown moments & known distribution

Assume now that the **demand distribution** is known, but one or more **distribution parameters** are unknown.

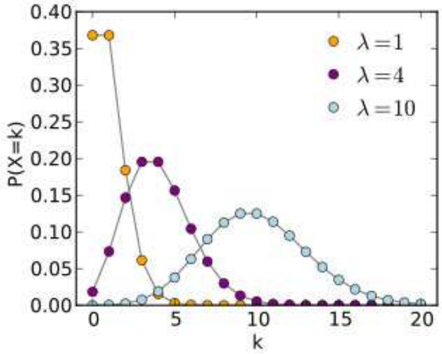
The decision maker has access to a set of M **past realizations of the demand**.

From these she has to estimate the **optimal order quantity** (or quantities) and the **associated cost.**

Assumptions on demand distribution

Unknown moments & known distribution

Poisson demand, probability mass function:



λ has to be **estimated** from past realizations.

A frequentist approach

Point estimates of the parameter(s)

Point estimates of the unknown parameters may be obtained from the available samples by using:

- **maximum likelihood estimators**, or
- the **method of moments.**

Point estimates for the parameters are then used **in place** of the unknown demand distribution parameters to compute:

- the estimated **optimal order quantity** $\widehat{Q}^*$, and
- the associated estimated **expected total cost** $G(\widehat{Q}^*)$.

A frequentist approach

Point estimates: example

M observed **past demand data** $d_1, \ldots, d_M$.

Demand follows a **Poisson distribution** $Poisson(\lambda)$, with demand rate λ.

We estimate λ using the **maximum likelihood estimator** (sample mean):

$$\widehat{\lambda} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i.$$

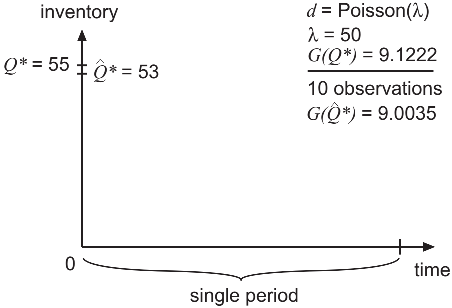
The decision maker employs the distribution $Poisson(\widehat{\lambda})$ **in place** of the actual unknown demand distribution.

A frequentist approach

Point estimates: example

Holding cost: $h = 1$; penalty cost: $p = 3$; observed past demand data: $\{51, 54, 50, 45, 52, 39, 52, 54, 50, 40\}$.

$\widehat{\lambda} = 48.7$, $\widehat{Q}^* = 53$ and $G(\widehat{Q}^*) = 9.0035$.



Bayesian approach

The bayesian approach **infers** the distribution of parameter λ given some past observations d by applying **Bayes' theorem** as follows

$$p(\lambda|d) = \frac{p(d|\lambda)p(\lambda)}{\int p(d|\lambda)p(\lambda)\mathrm{d}\lambda}$$

where

$p(\lambda)$ is the **prior distribution** of λ, and

$p(\lambda|d)$ is the **posterior distribution** of λ given the observed data d.

Bayesian approach

The **prior distribution** describes **an estimate** of the likely values that the parameter λ might take, without taking the data into account. It is based on **subjective assessment** and/or **collateral data**.

A number of methods for constructing "**non-informative priors**" have been proposed (i.e. maximum entropy). These are meant to reflect the fact that the decision maker **ignores** of the prior distribution.

If prior and posterior distributions are in the same family, then they are called **conjugate distributions**.

Bayesian approach
[Hill, 1997]

Hill [EJOR, 1997] proposes a **bayesian approach to the Newsvendor problem**.

He considers a number of distributions (Binomial, Poisson and Exponential) and **derives posterior distributions for the demand** from a set of given data.

He adopts **uninformative priors** to express an initial state of **complete ignorance** of the likely values that the parameter might take.

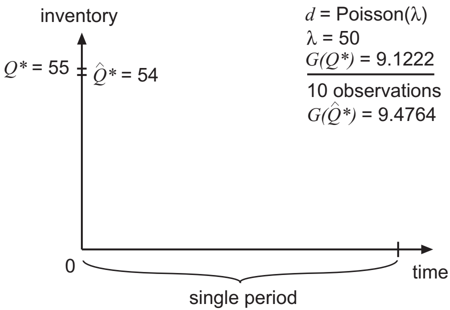
By using the posterior distribution he obtains an **estimated optimal order quantity** and the respective **estimated expected total cost**.

Bayesian approach

[Hill, 1997] example

Holding cost: $h = 1$; penalty cost: $p = 3$;
observed past demand data:
$\{51, 54, 50, 45, 52, 39, 52, 54, 50, 40\}$.

$\widehat{Q}^* = 54$ and $G(\widehat{Q}^*) = 9.4764$.



Drawbacks of existing approaches

Only provide point estimates of the order quantity and of the expected total cost.

Do not quantify the uncertainty associated with this estimate.

- How do we distinguish a case in which we only have 10 past observations vs a case with 1000 past observations?

The bayesian approach produces results that, for small samples, are “**biased**” by the selection of the prior; further drawbacks are outlined in

J. Neyman. Outline of a theory of statistical estimation based on the classical theory of probability. Philosophical Transactions of the Royal Society of London, 236:333—380, 1937

An alternative approach

We propose a solution method based on **confidence interval analysis** [Neyman, 1937].

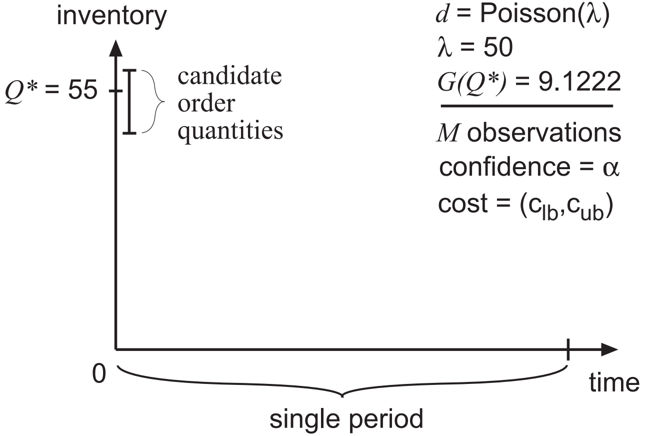
Observation

Since we operate under partial information, it may not be possible to uniquely determine “the” optimal order quantity and the associated exact cost.

We argue that a possible approach consists in **determining a range** of “candidate” optimal order quantities and **upper and lower** bounds for the **cost** associated with these quantities.

This range will contain the real optimum according to a **prescribed confidence probability** α.

An alternative approach



d = Poisson(λ)
$\lambda = 50$
$G(Q^*) = 9.1222$

M observations
confidence = α
cost = (c_{lb}, c_{ub})

Confidence interval for λ

Consider a set of M random variates d_i drawn from a random demand d that is distributed according to a Poisson law with unknown parameter λ.

We construct a **confidence interval** for the unknown demand rate λ as follows

$$\lambda_{lb} = \min\{\lambda | \Pr\{Poisson(M\lambda) \geq \bar{d}\} \geq (1-\alpha)/2\},$$
$$\lambda_{ub} = \max\{\lambda | \Pr\{Poisson(M\lambda) \leq \bar{d}\} \geq (1-\alpha)/2\},$$

where $\bar{d} = \sum_{i=0}^{M} d_i$.

A **closed form expression** for this interval has been proposed by Garwood [1936] based on the chi-square distribution.

Confidence interval for λ: example

Consider the set of 10 random variates

$$\{51, 54, 50, 45, 52, 39, 52, 54, 50, 40\},$$

and $\alpha = 0.9$.

The confidence interval for the unknown demand rate λ is

$$(\lambda_{lb}, \lambda_{ub}) = (45.1279, 52.4896),$$

Note that, by chance, this interval covers the actual demand rate $\lambda = 50$ used to generate the sample.

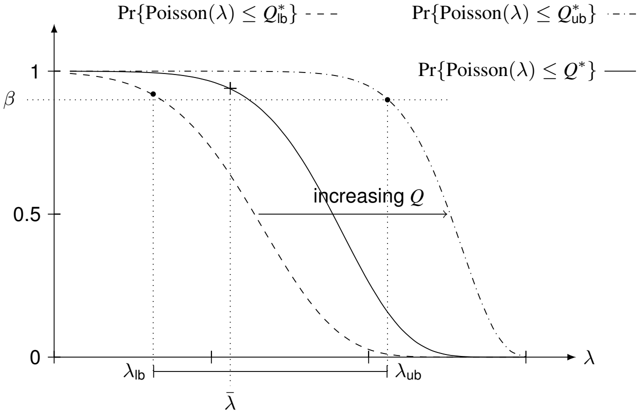
Candidate order quantities

Let Q^*_{lb} be the **optimal order quantity** for the Newsvendor problem under a $Poisson(\lambda_{lb})$ demand.

Let Q^*_{ub} be the **optimal order quantity** for the Newsvendor problem under a $Poisson(\lambda_{ub})$ demand.

Since $\Delta G(Q)$ is **non-decreasing** in Q, according to the available information, **with confidence probability** α, the optimal order quantity Q^* is a **member** of the set $\{Q^*_{lb}, \ldots, Q^*_{ub}\}$.

Candidate order quantities



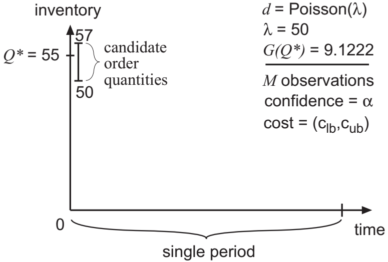
Candidate order quantities: example

Consider the set of 10 random variates

$$\{51, 54, 50, 45, 52, 39, 52, 54, 50, 40\},$$

and $\alpha = 0.9$.

The candidate order quantities are



Confidence interval for the expected total cost

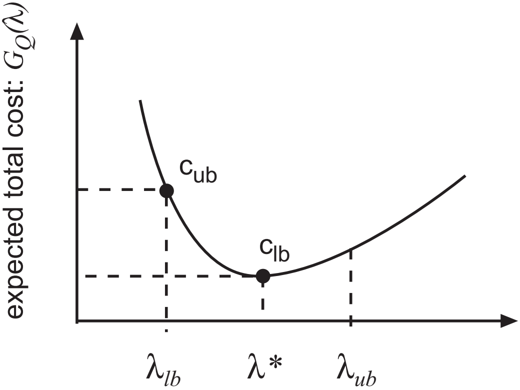
For a **given order quantity** Q we can prove that

$$
\begin{aligned}
G_Q(\lambda) = \quad & h\sum_{i=0}^{Q} \Pr\{Poisson(\lambda)=i\}(Q-i)+\\
& p\sum_{i=Q}^{\infty} \Pr\{Poisson(\lambda)=i\}(i-Q),
\end{aligned}
$$

is **convex** in λ.

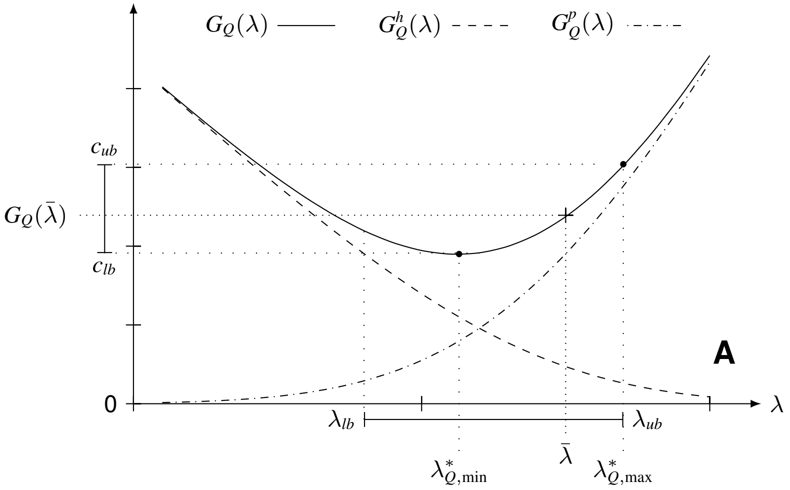
Upper (c_{ub}) and **lower** (c_{lb}) **bounds** for the cost associated with a solution that sets the order quantity to **a value in the set** $\{Q^*_{lb},\dots,Q^*_{ub}\}$ can be easily obtained by using **convex optimization** approaches to find the λ^* that maximizes or minimizes this function over $(\lambda_{lb},\lambda_{ub})$.

Confidence interval for the expected total cost

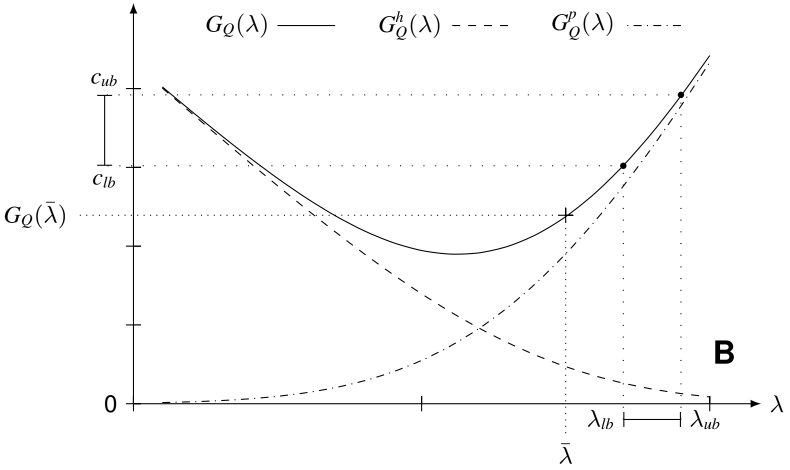


for $Q \in \{Q^*_{lb}, \ldots, Q^*_{ub}\}$.

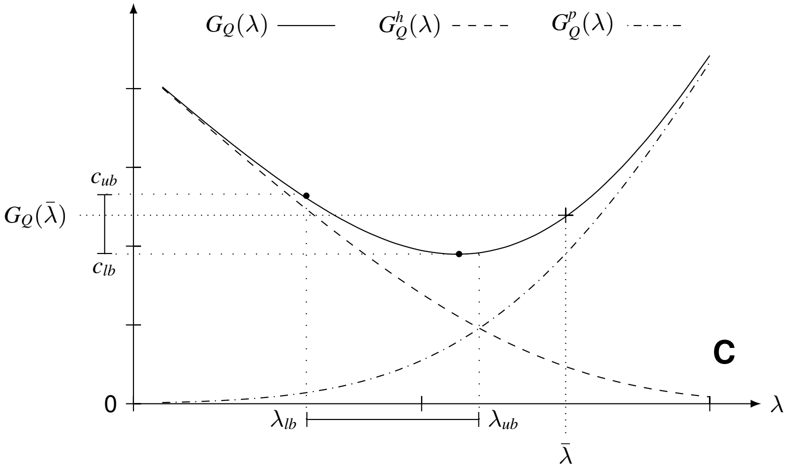
Confidence interval for the expected total cost



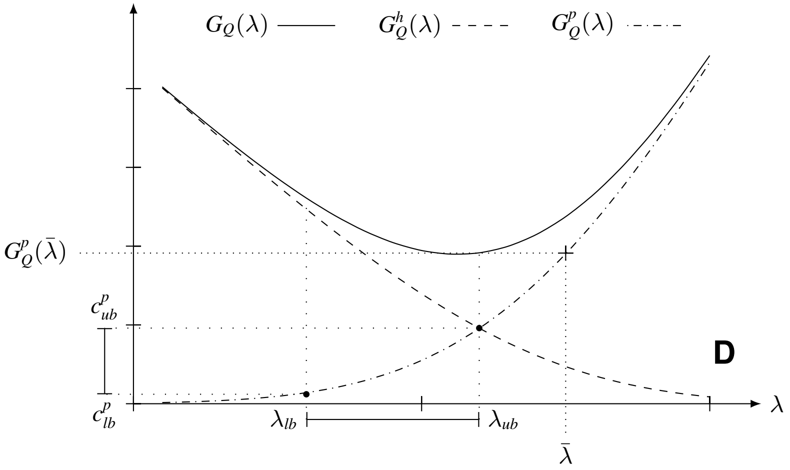
Confidence interval for the expected total cost



Confidence interval for the expected total cost



Confidence interval for the expected total cost



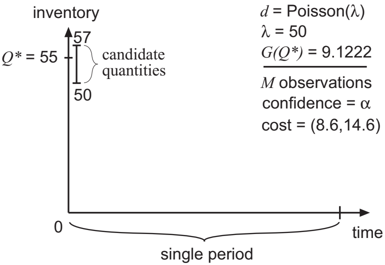
Expected total cost: example

Consider the set of 10 random variates

$$\{51, 54, 50, 45, 52, 39, 52, 54, 50, 40\},$$

and $\alpha = 0.9$.

The upper and lower bound for the expected total cost are



Expected total cost: example

Assume we decide to order 53 items, according to what a **MLE approach** suggests.

As we have seen, **MLE estimates** an expected total cost of 9.0035 (note that the real cost we would face is 9.3693).

If we compute $c_{lb} = 8.9463$ and $c_{ub} = 11.0800$, then we know that with $\alpha = 0.9$ confidence this interval **covers the real cost** we are going to face by ordering 53 units.

Similarly, the Bayesian approach **only prescribes** $\widehat{Q}^* = 54$ and estimates $G(\widehat{Q}^*) = 9.4764$ (real cost is 9.1530), while we know that $c_{lb} = 9.0334$ and $c_{ub} = 10.3374$.

Lost sales

Consider the case in which **unobserved lost sales** occurred and the M observed past demand data, $d_1, \ldots, d_M$, **only reflect** the number of customers that purchased an item **when the inventory was positive.**

The analysis discussed above **can still be applied** provided that the confidence interval for the unknown parameter λ of the $Poisson(\lambda)$ demand is computed as

$$\lambda_{lb} = \min\{\lambda | \Pr\{Poisson(\widehat{M}\lambda) \geq \bar{d}\} \geq (1-\alpha)/2\},$$
$$\lambda_{ub} = \max\{\lambda | \Pr\{Poisson(\widehat{M}\lambda) \leq \bar{d}\} \geq (1-\alpha)/2\}.$$

where $\widehat{M} = \sum_{j=1}^{M} T_j$, and $T_j \in (0,1)$ denotes **the fraction of time** in day j — for which a demand sample d_j is available — during which the **inventory was positive.**

Binomial demand

N customer enter the shop on a given day, the **unknown purchase probability** of the Binomial demand is $q \in (0, 1)$.

The analysis is **similar** to that developed for a Poisson demand.

Also in this case we prove that $G_Q(q)$ is **convex** in q.

Lost sales can be **easily incorporated** in the analysis.

Exponential demand

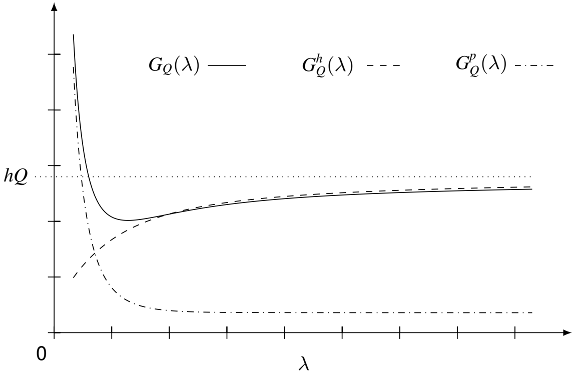
The interval of candidate order quantities can be **easily identified.**

The analysis on the expected total cost is **complicated** by the fact that $G_Q(\lambda)$ is **not convex.**

Extension to include lost sales is **difficult.**

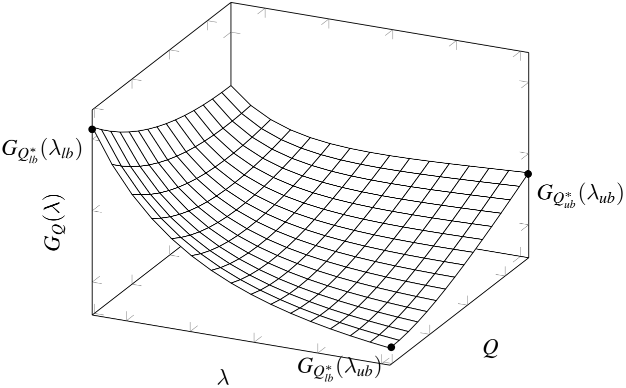
Exponential demand

A **number of properties** of $G_Q(\lambda)$ can be exploited to find **upper and lower bounds for the expected total cost.**



Exponential demand

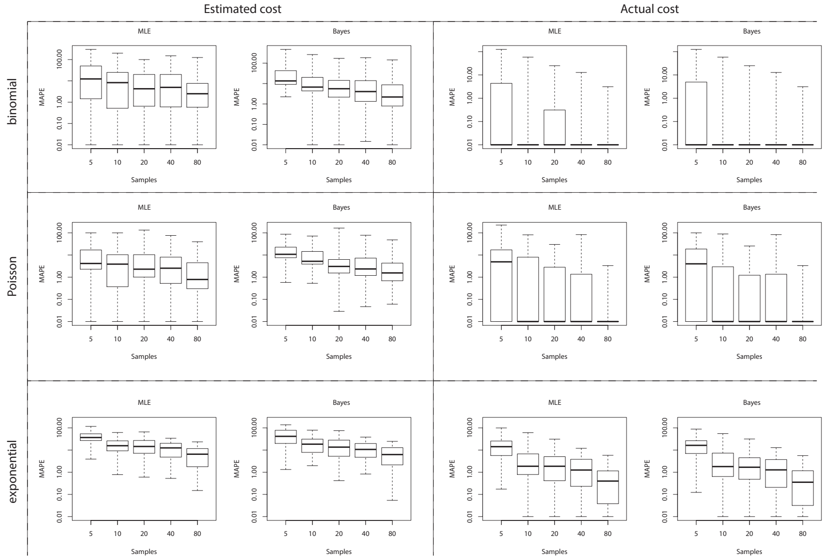
A **number of properties** of $G_Q(\lambda)$ can be exploited to find **upper and lower bounds for the expected total cost.**



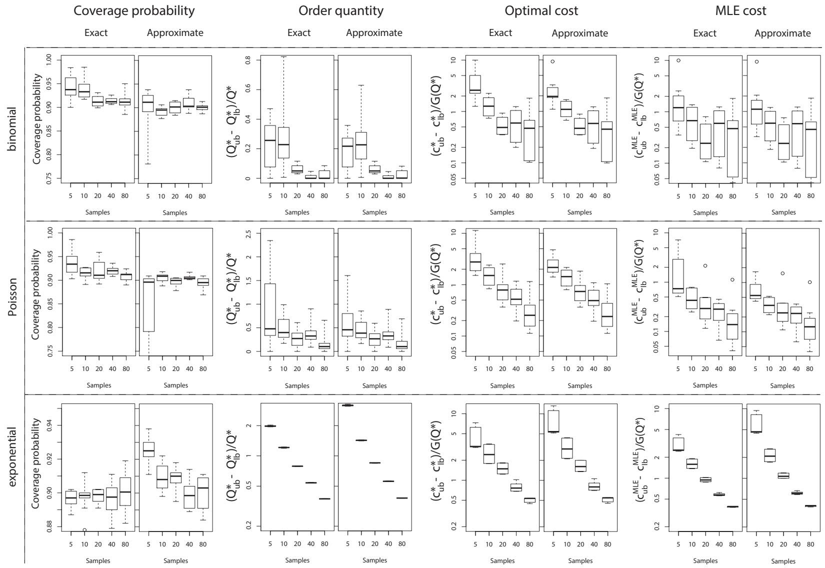
Experiment setup

Parameter	Values
h	1
p	$2, 4, 8, 16$
M	$5, 10, 20, 40, 80$
α	0.9
N	$1, 2, 4, 8, 16, 32, 64$
p	$0.5, 0.75, 0.95$
λ	$0.125, 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64$

Comparison with MLE and Bayesian approaches



Confidence-based optimisation results



Discussion

We presented a **confidence-based optimization** strategy to the Newsboy problem with **unknown demand distribution parameter(s)**.

We applied our approach to three **maximum entropy** probability distributions of the **exponential family.**

We showed the **advantages of our approach** over two existing strategies in the literature.

For two demand distributions we extended the analysis to include **lost sales.**

Future works

Consider **other probability distributions** (e.g. Normal, LogNormal, Multinomial etc.).

Further **develop the analysis on lost sales** for the Exponential distribution.

Extend the methodology to a **non-parametric** setting.

Apply confidence-based optimization to **other stochastic optimization problems**.

Questions