Estimation of Discrete Cumulative Distributions by Resampling

John E Boylan & M Zied Babai

Lancaster University, Kedge Business School

International Symposium on Inventories

Budapest, 21 August 2018

Marketing Analytics & Forecasting | Lancaster University Management School

KEDGE BUSINESS SCHOOL
CREATED BY BEM & EUROMED MANAGEMENT

Agenda

- Introduction
- Motivation of this Research
- Properties of Parametric Methods
- Properties of Overlapping and Non-Overlapping Blocks
- Resampling Approaches: Bias Properties
- Numerical Investigation
- Potential Bias Remedies
- Conclusions

Introduction: Parametric Approaches

Parametric distributions often recommended for inventory models

- Normal, Gamma
- Bernoulli (and Compound Bernoulli)
- Poisson (and Compound Poisson, including Negative Binomial)

Empirical evidence

- Good support for Bernoulli / Poisson models of demand incidence for intermittent demand items.
- Quite good support for Negative Binomial Distribution
- BUT: many demands not well modelled by ***any*** parametric distribution.

CMAF

Introduction: Non-Parametric Approaches

Example (History = 10 periods)

P1	P2	P3	P4	P5	P6	P7	P8	P9	P10

Suppose LT = 3 and Block-Size (m) = 3 [Natural application in inventory context]

1. Non-Overlapping Blocks

Block 1 = {P2, P3, P4}, Block 2 = {P5, P6, P7}, Block 3 = {P8, P9, P10}

2. Overlapping Blocks

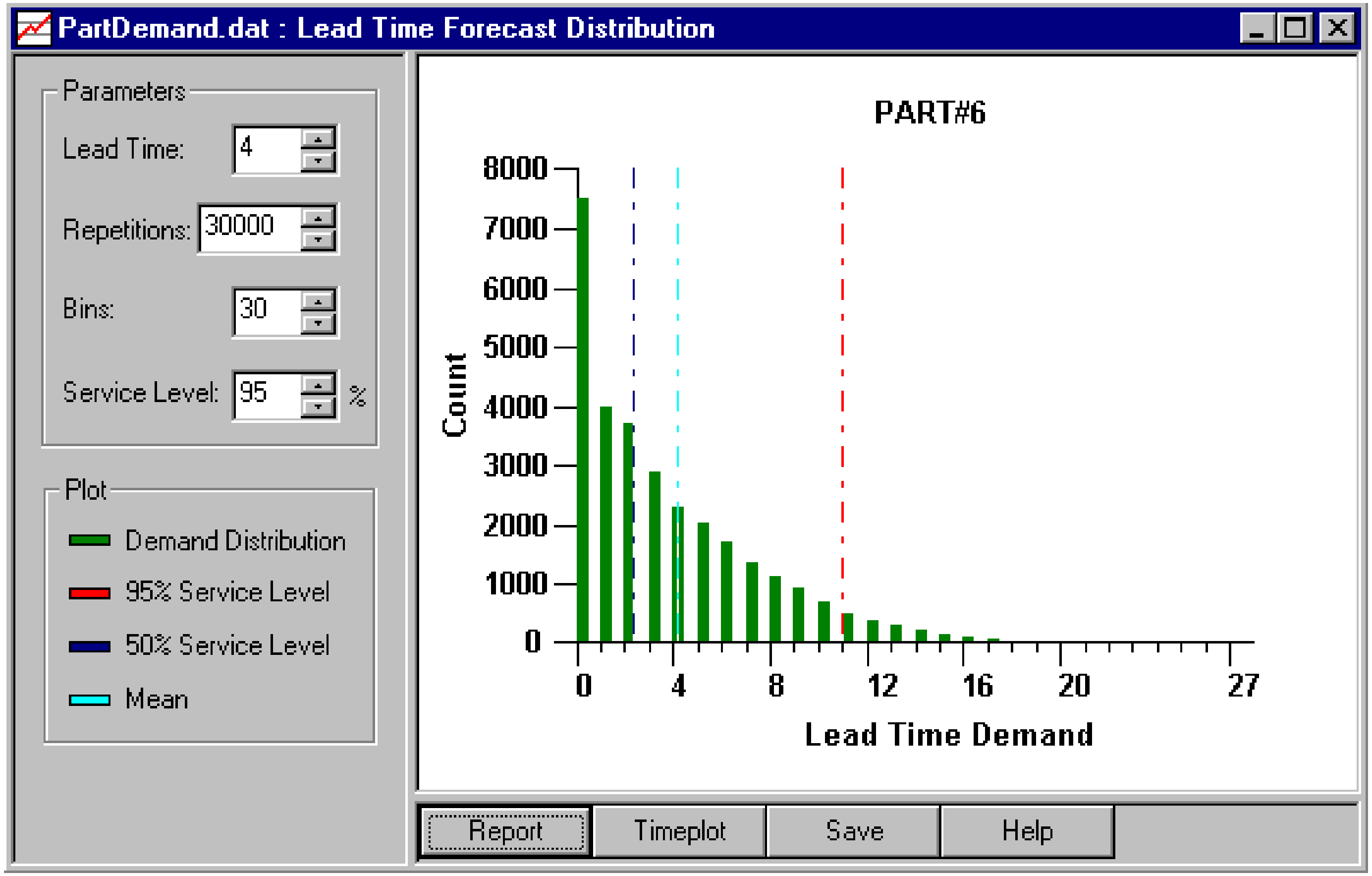
Block 1 = {P1, P2, P3}, Block 2 = {P2, P3, P4} , ... , Block 8 = {P8, P9, P10}

3. Resampling (Bootstrapping)

Randomly drawn from all combinations (with replacement)
eg: Block 1 = {P7, P2, P9}, Block 2 = {P2, P3, P3}, ...

CMAF

Introduction: Bootstrapping Software



- Smart Software incorporates Markov Chain switching (between ‘demand ‘ and ‘non-demand’ states), and “jittering”.
- These extensions not addressed in this research so far.

Motivation of this Research

Resampling approaches for intermittent demand have some empirical evidence in their support (Willemain, 2004).

Theoretical properties not investigated in an inventory context in which cumulative distributions need to be estimated.

This study aims to investigate the bias and variance: :

- With replacement
- Without replacement

and to suggest ways of de-biasing With Replacement.

Properties of Non-Overlapping Blocks

Estimation Problem

- Suppose History Length = ***n*** and we are given an (integer-valued) quantity ***y*** and the population cumulative distribution function is $P_m(y)$ for total demand over ***m*** periods
- We produce an estimate, $\hat{P}_m(y)$, based on ***k=Int(n/m)*** Non-Overlapping Blocks
- Statistical properties are well established for stationary i.i.d. demand:

Unbiased Estimate

$$E[\hat{P}_m(y)] = P_m(y)$$

Variance of the Estimate

$$Var[\hat{P}_m(y)] = \frac{P_m(y)}{k} - \frac{P_m(y)^2}{k}$$

- Variance depends only on **number of blocks** (*k*) and **CSL** (*P_m(y)*)

CMAF

Properties of Overlapping Blocks
(Boylan and Babai, IJPE, 2016)

Number of Blocks

- This increases from n/m to $n-m+1$
  (eg from 12 to 34 for History Length=n=36 and Block Size=m=3).

Unbiased Estimate .

$$E[\hat{P}_m(y)] = P_m(y)$$

Variance of the Estimate

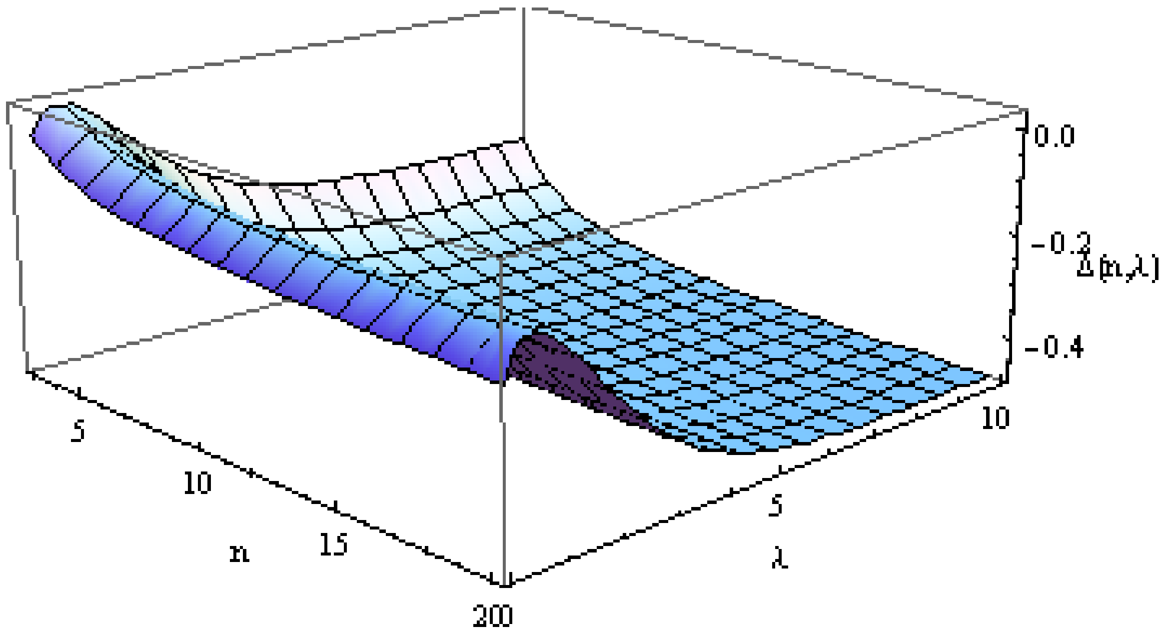
$$Var[\hat{P}_m(y)] = \frac{1}{k}P_m(y) + \left[\frac{(k-m)(k-m+1)}{k^2} - 1\right]P_m(y)^2$$

$$+\frac{2}{k^2}\sum_{s=1}^{m-1}(k-m+s)\overbrace{\sum_{Y_1=0}^{y}\cdots\sum_{Y_m=0}^{y-Y_1-\ldots-Y_{m-1}}\;\sum_{Y_{m+1}=0}^{Y-Y_{m-s+1}-\ldots Y_m}\cdots\sum_{Y_{2m-s}=0}^{y-Y_{m-s+1}-\ldots-Y_{2m-s-1}}\;\prod_{j=1}^{2m-s}p(Y_j)}^{\Theta_s(y)}$$

- s : number of overlapping observations between two blocks
- p : probability mass function (pmf) of the demand in one period
- Performance also depends on block size (*m*) and the pmf *(p)*

Variance Reduction using OB instead of NOB

- Poisson distributed demand with $\lambda \in [0,10]$; $n \in [3,20]$ and $m = 2$



- In almost all cases, the OB approach leads to a variance reduction.
- For very low values of λ (i.e. slow moving demand), the benefit from using OB instead of NOB is very low when *n* is low.
- There are very few values where NOB outperforms OB; these cases occur when both λ and *n* are very low.

Resampling with Replacement: Bias Properties

Biased Estimate

$$E[\hat{P}_m^{Boot(R)}(y)] \neq P_m(y)$$

Special Case (m=2)

$$Bias = [P_1(\lfloor y/2 \rfloor) - P_2(y)]$$

Special Case (m=3)

Λ

$$Bias = \frac{1}{n^2} P_1(\lfloor y/3 \rfloor) + 3\left(1 - \frac{1}{n}\right)\Lambda_{21}(y) - \left(\frac{3}{n} - \frac{2}{n^2}\right)P_3(y)$$

$$\Lambda_{21}(y) = \sum_{Y_1=0}^{\lfloor y/2 \rfloor} \sum_{Y_2=0}^{y-Y_1} p(Y_1)\,p(Y_2)$$

- Different factor (Λ) than in the Overlapping Blocks approach.
- New factor takes into account repeated selection of same time index.

CMAF

Resampling with Replacement: Variance Properties

$$Var(\hat{P}_m(y)) = \frac{1}{k} P_m(y) + \frac{1}{k^2} \sum_{|Si \cap S_j = l|} \sum_{l=1}^{m-1} E[I_{[0,y]}(D_i) I_{[0,y]}(D_j)] P_m(y)^2 + \frac{1}{k^2} \sum_{|S_i \cap Sj| = m} P_m(y)$$

- First term – selection of same indices in same sequence
- Second term – some indices different
- Third term – same indices in different sequence

Special Case (m=2)

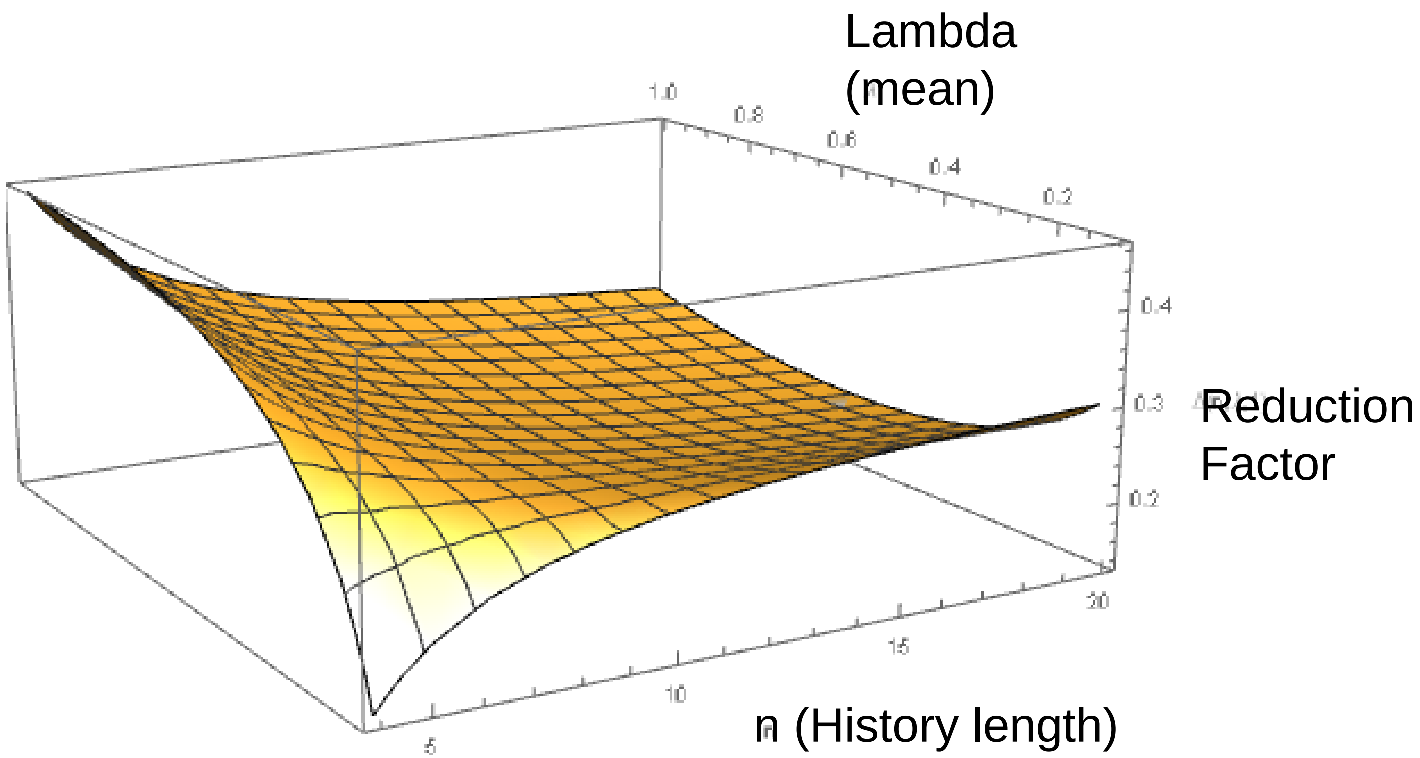
$$n^4 E(\hat{P}_2^{Boot}(y)^2) = n(n-1)[P_1(\lfloor y/2 \rfloor)^2 + 2(n-2)P_1(\lfloor y/2 \rfloor)P_2(y) + (n-2)(n-3)P_2(y)^2]$$
$$+ 4n(n-1)\Lambda_1(y) + 4n(n-1)(n-2)\Theta_1(y) + nP_1(\lfloor y/2 \rfloor) + 2(n^2 - n)P_2(y)$$

$$Var[\hat{P}_2^{Boot(R)}(y)] = E[P_2^{Boot(R)}(y)^2] - E[P_2^{Boot(R)}(y)]^2$$

- Special case for m=3 has been derived – very messy!

CMAF

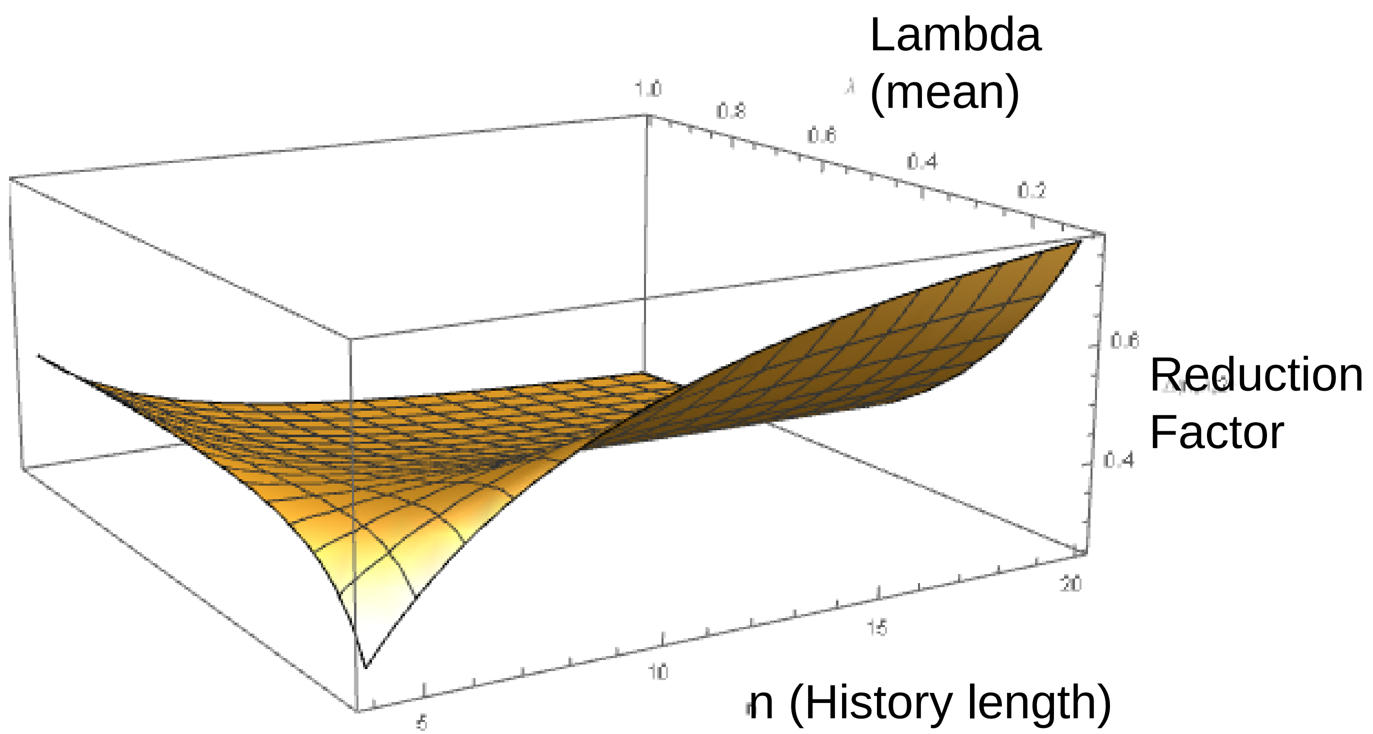
Resampling with Replacement: compared with OB for y=1 (Poisson demand)



- Variance always reduced by using Resampling with Replacement
- Strong reduction across most parameters

. .

Resampling with Replacement: compared with OB for y=2 (Poisson demand)

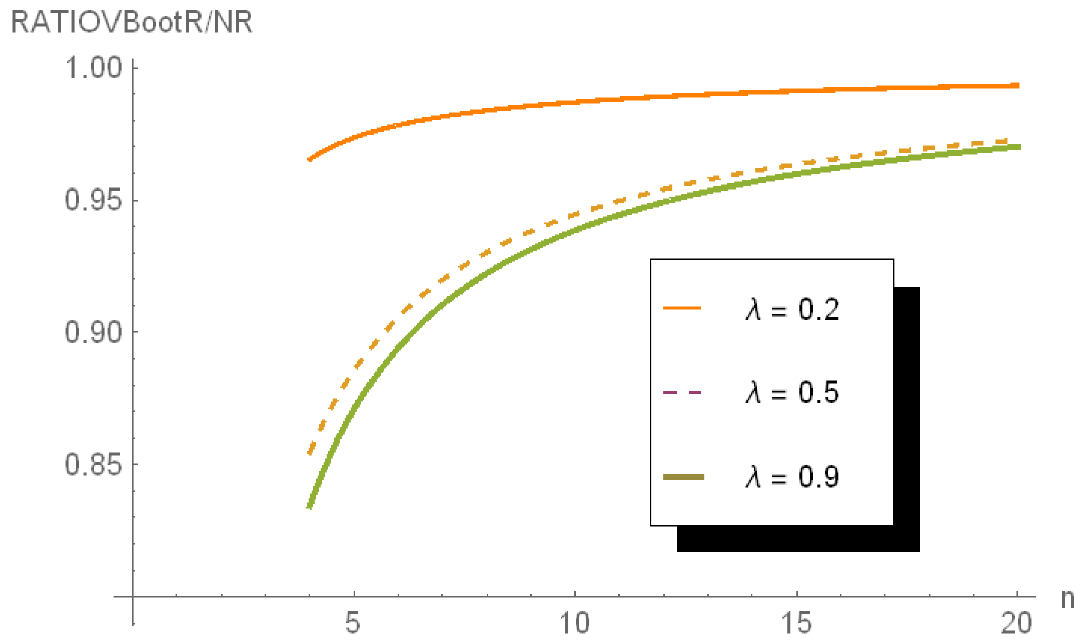


- Variance always reduced by using Resampling with Replacement
- Stronger reductions for y=2 compared with y=1
- Stronger effect of the length of demand histories

CMAF

Comparison of Resampling With and Without Replacement (y=1)

Poisson Demand



- With Replacement always gives lower variance for y=1
- This can be reversed for higher values of y and high intermittence
- For shorter histories, difference may become significant for less intermittent demand .

CMAF

Bias Issue

- With Replacement generally gives lower variance of CDF estimates than Without Replacement, except for very high Service Levels.

BUT

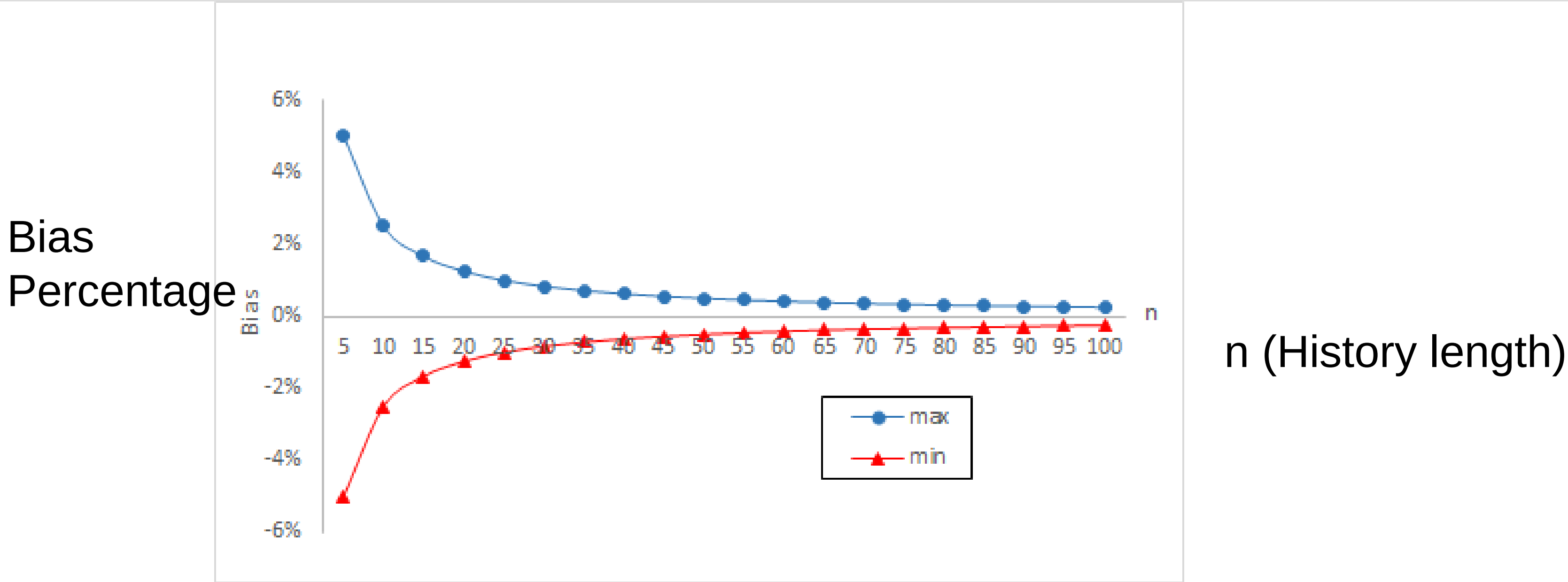
- With Replacement gives biased CDF estimates.

SO

- Can we quantify the bias?
- Can we identify when it occurs?
- Can we remove it?

CMAF

Maximum and Minimum Bias (m=2)



- **Allowing for any demand distribution**, we can treat p(0), p(1), p(2), … as decision variables, Then, it is non-linear optimisation problem to determine the maximum and minimum bias.
- Simple result for m=2, *and any value of y*, namely:
  - Max Bias = 1/4n
  - Min Bias = -1/4n

CMAF

Conditions for Greatest Positive Bias (m=2)

Can be derived from the KKT conditions for optimality.

$$\sum_{i=0}^{Int(y/2)} p(i) = \frac{1}{2}$$

For i=Int(y/2),…y:

$$p(i)\sum_{j=0}^{y-i} p(j) = 0$$

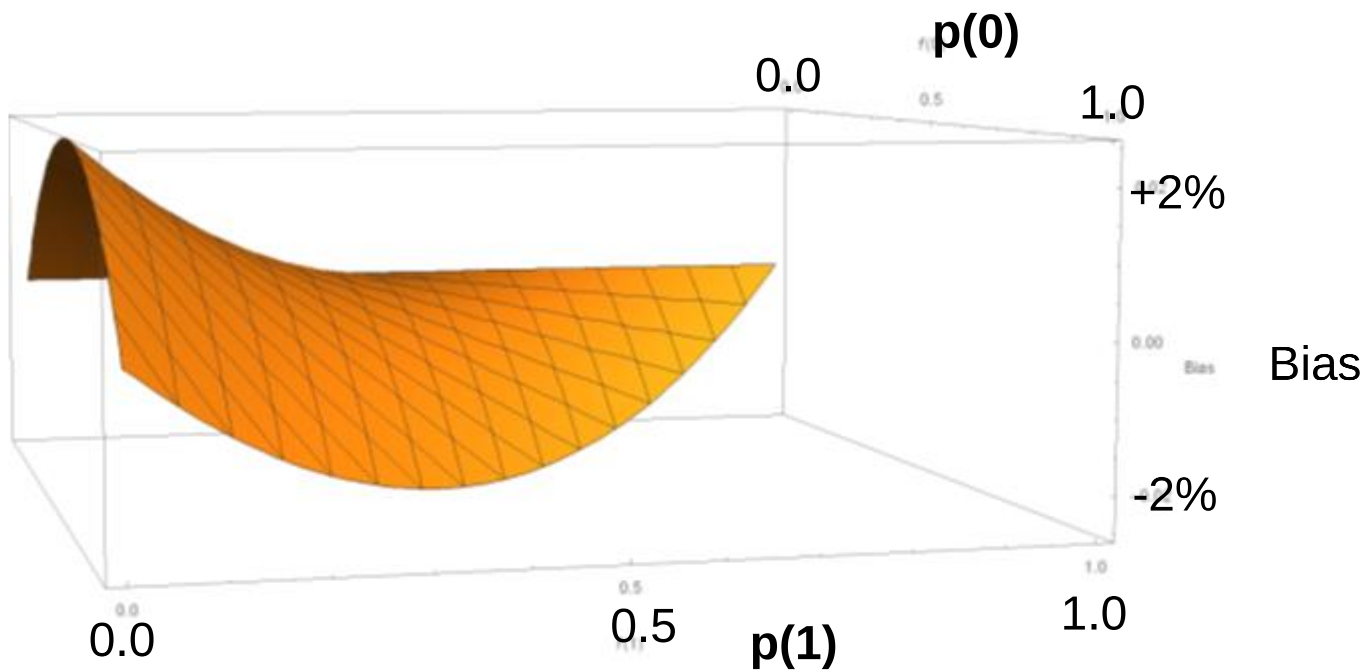
with any probabilities for i=y+1,… (satisfying a pmf).

Example

For y=1, maximum bias is for p(0)= 1/2, p(1) =0, and any other remaining probabilities summing to 1/2.

CMAF

Sensitivity of Bias to Distribution (m=2, y=1; n=10)



- Maximum bias at p(0)=0.5 and p(1)=0, as anticipated. This situation may arise with ‘clumped demand’.
- Can still have significant bias when p(1)>0 [p(1)=0 not so common in practice]

Conditions for Greatest Negative Bias (m=2)

Again, can be derived from the KKT conditions for optimality.

$$\sum_{i=0}^{Int(y/2)} p(i) = \frac{1}{2} \qquad \sum_{i=Int(y/2)+1}^{y} p(i) = \frac{1}{2}$$

For i=0,…,Int(y/2):

$$p(i)\sum_{j=0}^{y-i} p(j) = p(i)$$

For i=Int(y/2)+1,…,y:

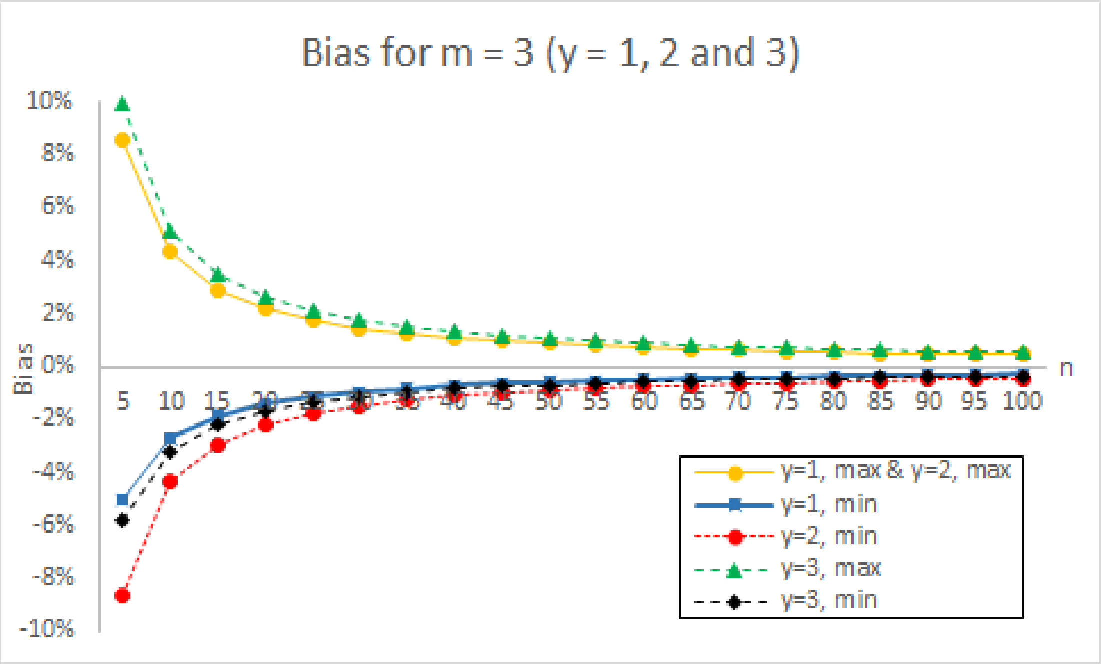
$$p(i)\sum_{j=0}^{y-i} p(j) = \frac{p(i)}{2}$$

Example

For y=1, minimum bias is for p(0)= 1/2, p(1) =1/2.

Maximum and Minimum Bias (m=3)

Bias Percentage

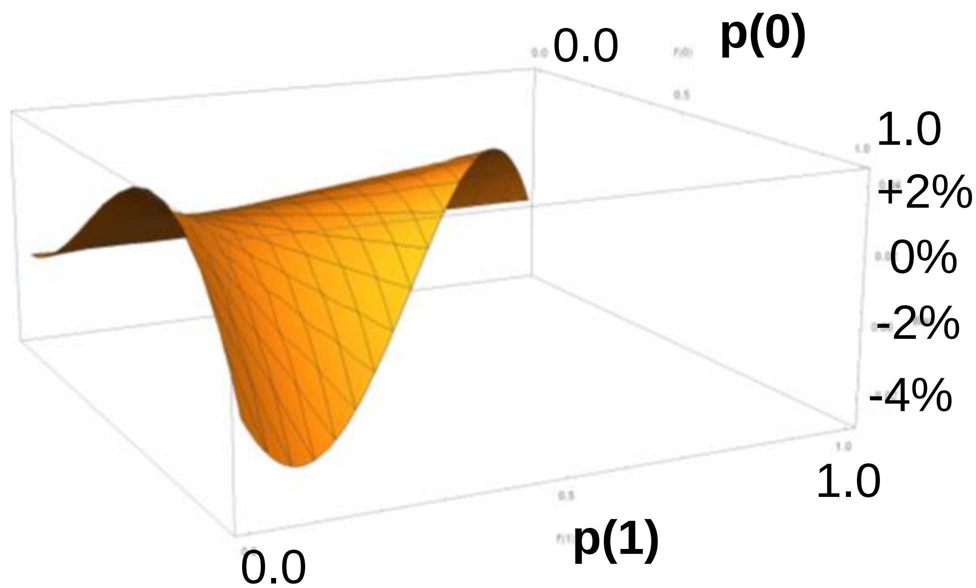


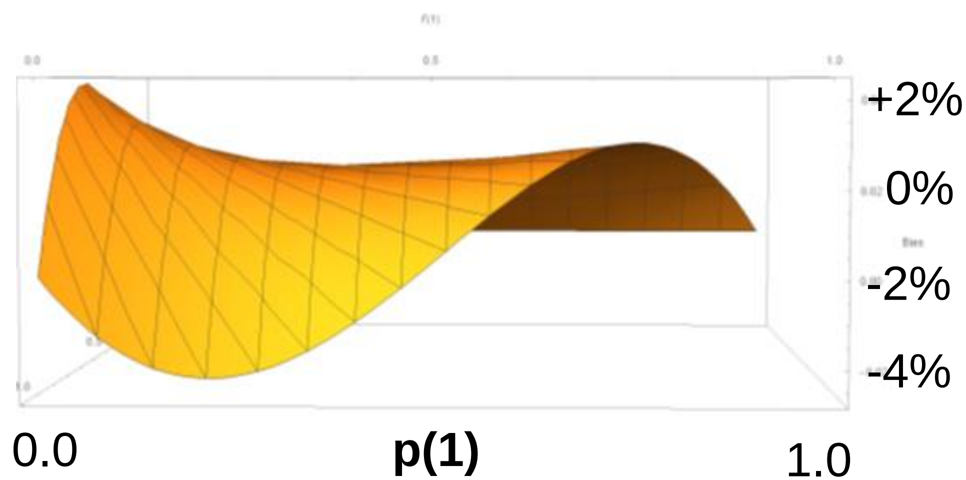
n (History length)

- Bias can be higher than for m=2.
- No longer symmetry in maximum and minimum bias
- Results no longer general for any value of y
- KKT conditions become intractable, except for simplest case (y=1)

CMAF

Sensitivity of Bias to Distribution (m=3, y=1; n=10)

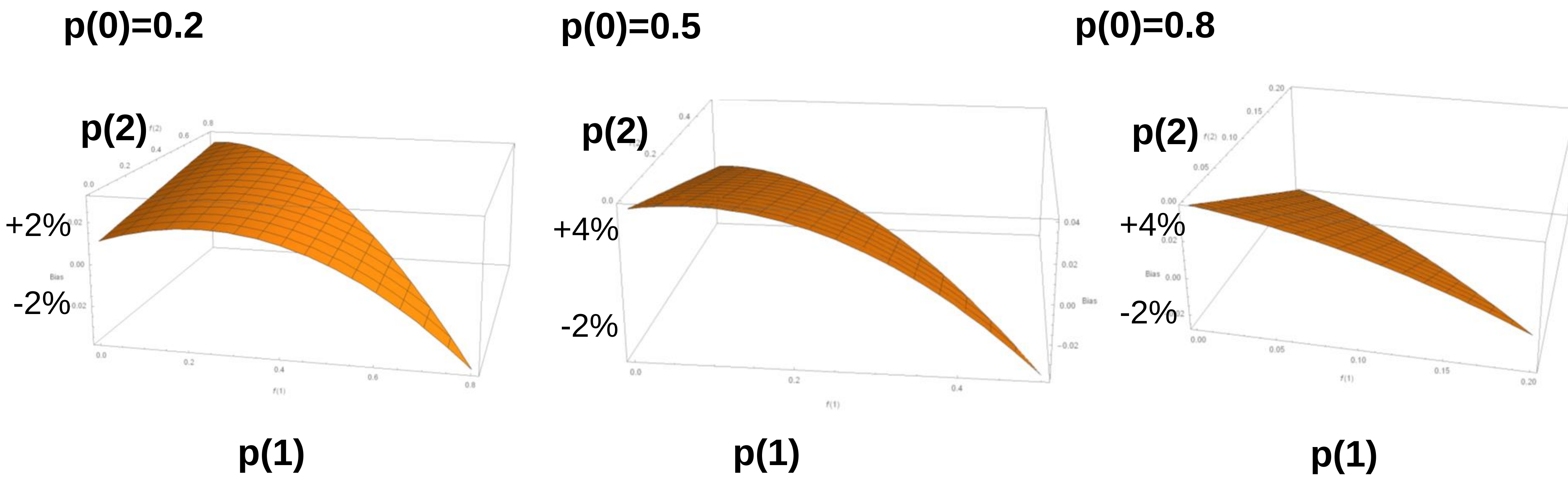




- Maximum Bias occurs at p(0)=0.79 and p(1)=0.
- Minimum Bias occurs at p(0)=0.79 and p(1)=0.21.
- Graphs show bias highly sensitive to specification of p(0) and p(1).

CMAF

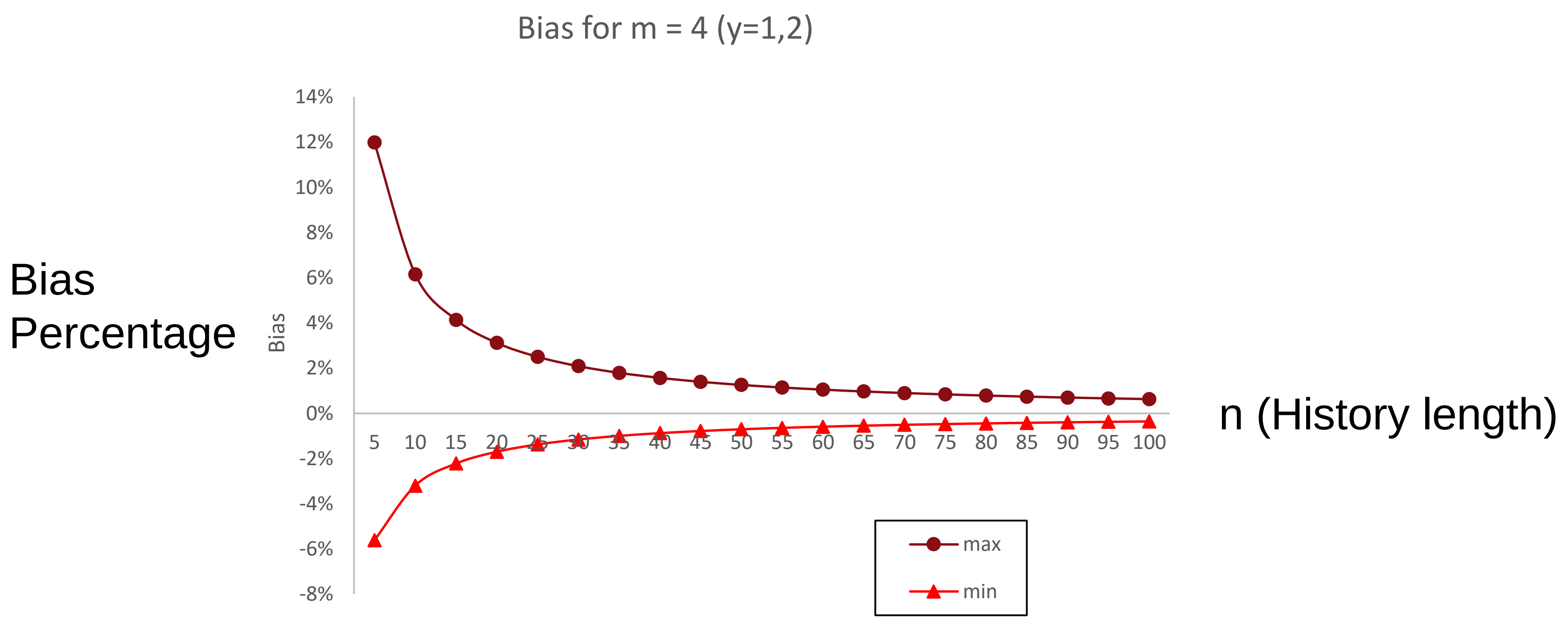
Sensitivity of Bias to Distribution (m=3, y=2; n=10)



- Admissible regions for p(1) and p(2) shrink as the degree of intermittence, p(0), increases.
- Lower maximum biases for p(0)=0.2 (mildly intermittent) but significant biases for p(0)=0.5 and p(0)=0.8.

CMAF

Maximum and Minimum Bias (m=4)



- Bias even higher than for m=3.
- No longer symmetry in maximum and minimum bias
- KKT conditions remain intractable.
- Have shown, mathematically, asymptotic bias is zero, *for any m*.

CMAF

Bias Correction

Special Case (m=2)

$$P_2^{Corr}(y) = \frac{n}{n-1}\hat{P}_2^{Boot}(y) - \frac{1}{n-1}\hat{P}_1(Int(y/2))$$

[$\hat{P}_1(Int(y/2))$ is estimated using the empirical CDF (unbiased for m=1)]

This removes the bias from the original CDF estimate.

Higher Values of m

Formulae become (much) more complex

Alternative is to construct CDF estimate from all combinations of values for each of the steps-ahead which satisfy the CDF constraint, continuing to assume distribution is iid.

CMAF

Conclusions

- Bias and variance properties of Resampling (with and without Replacement) have been derived for CDF estimates.
- Resampling with Replacement is generally biased in its estimate of the CDF.
- Bias stronger for shorter demand histories (length n).
- For m=2, Max Bias = 1/4n, Min Bias = -1/4n.
- For m =3 and m=4, bias can be even more significant.
- Bias correction formula available for m=2
- For higher values of m, can construct the CDF from the probabilities of all viable combinations.

CMAF

Thank you for your attention!

Q&A?

John Boylan

Lancaster University

j.boylan@Lancaster.ac.uk

Marketing Analytics & Forecasting

Lancaster University Management School