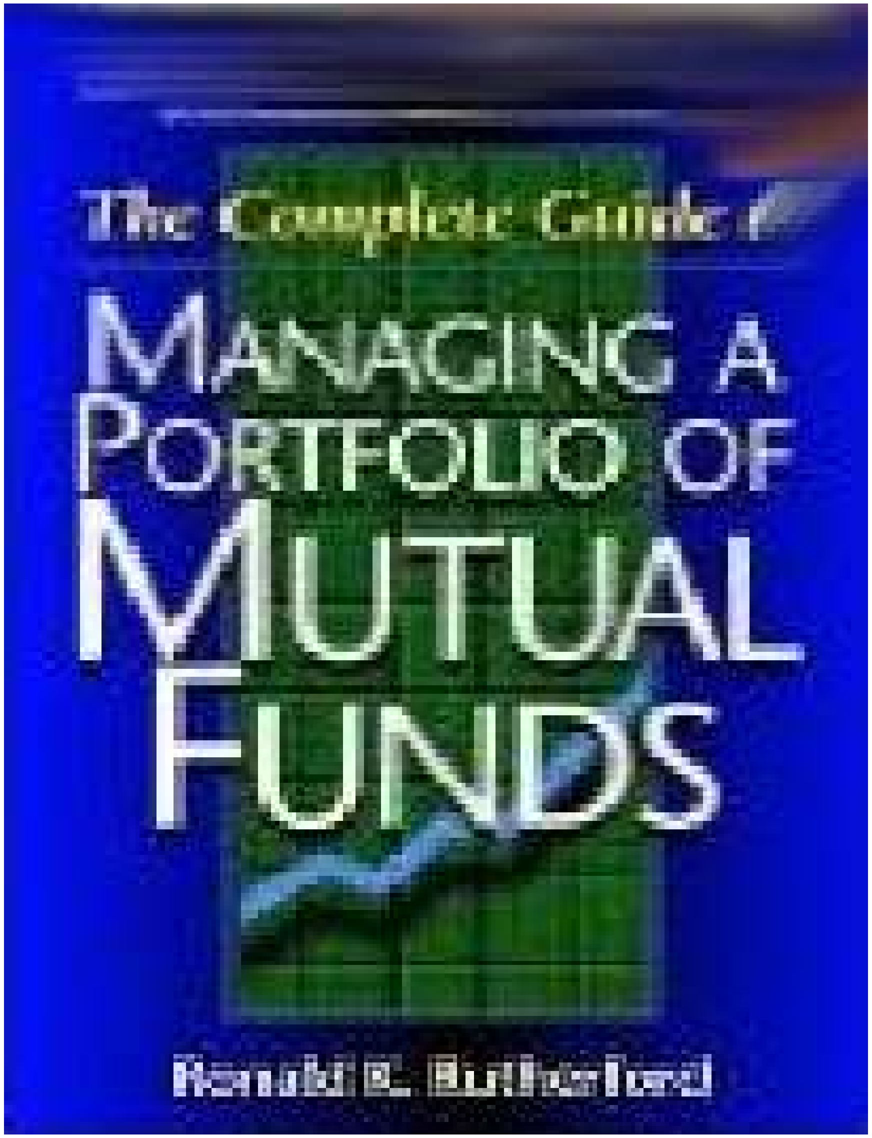
Lecture outline

- Applications of the newsvendor problem
- The newsvendor problem
- Estimating the distribution and censored demands
- The newsvendor problem and risk
- The newsvendor problem with an unknown distribution

The newsvendor problem

- How large should your cash reserve be for redemptions?
  - » Too big, and you have money that is underperforming.
  - » Too small, and you may have to dump stocks at a reduced price.



The newsvendor problem

- Water reservoirs can be used to balance the variability of energy from wind turbines. How much water should we store?
  - » Too much, and we are storing energy now when we could be using it more productively (you lose energy when you store it).
  - » Too little, and we may run out of water, which exposes us to the volatility of wind.

The newsvendor problem

- Examples (physical):
  - » How many jets should you order for your business jet fleet?
    - Too many: have to pay for them when they are not being used.
    - Too few: have to charter flights.
  - » How fast should you drive?
    - Too fast: risk of a ticket
    - Too slow: takes a while to get there.
  - » What should you sell your house for?
    - Too high: takes a long time to sell.
    - Too low: lose money on the deal.

The newsvendor problem

- Examples (physical)
  - » How many offers should J.P. Morgan make to ORFE students?
    - Assume J.P. Morgan wants to hire 10 ORFE majors. After the first round of interviews, JPM may make 10 offers but get only six, after which they make additional offers.
    - By the time they learn of the refusals, other students on their short list may have accepted offers from competing companies. JPM could make 14 offers and hope to get 10.
      - – If more than 10 accept, JPM has to create additional jobs.
      - – If less than 10 accept, JPM may be losing people to different companies.

The newsvendor problem

- Examples (financial)
  - » You are starting a new company.
    - You have to raise initial capital to get the company started. You never know exactly how much you will need before the company is profitable.
    - If you raise too little, you have to raise additional capital (which is more expensive) or risk the company. If you raise too much, then you have to pay for this.
  - » Pricing an IPO:
    - What if you price too high?
    - What if you price too low?

The newsvendor problem

- Examples (time)
  - » You have to allocate time to finish a project.
    - If you allocate too much time, then you have lost utilization of resources.
    - If you allocate too little time, then you face the penalties of not finishing the project on time.
  - » How many minutes should you commit to in your monthly cell phone plan?
    - Too few: you pay the per minute cost of overage.
    - Too much: you are paying for unused minutes.
  - » How early to wake up to go the 9am class?
    - 8am: likely to waste time waiting for class
    - 8:45am: likely to arrive late

The newsvendor problem

- The classic newsvendor problem (a.k.a. newsboy problem)
  - » You have to decide how many newspapers to put in the newsstand at the beginning of the day.
  - » At the end of the day you learn how many newspapers are left over. If you have any, you have to dispose of them (excess resource is assumed lost). Otherwise you may have lost demand.

The newsvendor problem

- Essential elements:
  - » Make a decision to allocate a resource of some form.
  - » Later learn the demand for the resource.
  - » Earn a net contribution from your decision.
  - » Game ends.
- Observations:
  - » The newsvendor problem is an elegant exercise in the study of sequencing information and decisions.
  - » It is imbedded in almost all resource allocation problems. It is imbedded in decisions you make every day. In fact, you solved this problem when you decided when to leave to arrive to this class!

The newsvendor problem

- Dimensions of a newsvendor problem
  - » Repeatability:
    - The one-shot version: no ability to learn from the result of an experiment.
    - The repeated newsvendor problem
  - » Time step (repeated version)
    - Short – e.g. daily
    - Long – update from one year to the next
  - » "Demand" distribution
    - Known (distribution known from exogenous sources)
    - Unknown (depend on data as it is coming in)
  - » Feedback
    - We know how many resources were used.
    - Imperfect information on resource utilization.

Professor,

As discussed briefly during the break in Saturday's class I am trying to apply the newsvendor inventory service level formula to optimize the level of cash that is held by my mutual fund in order to meet redemption requests from investors.

Problem

Mutual funds hold a certain percentage of their assets in cash in order to meet redemptions from investors. The exact amount is determined more as a guesstimate rather than systematically. I believe the problem is a variation of the Newsvendor problem discussed in class:

Cost of shortfall: If not enough cash is held, the fund must sell some of its holdings and will experience transaction costs. These costs are deterministic and can be assumed to amount to 0.2% of the transaction volume. We can assume that there are no financing costs if there is a cash shortfall. Cost of excess cash [Ce]: Holding to much cash leads to an opportunity cost of not participating in the market. Daily returns on the fund's portfolio are stochastic, and therefore I am not sure whether the Newsvendor formula we saw in class can be applied. The 'cost' of excess cash may even be a gain on some days when the portfolio is down.

Complications

I think reducing the problem to a Newsvendor problem is a good first approximation, but I see several complications:

(i) There will be some correlation between the error term in the daily return on the portfolio and the probability function of getting redemptions.

(ii) Reducing the problem to a single-period is a simplification for which I don't have a good feeling whether it is significant or not.

(iii) The demand function is likely to be skewed by a few large redemptions. Although there is a large number of atomistic retail investors who would redeem small amounts each, a few large institutional investors might redeem large amounts at a time.

(iv) The zero financing cost assumption does not apply for all redemptions, and in particular may not apply to large redemptions by institutions. Their redemptions proceeds need to be wired the following business day, while corresponding sales of portfolio securities take three business days to settle. However, small redemptions by retail investors are paid by check which take several days to mail and clear, by which time any securities sales will have settled.

I would appreciate if you could point me to some literature that treats Ce as a stochastic variable. Maybe there is already a published solution to my problem (I'm not aware of any)?

Thank you for your help.

Regards,

Part-time MBA student at the University of Chicago

President & Portfolio Manager

Lecture outline

- Applications of the newsvendor problem
- The newsvendor problem
- Estimating the distribution and censored demands
- The newsvendor problem and risk
- The newsvendor problem with an unknown distribution

The newsvendor problem

Let:

Parameters

c^o = Cost of overage (cost of ordering one unit too many)

c^u = Cost of underage (cost of ordering one unit too few)

Decision variables

x = Order quantity

Activity variables

$D(\omega)$ = Realization of random demand (assume it is continuous)

$p(\omega)$ = Probability of an outcome ω

$S^o(\omega)$ = Overage = $[x - D(\omega)]^+$

$S^u(\omega)$ = Underage = $[D(\omega) - x]^+$

The newsvendor problem

Objective function:

$$F(x,\omega) = c^o S^o(x,\omega) + c^u S^u(x,\omega)$$
$$= c^o[x - D(\omega)]^+ + c^u[D(\omega) - x]^+$$

$$F(x) = E\{F(x,\omega)\}$$
$$= \int_\omega \left(c^o[x - D(\omega)]^+ + c^u[D(\omega) - x]^+\right) p(\omega)\, d\omega$$

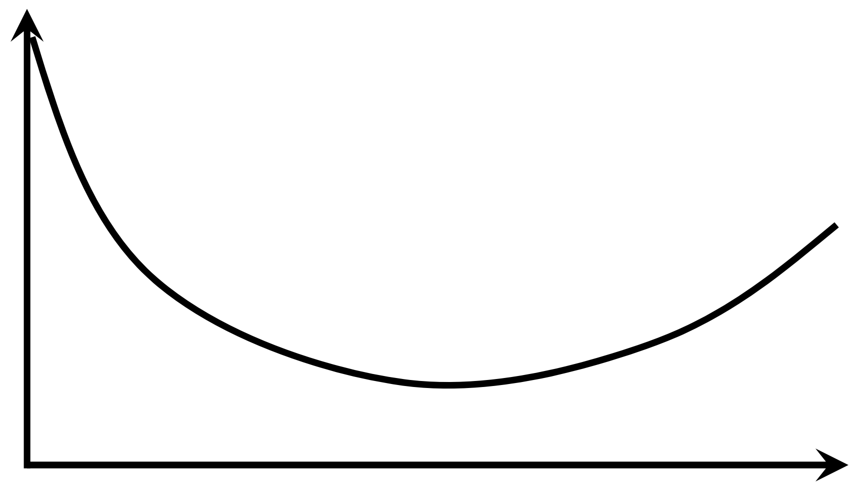
We want to solve:

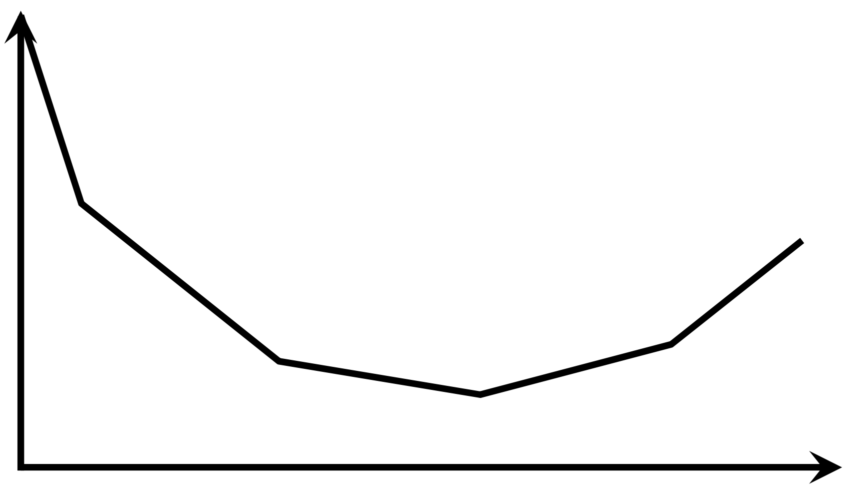
$$\min_x F(x) = E\{F(x,\omega)\}$$

The newsvendor problem

With just a couple assumptions (we will figure these out later), $F(x)$ will look like:

Continuous distribution



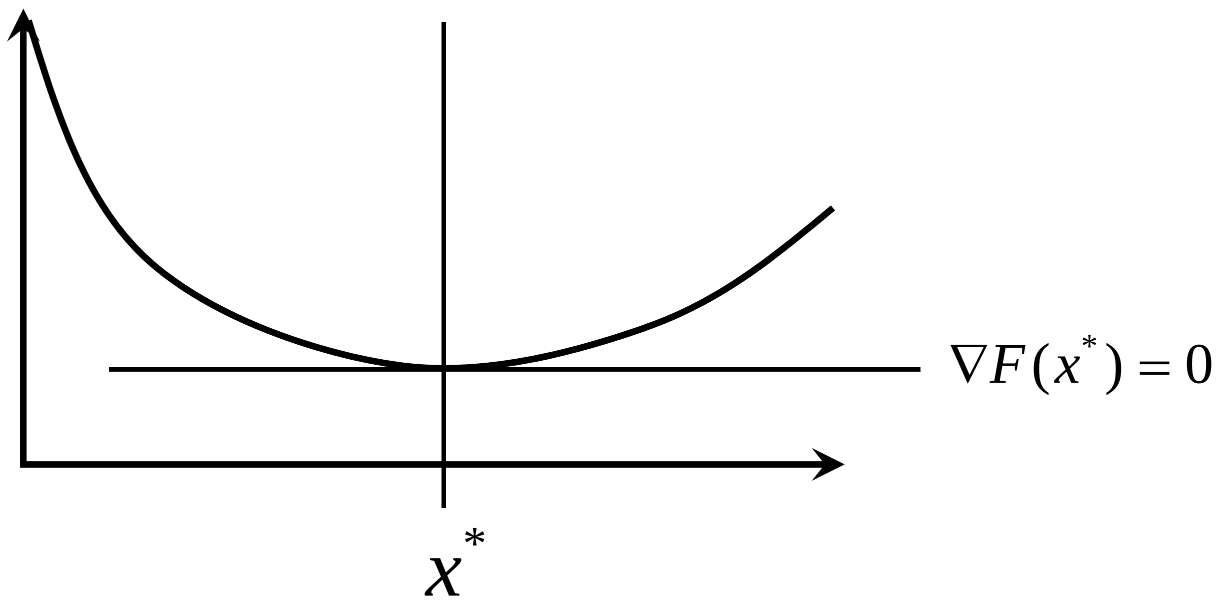


Discrete distribution

The newsvendor problem

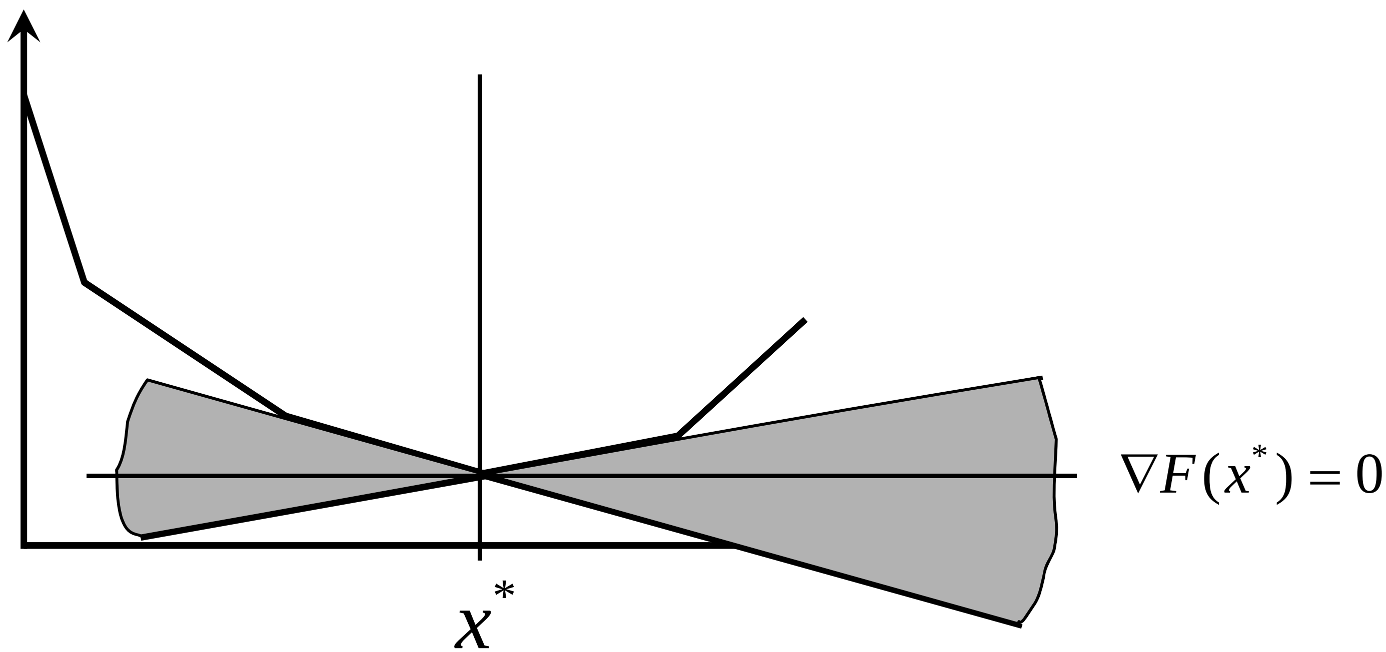
When the underlying random variable is continuous, the function $F(x)$ is continuous.

At the optimum, the gradient will equal to zero



The newsvendor problem

When the underlying random variable is discrete, the function $F(x)$ is piecewise linear. At the break points, there is more than one gradient. At $x = x^*$, there will be *one* gradient equal to zero:



The newsvendor problem

Let's say that we have found the optimal solution x^*. Normally, we could claim that the derivative of a function at the optimum would equal zero. But it will not generally be the case that:

$$\nabla F(x^*, \omega) = 0$$

In fact, we may find that this is *never* true for any ω. But, we should find that:

$$\mathbb{E}\nabla F(x^*, \omega) = 0$$

If:

$$F(x,\omega) = c^o[x - D(\omega)]^+ + c^u[D(\omega) - x]^+$$

then:

$$\nabla F(x^*, \omega) = \begin{cases} c^o & \text{if } D(\omega) \le x^* \\ -c^u & \text{if } D(\omega) > x^* \end{cases}$$

The newsvendor problem

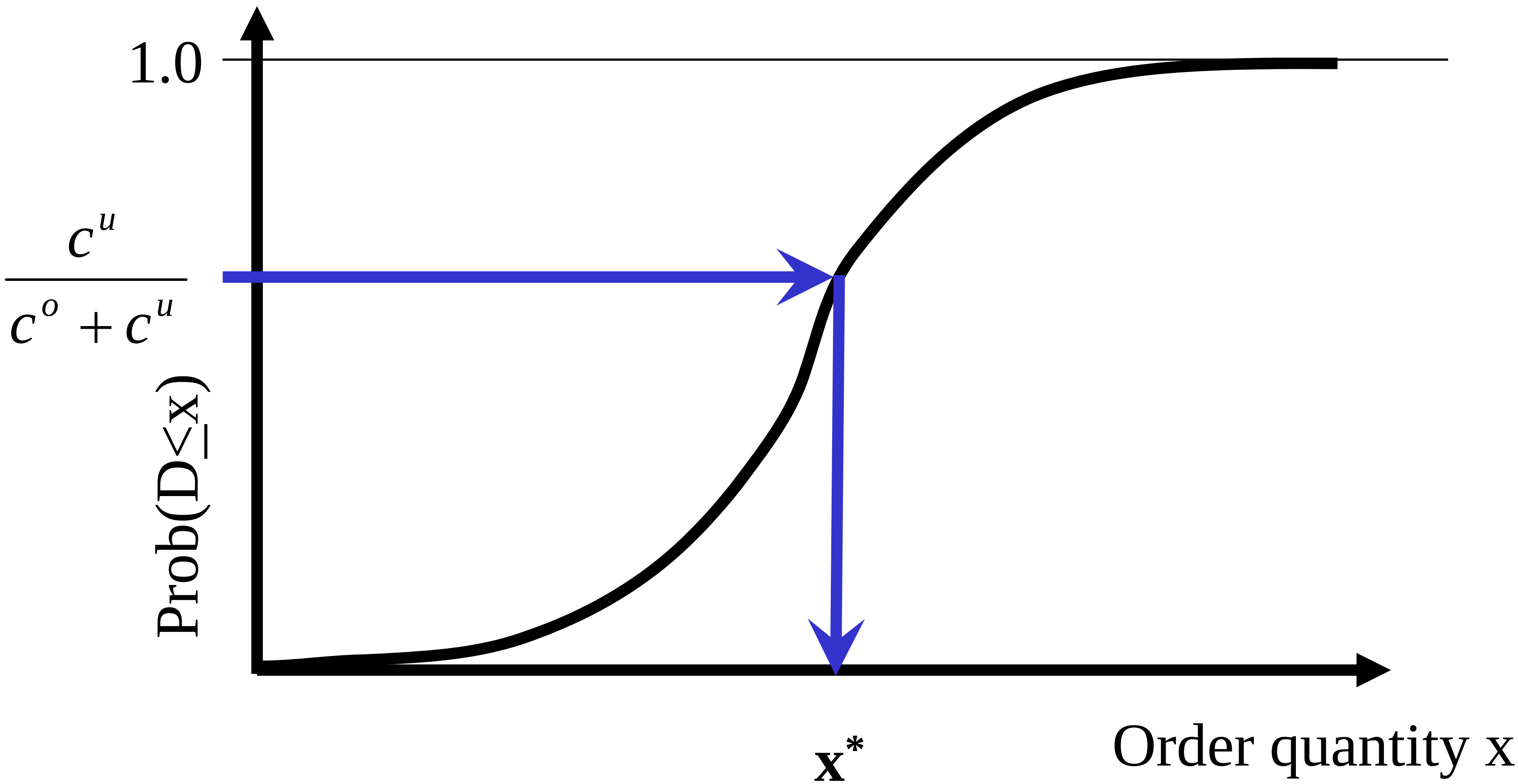
So:

$$
\begin{aligned}
\mathrm{E}\nabla F(x^*, D) &= c^o \text{Prob}[\, D_t(\omega) \le x^* \,] - c^u \text{Prob}[\, D_t(\omega) > x^* \,] \\
&= c^o \text{Prob}[\, D_t(\omega) \le x^* \,] - c^u \left(1\text{-Prob}[\, D_t(\omega) \le x^* \,] \right) \\
&= \left(c^o + c^u \right) \text{Prob}[\, D_t(\omega) \le x^* \,] - c^u \\
&= 0
\end{aligned}
$$

Rearranging gives us:

$$
\text{Prob}[\, D_t(\omega) \le x^* \,] = \frac{c^u}{c^o + c^u} \qquad \text{The "critical ratio"}
$$

The newsvendor problem



The newsvendor problem

- The profit maximizing version:

Let:

$c =$ Unit cost of purchasing product.

$p =$ Price we sell our product for.

$x =$ Quantity ordered.

$D =$ Random demand for product

Total conditional profits given demand $D(\omega)$:

$$F(x,\omega) = p\min\left(x, D(\omega)\right) - cx$$

Problem is to solve:

$$\max_x EF(x,\omega) = E\left\{p\min\left(x, D\right) - cx\right\}$$

Now find the critical ratio....

The newsvendor problem

- What cell phone plan should I use?

Plan Type	Map	Description	Monthly Access	Monthly Airtime Allowance (in minutes)	Per Minute Rate after allowance	Select Plans
America's Choice[sm]	View Map	Call anywhere coast to coast from anywhere on the America's Choice network with no roaming or long distance charges.	$34.99-$199.99	300- 3,200	$0.20-$0.45	○
America's Choice[sm] with Push to Talk	View Map	All the benefits of the America's Choice plans plus Push to Talk, Verizon Wireless' new walkie-talkie service. Unlimited Push to Talk one-to-one calls and only 0.15 per minute per participant for Push to Talk Group calls	$59.99-$219.99	400- 3,200	$0.20-$0.45	○
Promotional America's Choice[sm] Family SharePlan[sm]	View Map	Share your minutes plus, **great Buy one Phone Get another FREE offers available ONLINE !**	$34.99 - $199.99	300-3200	$0.20-$0.45	○
America's Choice[sm] Family SharePlan[sm]	View Map	Share the value. One account. One monthly bill. **Plus**, 250 included national mobile to mobile airtime minutes for each line of service.	$34.99 - $199.99	300- 3200	$.20-$.45	○
National SingleRate[sm]	View Map	Perfect if you talk and travel around the country. Domestic roaming and domestic long distance included.	$35-$300	150- 3,000	$0.20-$0.40	○

The newsvendor problem

- Cost structure of my Verizon plan:
  - » $0.14/minute of guaranteed time.
  - » $0.45/minute if exceed guaranteed minimum.

  - » Ex: 500 minutes per month
    - $70 fixed monthly fee
    - $0.45/minute over 500 minutes.

The newsvendor problem

- Finding the optimal plan:

Let:

x = Guaranteed minutes per month

M =random variable giving minutes used per month.

Objective function: ???

The newsvendor problem

Analysis:

Stochastic gradient:

$$\nabla F(x^*, M(\omega)) = \begin{cases} 0.14 & M(\omega) < x^* \\ -0.31 & M(\omega) \geq x^* \end{cases}$$

Find:

$$\mathrm{E}\nabla F(x^*, M(\omega)) = 0$$

$$= .14\text{Prob}(M(\omega) < x^*) - .31\text{Prob}(M(\omega) \geq x^*)$$

$$= .14\text{Prob}(M(\omega) < x^*) - .31\left(1 - \text{Prob}(M(\omega) < x^*)\right)$$

$$= -0.31 + 0.45\text{Prob}(M(\omega) < x^*)$$

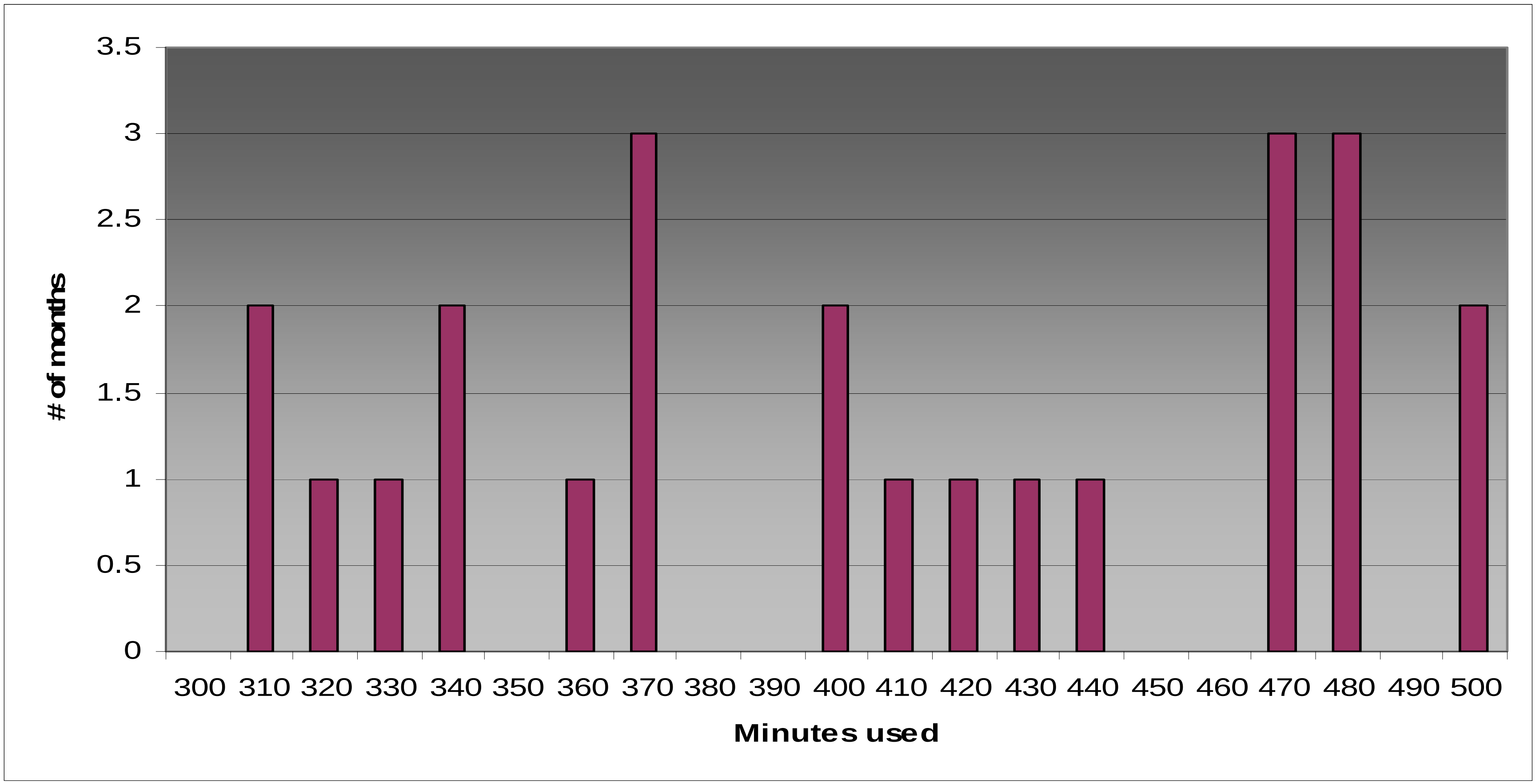
$$\text{Prob}(M(\omega) < x^*) = 0.31 / 0.45$$

$$\approx 0.70$$

I should exceed my minutes (roughly) 30 percent of the time.

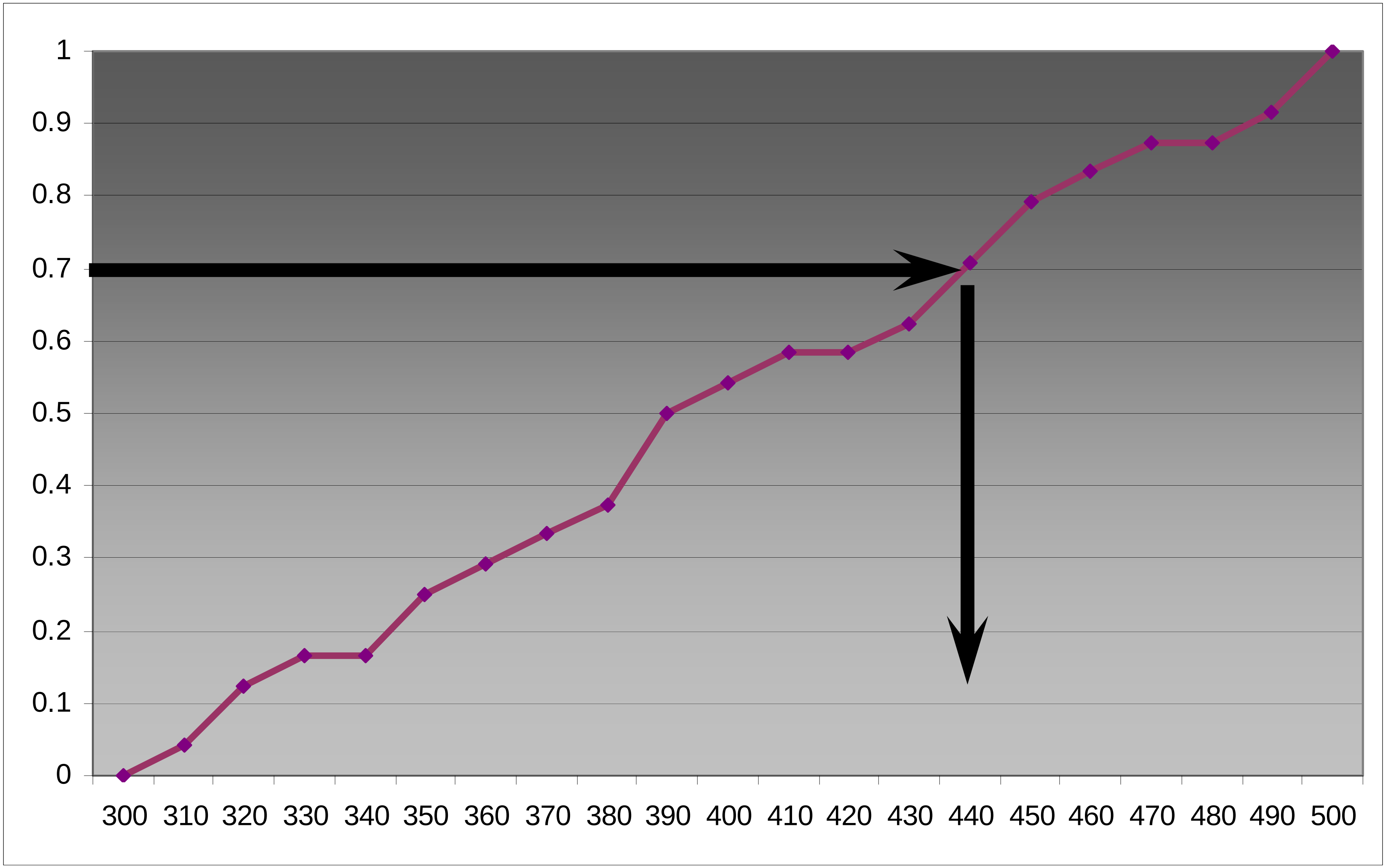
The newsvendor problem

- From past phone records, we can construct a histogram of the number of minutes used per month.



`
The newsvendor problem

- … from which we can compute a cumulative distribution of minutes used.



The newsvendor problem

- Practical challenges in applying the newsvendor problem:
  - » Estimating the probability distribution
    - Use history from similar experiences
    - Use judgment
    - (This part is hardest with "one-shot" problems).
  - » Estimating the cost of overage and underage
    - Usually one of these two costs is "soft"
      - – Example: raising capital for startup:
        - » Cost of raising too much: cost of capital
        - » Cost of raising too little: need to estimate the cost of returning to the financial markets.

The newsvendor problem

- Notes on solving newsvendor problems
  - Do *NOT* memorize formulas for overage and underage.
  - Start by writing out the objective function.
  - Take the derivative with respect to the order quantity. You will usually find you obtain two values for the derivative – one if you order too much, and one if you order too little.
  - Take the expectation of the gradient and set it equal to zero. Solve for the optimal order quantity in terms of the probability of being over or under.

Lecture outline

- Applications of the newsvendor problem
- The newsvendor problem
- Estimating the distribution and censored demands
- The newsvendor problem and risk
- The newsvendor problem with an unknown distribution

Censored demands

- Setup:
  - » We make an allocation for period t.
  - » Then learn requirement and compute cost.
  - » Process repeats over and over.
- What do we do?
  - » Use history to build probability distribution of “demand”
  - » Update distribution periodically (e.g. after each observation).
  - » Apply standard newsvendor logic.
- Challenge:
  - » We typically cannot observe actual demand when we ordered too little.

Censored demands

- Working with censored demands
  - » In general, we estimate demand based on actual sales, rather than real demand.
  - » This is often referred to as “censored data.” It means that we are not working with a complete dataset.

Censored demands

- A naïve approach:
  - » Estimate demand based on actual sales

Let:

$$x_t = \text{ Order quantity at time } t$$

$$D_t(\omega) = \text{ Sample realization of demand at time } t$$

$$\hat{D}_t(\omega) = \text{Observed demand}$$

$$= \min\{x_t, D_t(\omega)\}$$

We can estimate the mean and variance using observed demands:

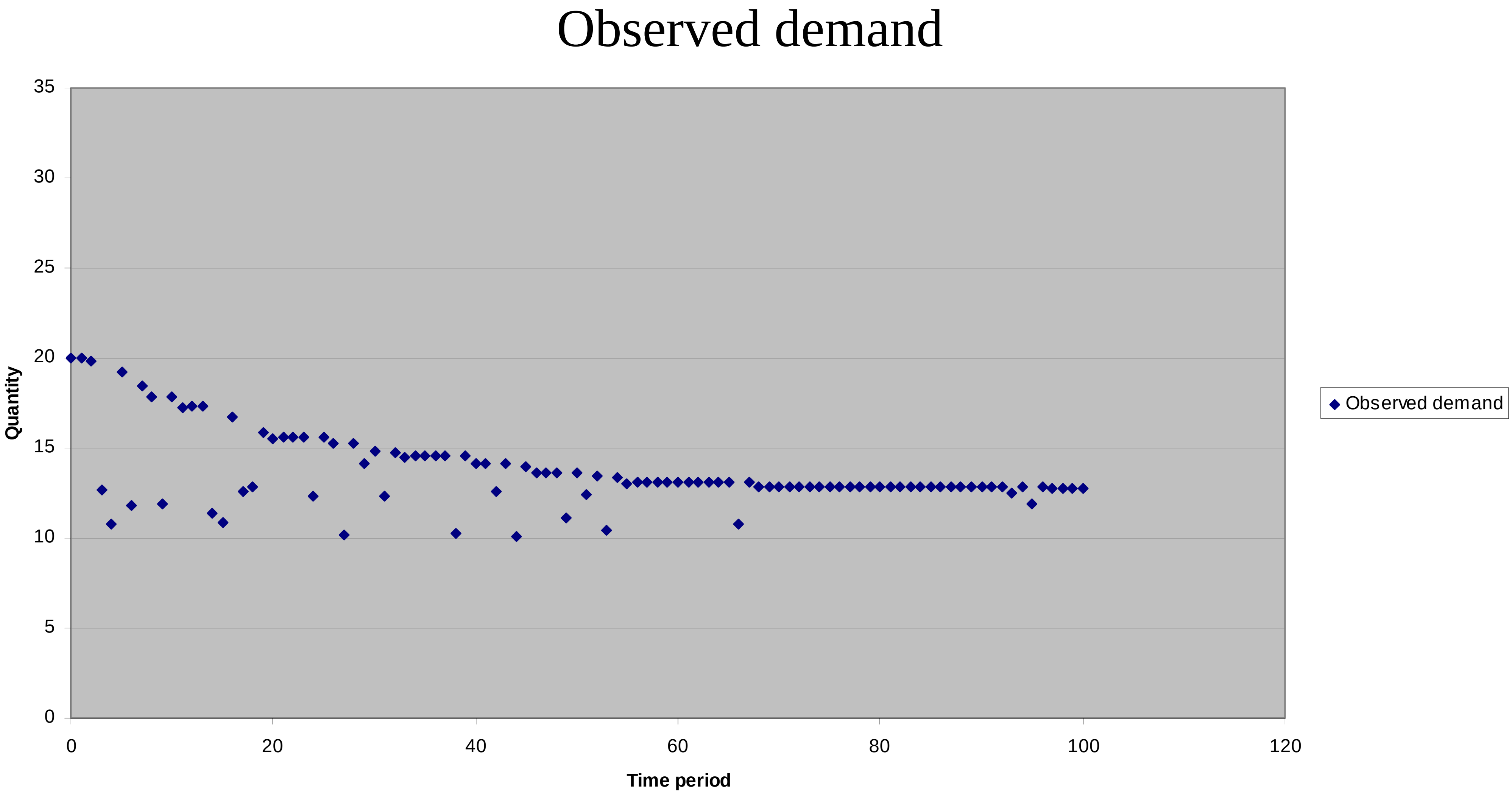
$$\mu_t = (1-\alpha_t)\mu_{t-1} + \alpha_t \hat{D}_t(\omega)$$

$$s_t^2 = (1-\alpha_t)s_{t-1}^2 + \alpha_t\left(\mu_{t-1} - \hat{D}_t(\omega)\right)^2$$

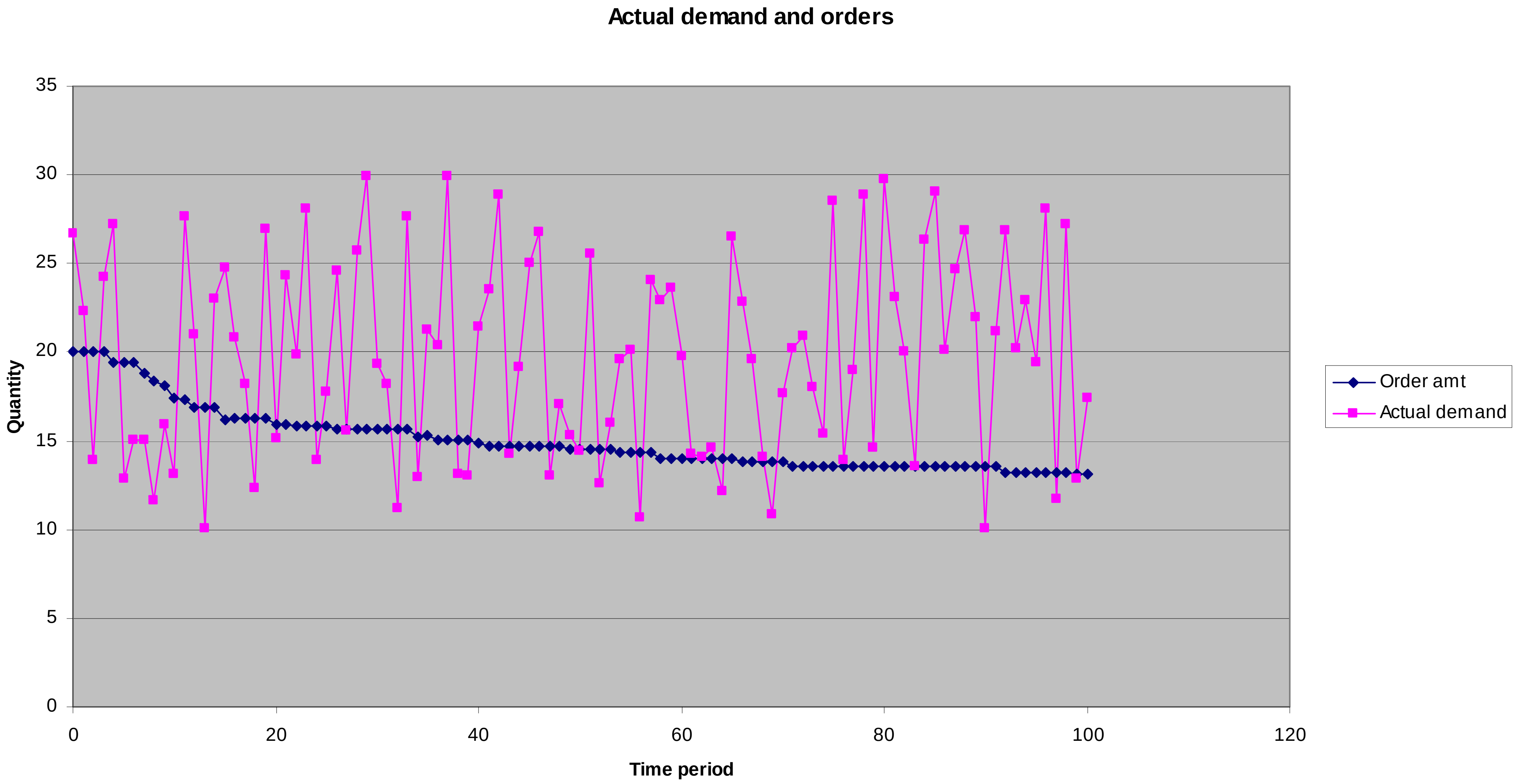
Set order quantity using:

$$x_t = \mu_t$$

Censored demands



Censored demands



Censored demands

- Alternative:
  - » Use "newsvendor" concept:

Let:

$s_t^2 =$ Estimate of the variance at time t

Compute a smoothed estimate of the variance:

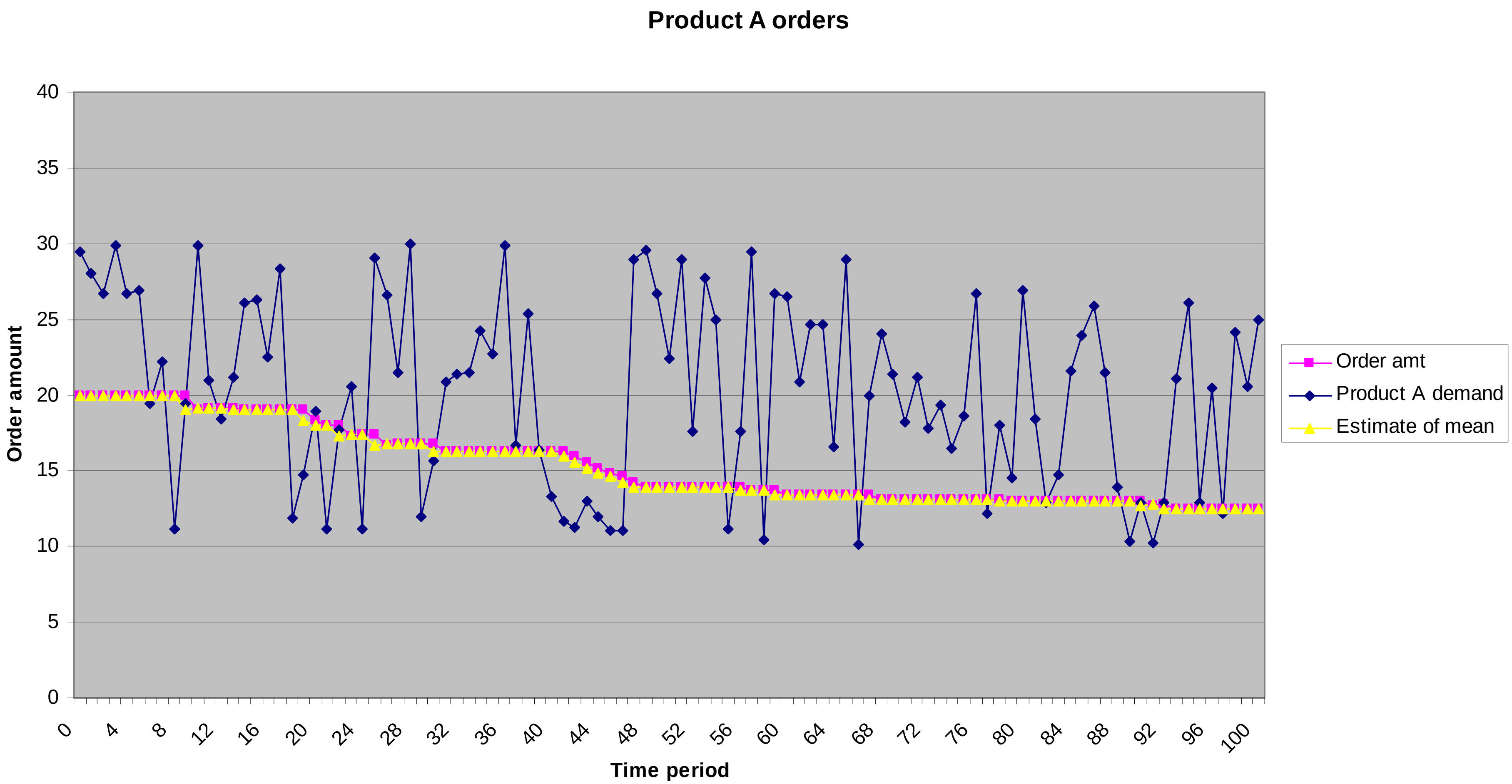
$$s_t^2 = (1-\alpha_t) s_{t-1}^2 + \alpha_t \left(x_{t-1} - \hat{D}_t(\omega) \right)^2$$

Set order quantity using:

$$x_t = \mu_t + z\sqrt{s_t^2}$$

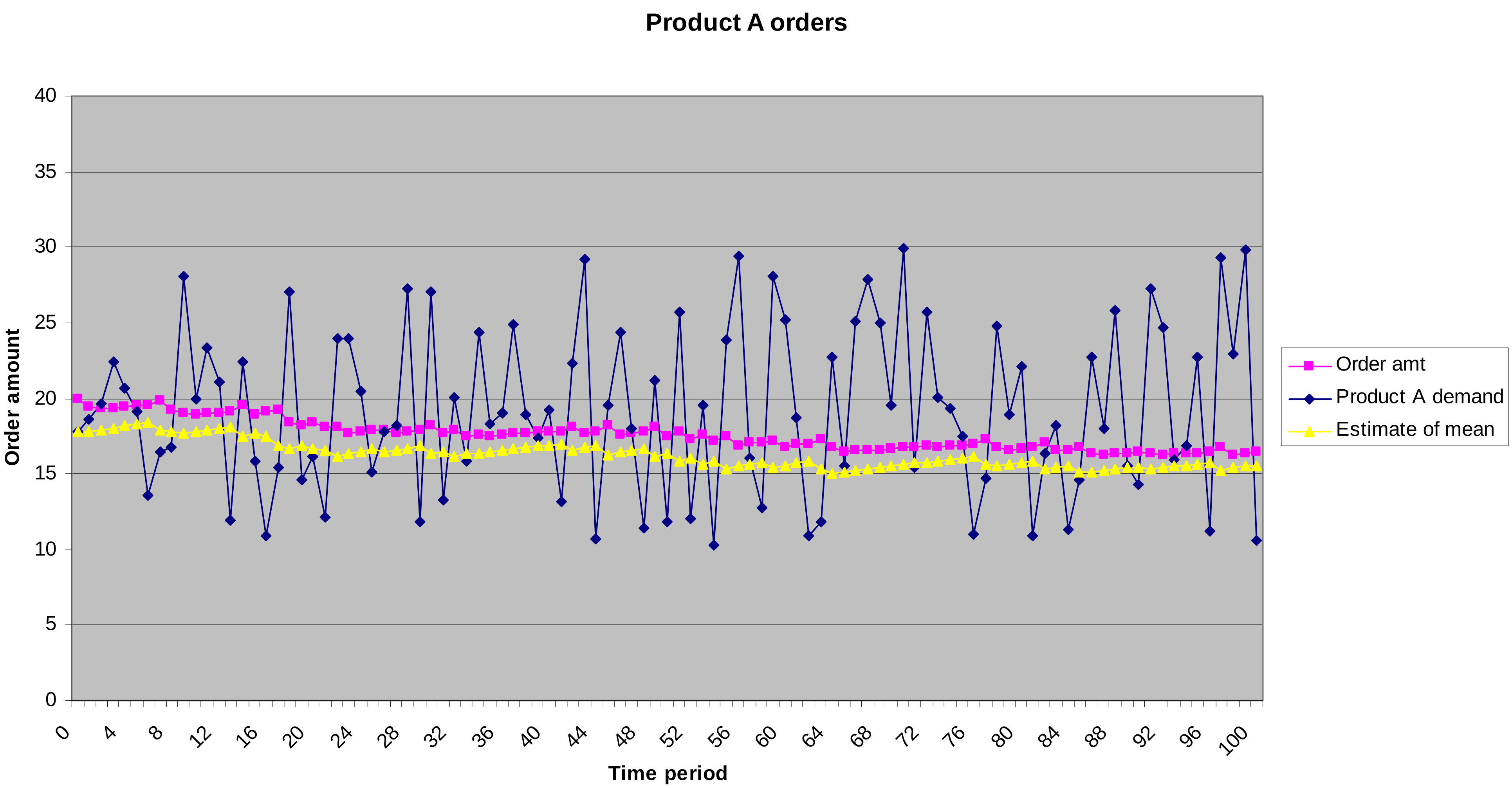
Where z is chosen based on the newsvendor problem.

Censored demands



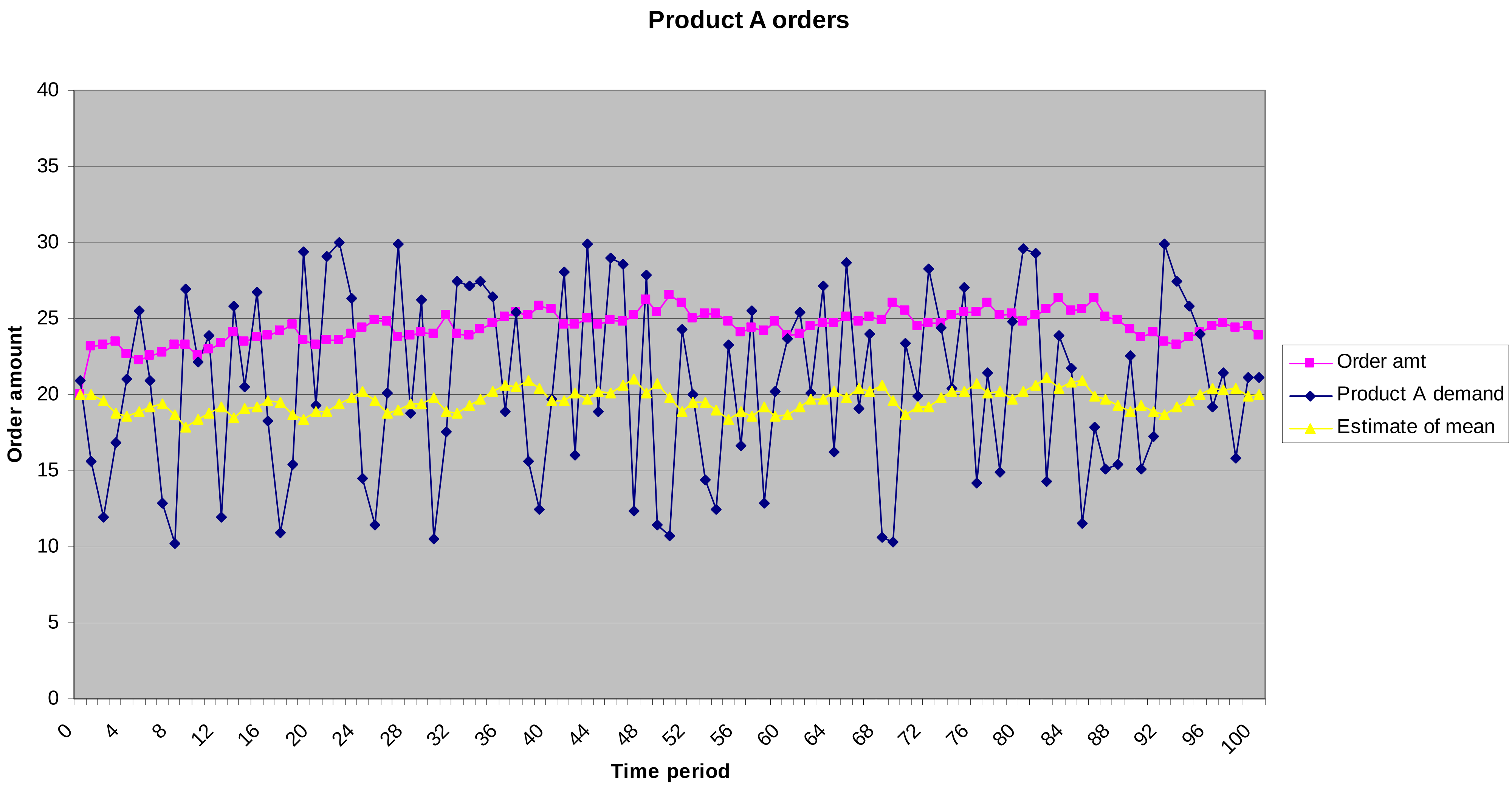
Using Z = 0

Censored demands



Using Z = .5

Censored demands



Using Z =1.0

Censored demands

- Notes:
  - » Ordering extra quickly reduces the downward bias.
  - » The price of this information is the cost of purchasing and holding additional product.
  - » If the optimal solution is to cover a high percentage of demand, the error is small.
  - » But if the optimal solution requires a significant amount of lost demand, you need to either:
    - Periodically over order just to estimate the demand.
    - Use specialized results that compensate for censored demands.

Lecture outline

- Applications of the newsvendor problem
- The newsvendor problem
- Estimating the distribution and censored demands
- The newsvendor problem and risk
- The newsvendor problem with an unknown distribution

Newsvendor and risk

- Using the optimal solution to the newsvendor problem can expose you to high levels of risk
  - » Assume that the demand follows a Pareto distribution:

$$F_D(y) = P\left[D^n \le y\right] = \begin{cases} 1 - \left(\frac{\beta}{y}\right)^{\alpha} & y \ge \beta \\ 0 & y < \beta \end{cases}$$

  - » Moments:
    - Mean:

$$\text{For } \alpha > 1: \quad \mathbb{E}X = \frac{\alpha\beta}{\alpha - 1} \qquad \text{For } \alpha \le 1: \ \mathbb{E}X = \infty$$

    - Variance:

$$\text{For } \alpha > 2: \quad Var\mathrm{X} = \frac{\alpha\beta^2}{(\alpha - 1)^2(\alpha - 2)} \qquad \text{For } \alpha \le 2: Var\mathrm{X} = \infty$$

Newsvendor and risk

- Let's see what happens when we choose parameters so that the mean and variance are infinite:
  - » If we use β=1,α=1/2:

$$F_D(y) = P\left[D^n \le y\right] = \begin{cases} 1-\left(\frac{1}{y}\right)^{1/2} & y \ge 1 \\ 0 & y < 1 \end{cases}$$

  - » We sample from this distribution using the relationship:

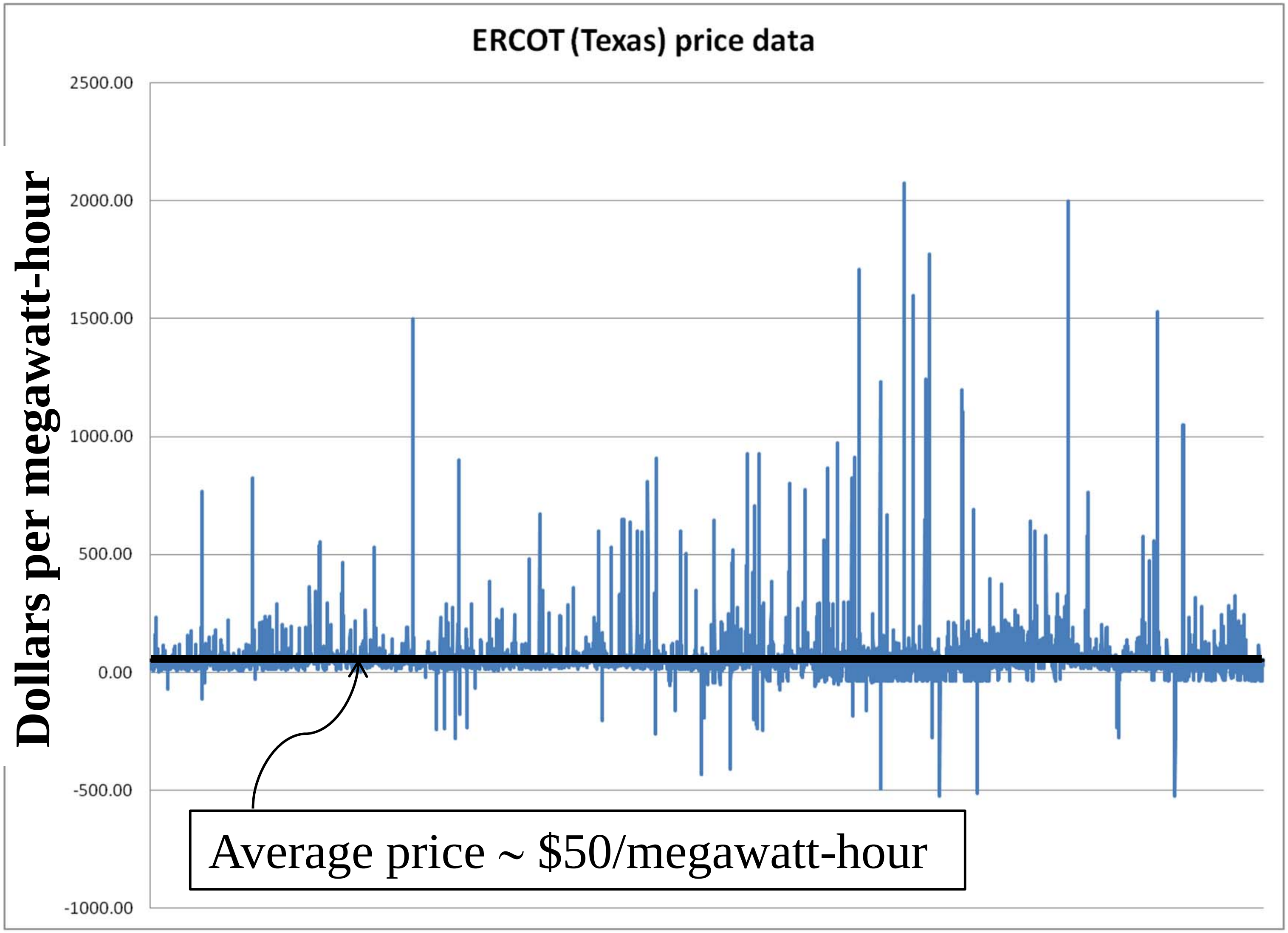
$D \sim F_D^{-1}(U)$ where U is uniformly distributed between 0 and 1.

So we can sample observations (in Excel) using:

$$D = \frac{1}{(1-rand())^2}$$

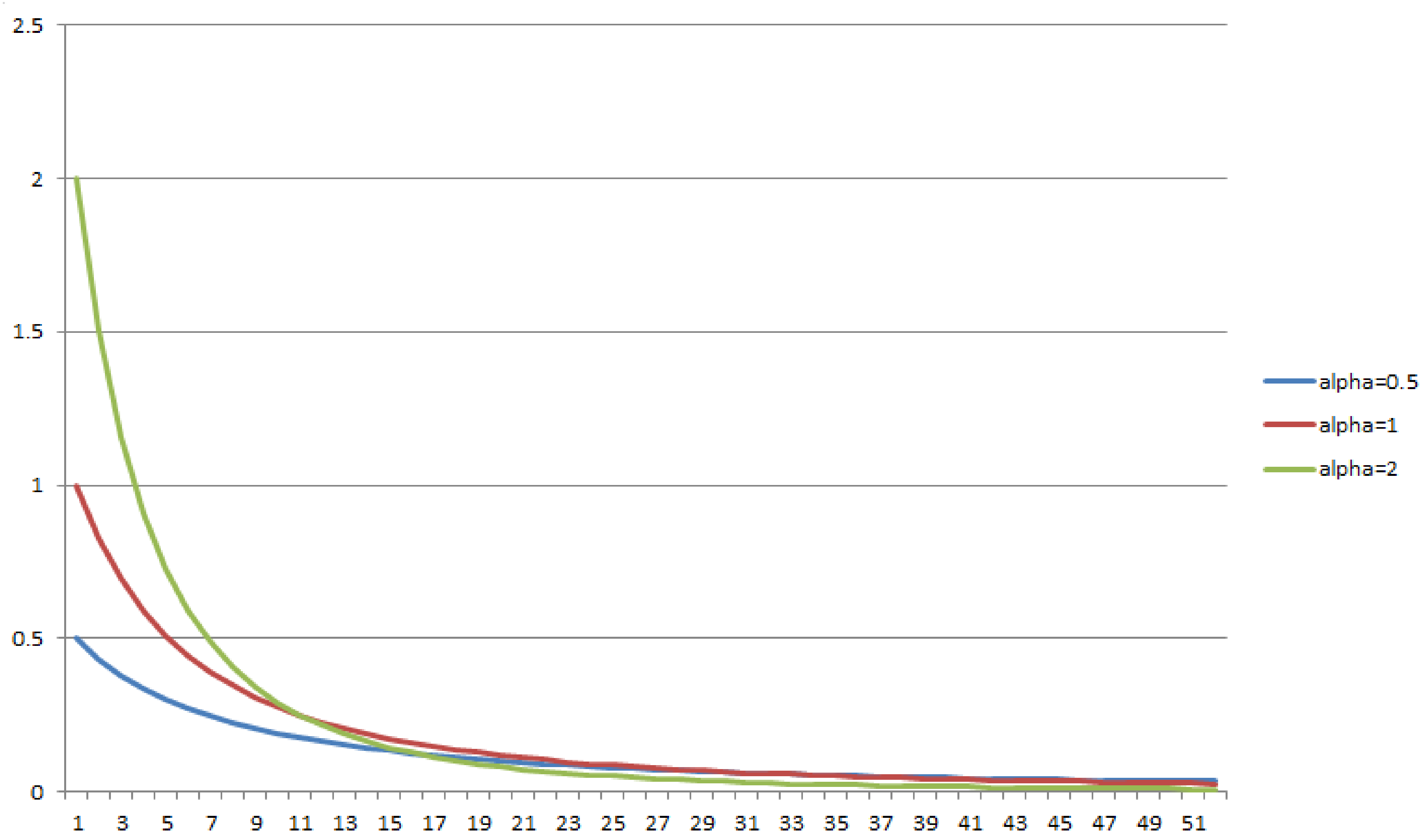
Newsvendor and risk

- Electricity spot prices



Newsvendor and risk

- Density:



Newsvendor and risk

- Expected profits from newsvendor:

$$\mathbb{E}\left[p\min\left\{x,D^n\right\}\right]-cx=\frac{p}{2}\int_{y=1}^{x}y^{-1/2}dy+10xP\left[D^n>x\right]-cx$$

$$=py^{1/2}\,|_{y=1}^{x}+px^{1/2}-cx$$

$$=px^{1/2}-p+px^{1/2}-cx$$

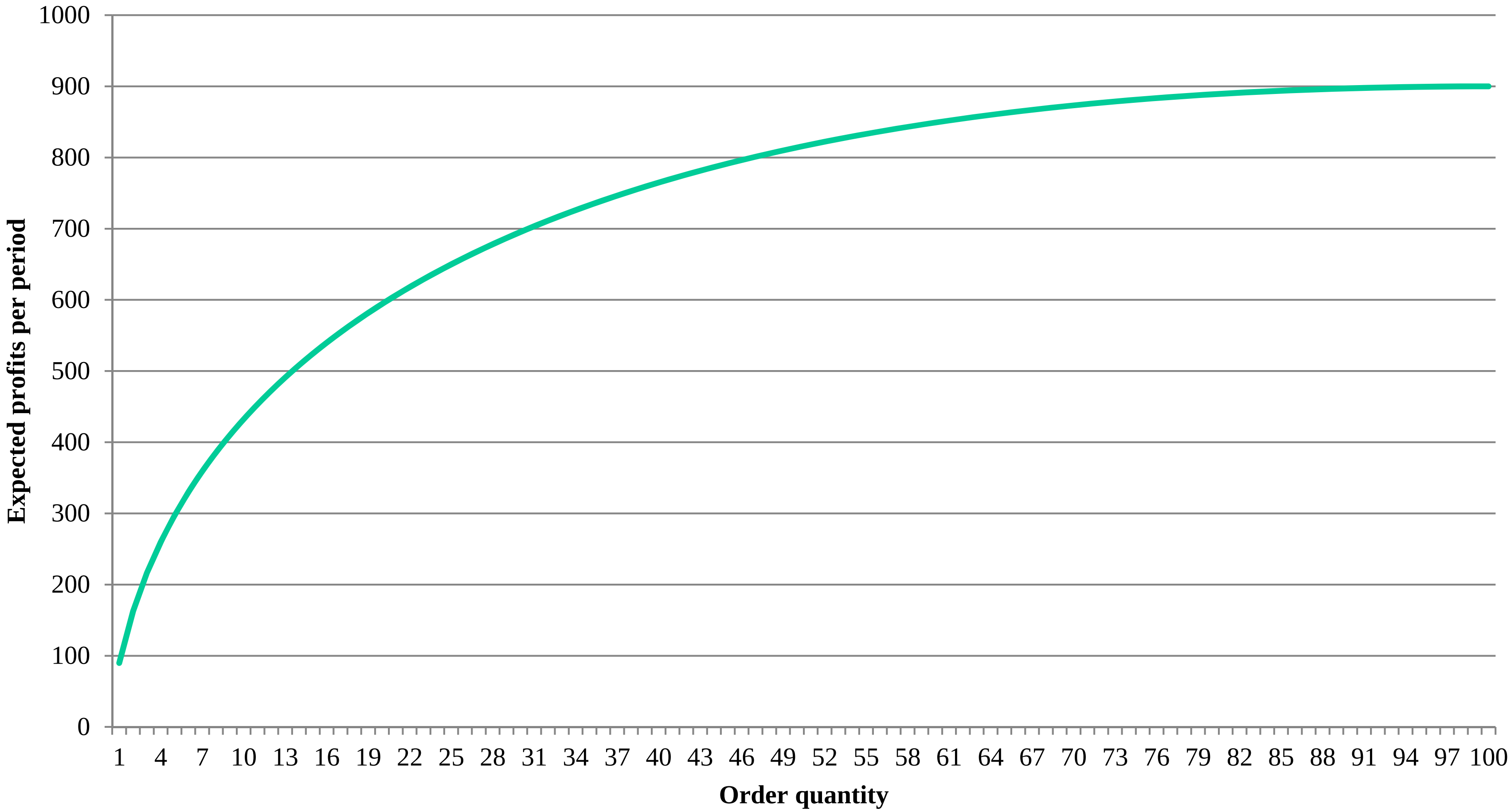
$$=2px^{1/2}-p-cx$$

» The optimal order quantity is found by differentiating:

$$px^{-1/2}-c=0 \quad \Rightarrow x=\left(\frac{p}{c}\right)^2$$

Newsvendor and risk

- Profits as a function of order quantity

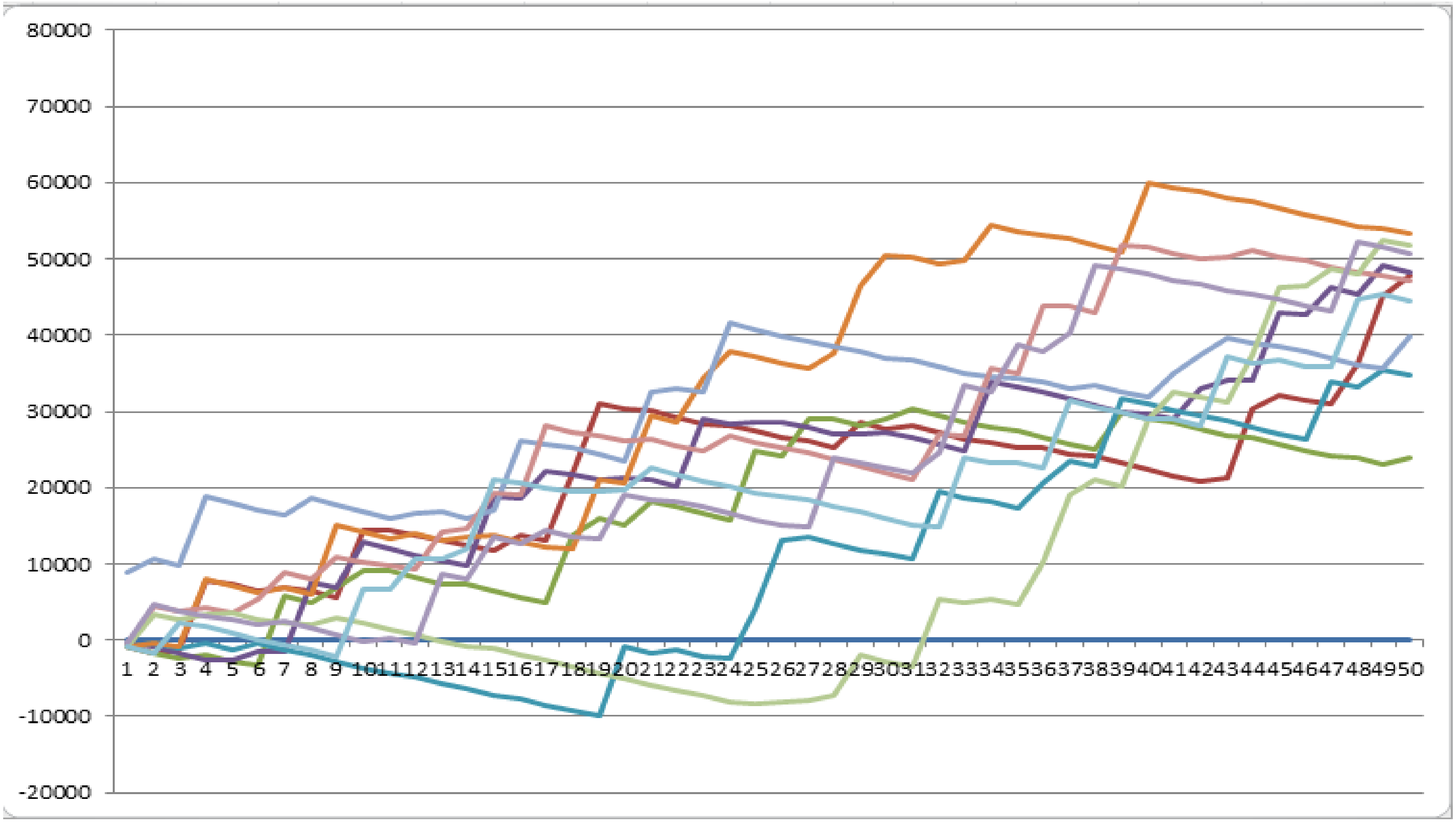


Newsvendor and risk

- Let's use our optimal solution and see how well it does.
  - » p=100
  - » c=10
  - » Optimal order quantity = 100

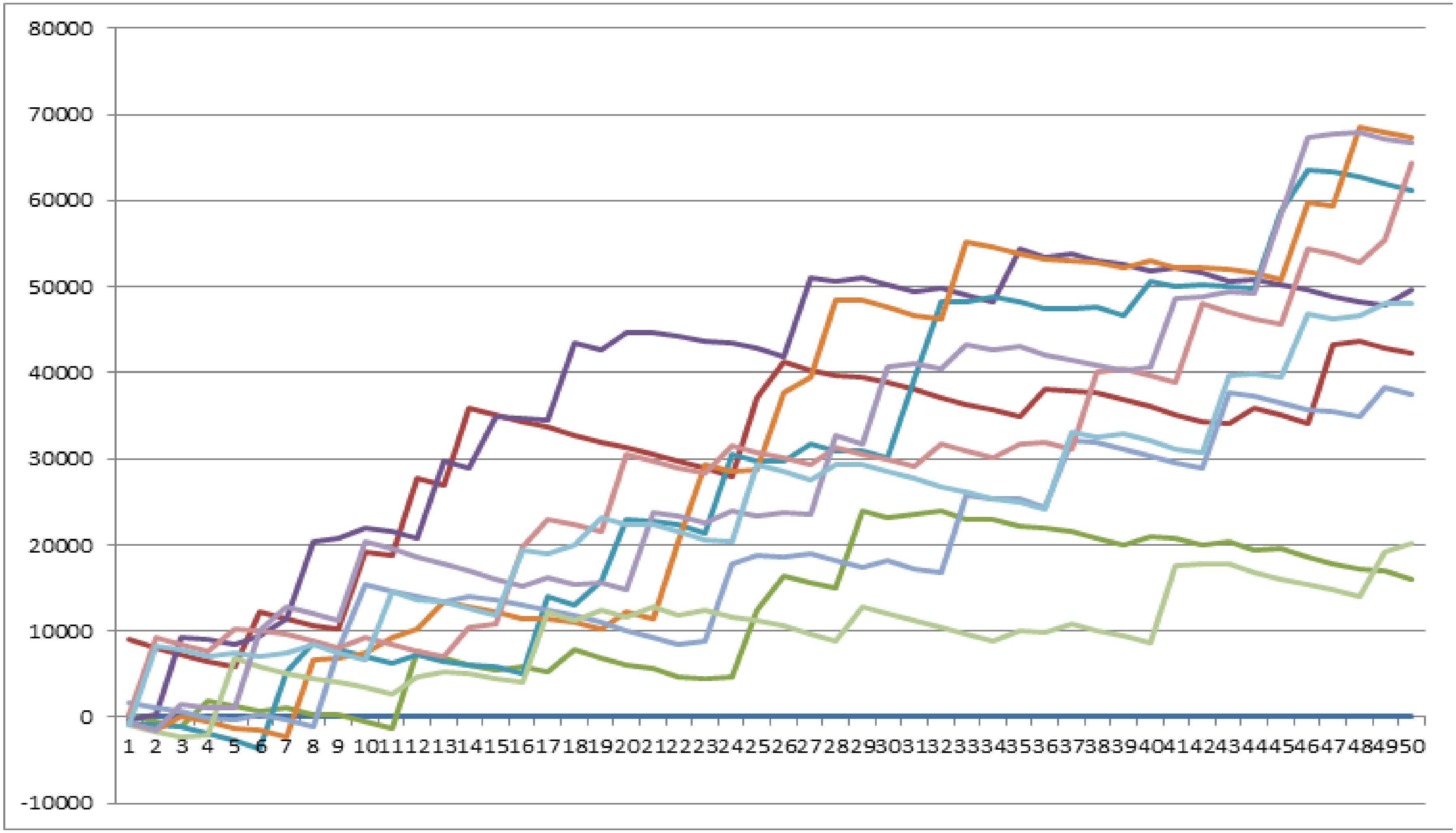
Newsvendor and risk

- Order quantity = 100



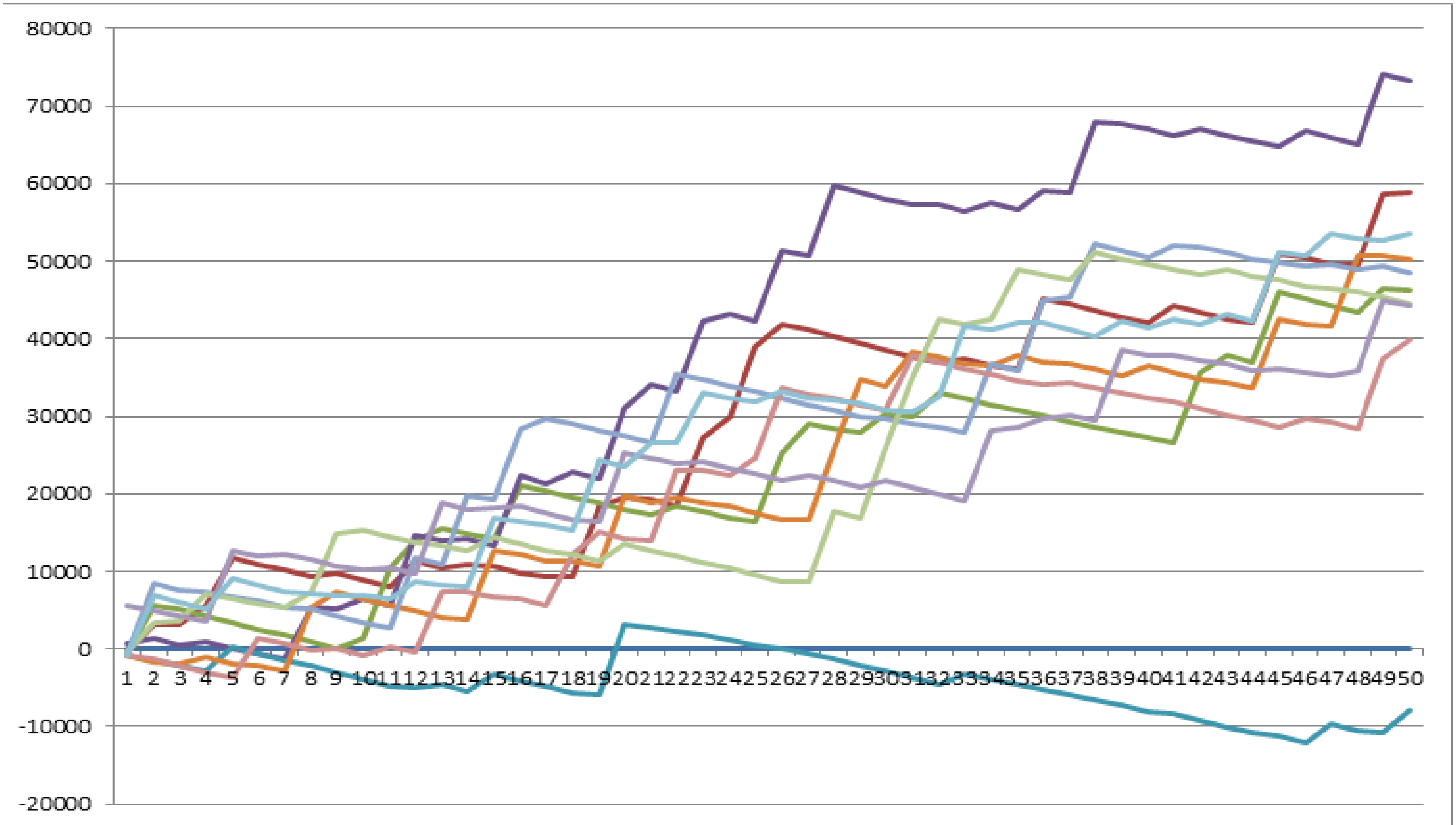
Newsvendor and risk

- Order quantity = 100



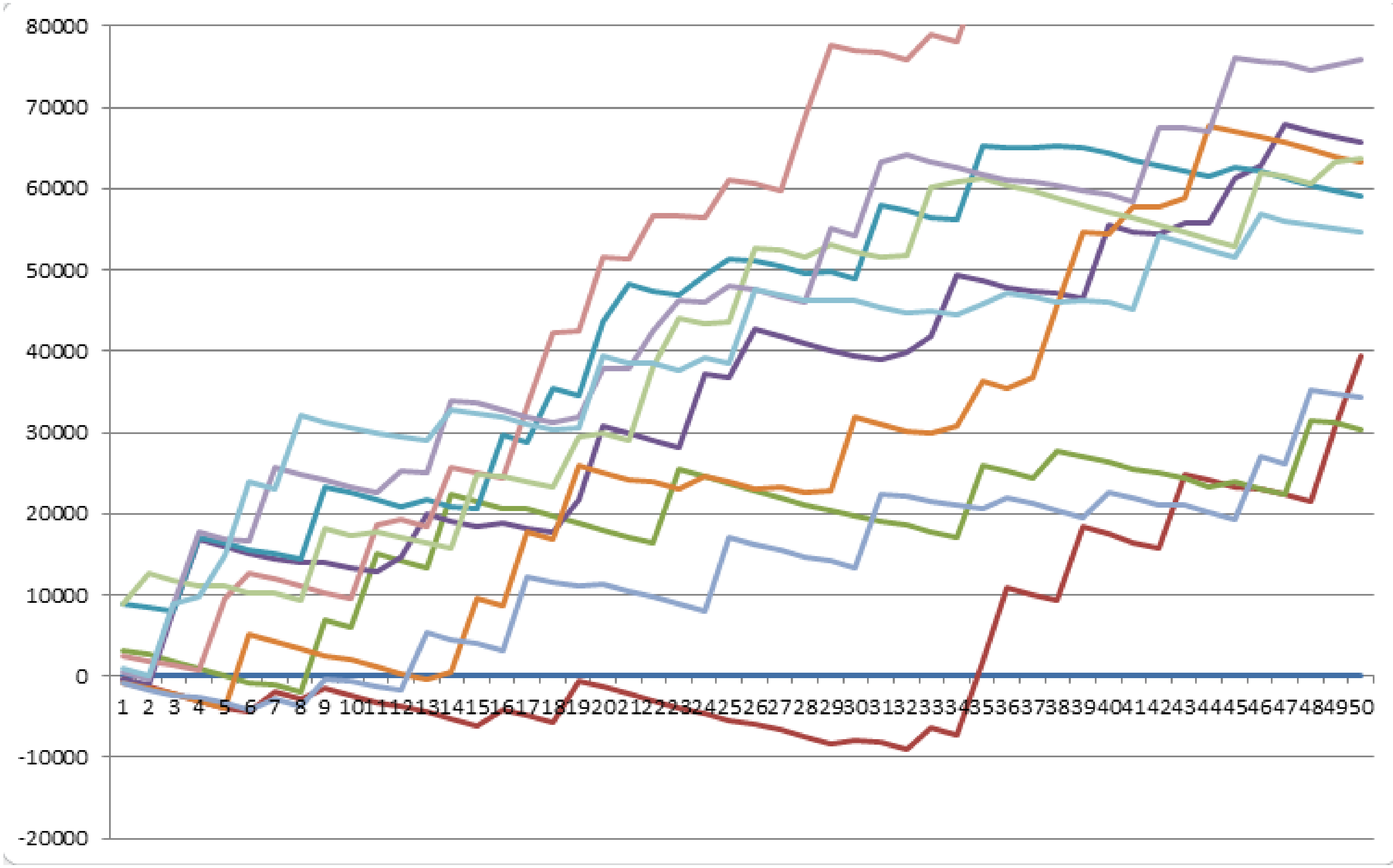
Newsvendor and risk

- Order quantity = 100



Newsvendor and risk

- Order quantity = 100

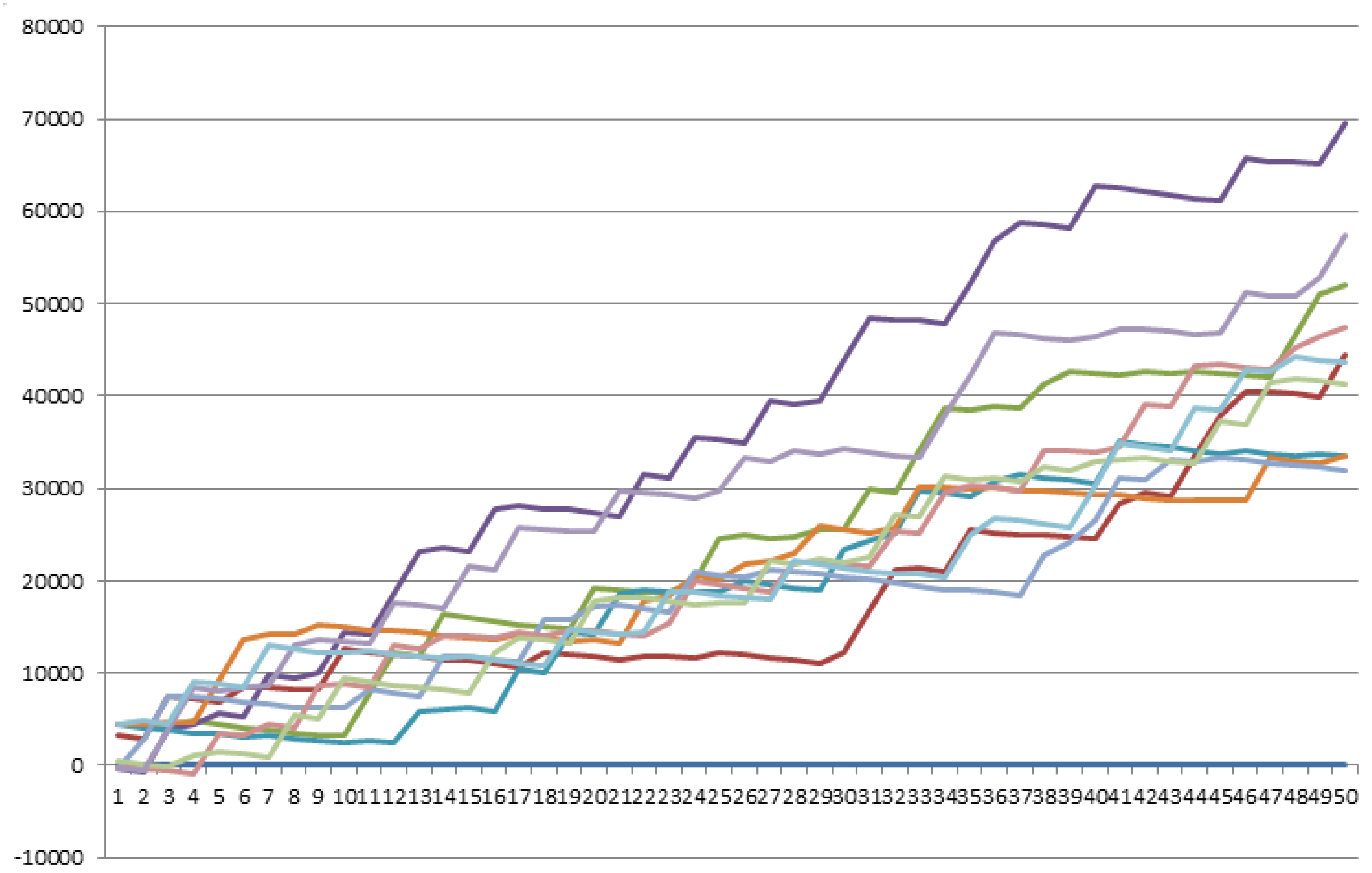


Newsvendor and risk

- Observations:
  - » On average, we are doing well.
  - » But there are frequently sample paths where we lose a lot of money.
  - » The problem is that our order quantity is too high – we are chasing the possibility that the demands might be quite high.
  - » Long stretches of low demand produce large losses.
  - » What if we try a lower order quantity?

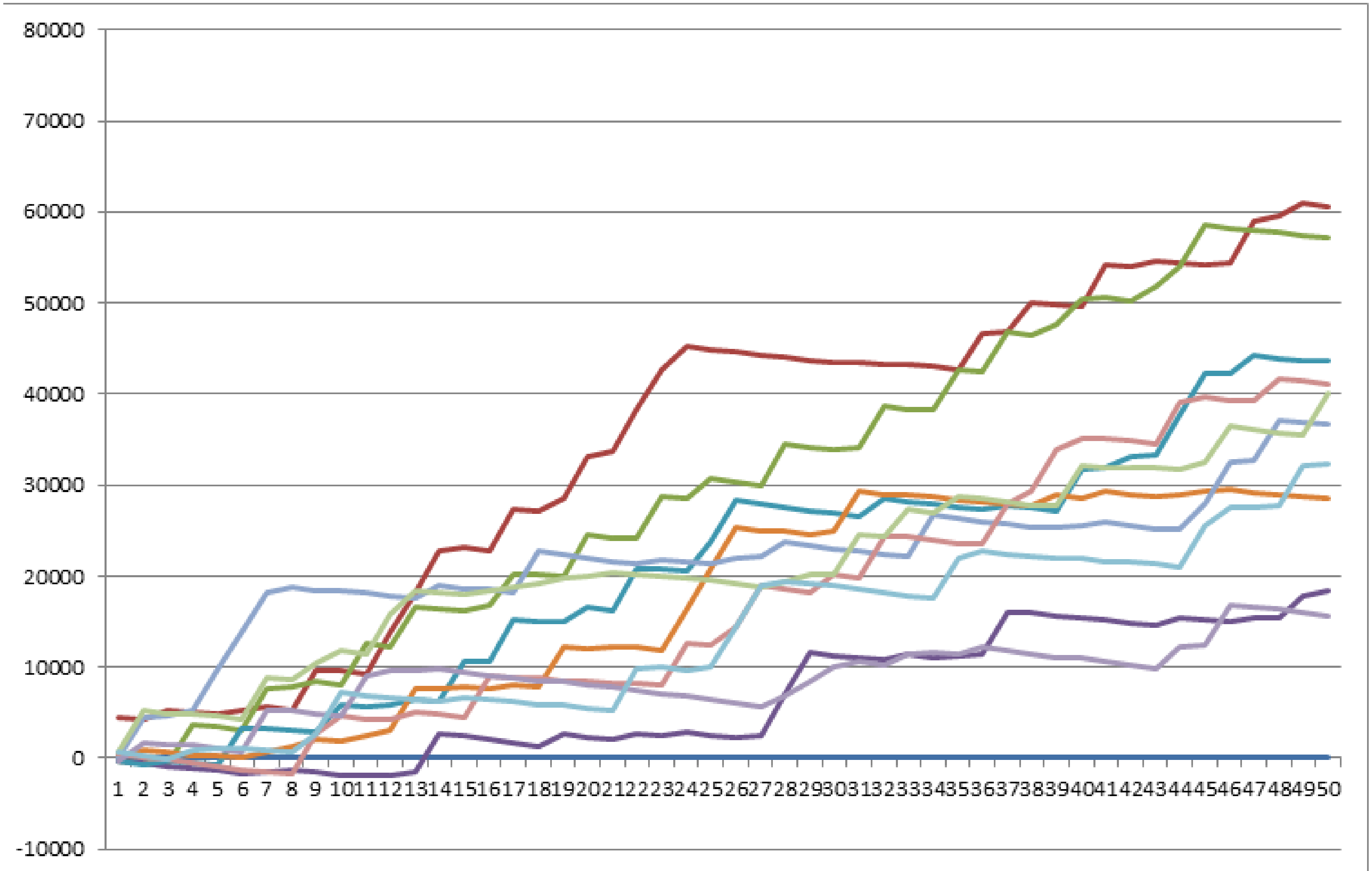
Newsvendor and risk

- Order quantity = 50



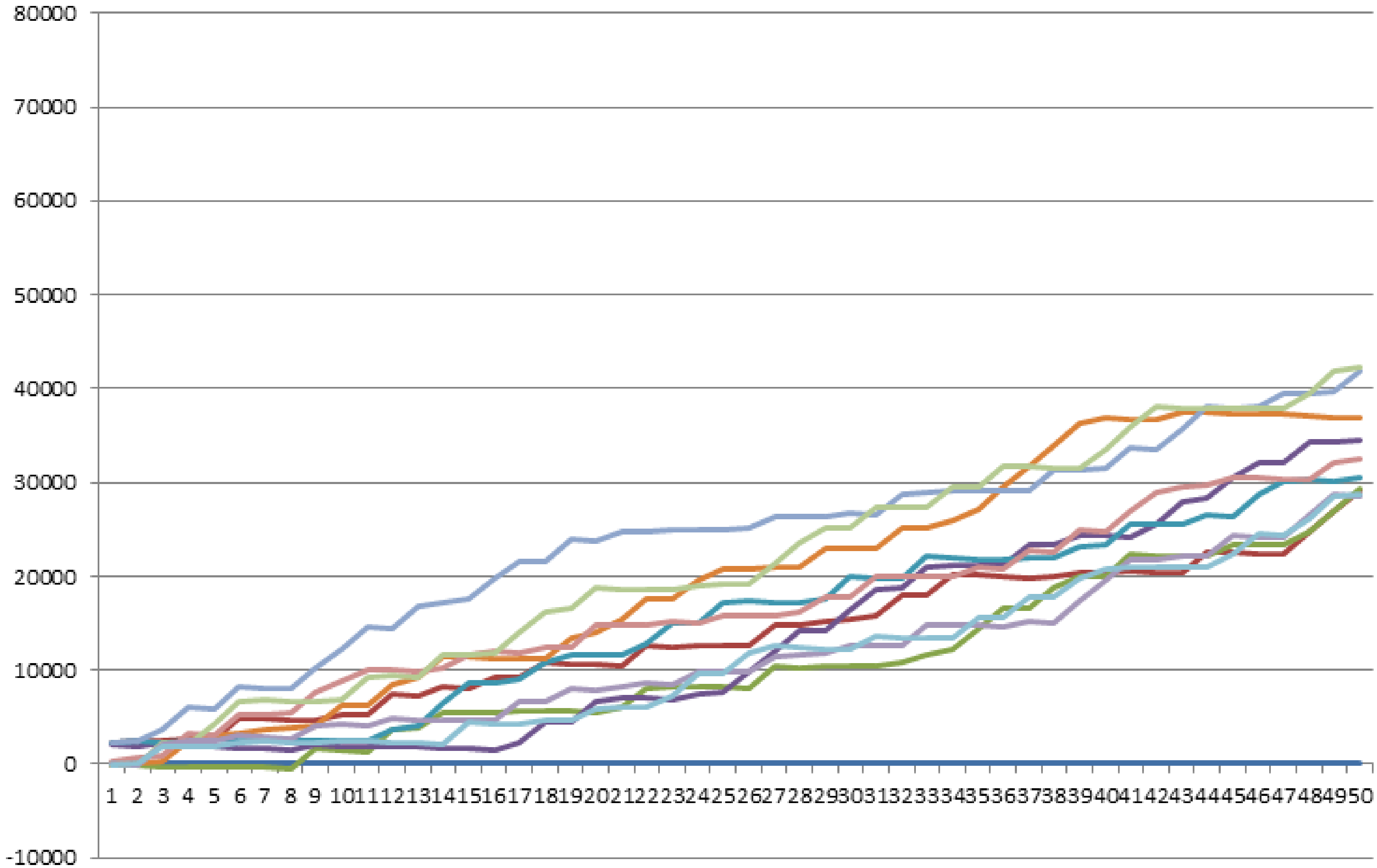
Newsvendor and risk

- Order quantity = 50



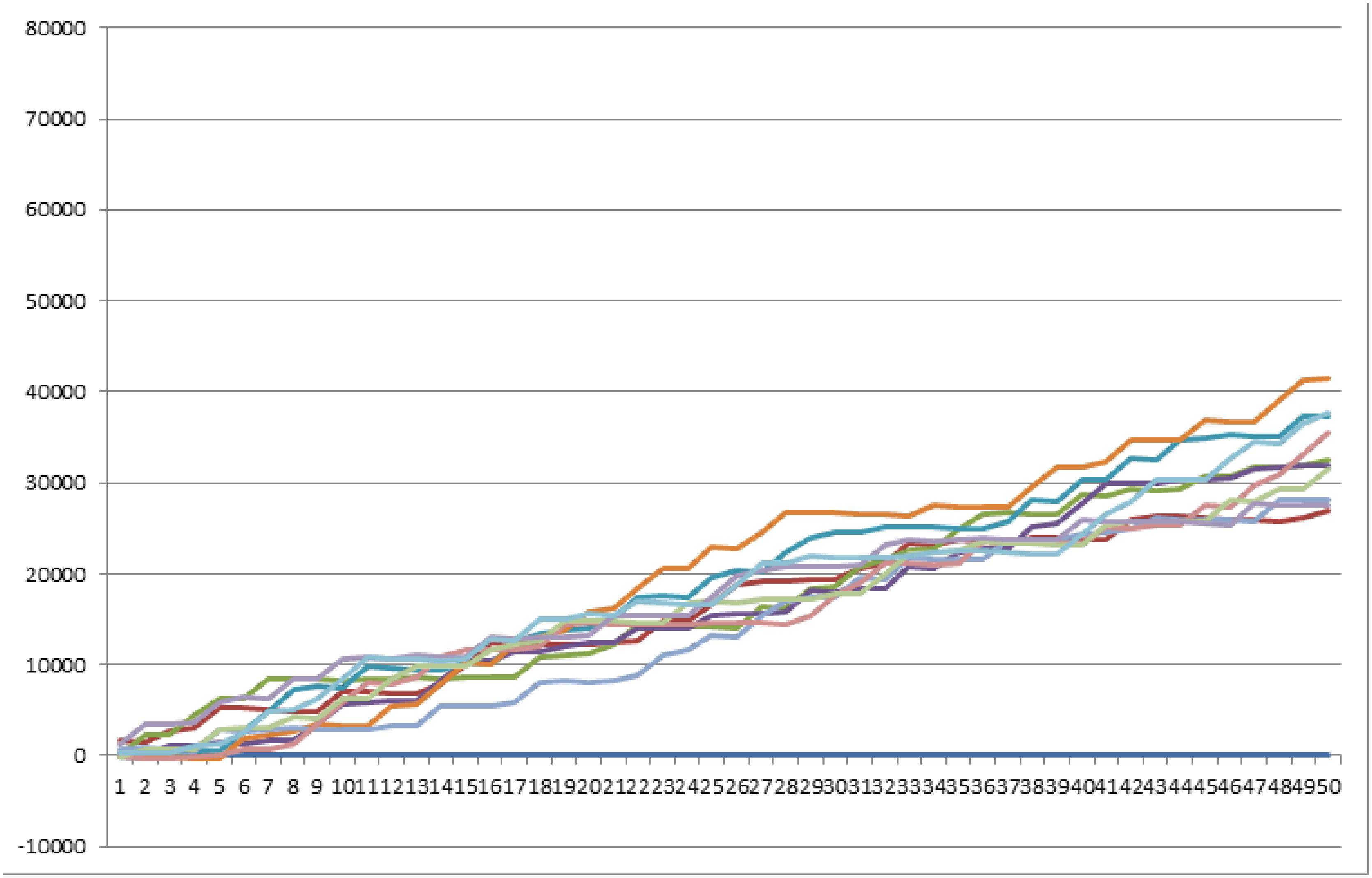
Newsvendor and risk

- Order quantity = 25



Newsvendor and risk

- Order quantity = 25



Newsvendor and risk

- Observations
  - » Smaller order quantities reduces profits, but reduces risk.
  - » The correct order quantity depends on your startup-capital and tolerance for losses.

Lecture outline

- Applications of the newsvendor problem
- The newsvendor problem
- Estimating the distribution and censored demands
- The newsvendor problem and risk
- The newsvendor problem with an unknown distribution

Stochastic gradient algorithm

A stochastic optimization problem involves a problem of the form:

$$\min E\left\{F(x,\omega)\right\}$$

(for the moment, assume the problem is unconstrained).

A stochastic approximation procedure (or stochastic gradient algorithm) is an algorithm of the form:

$$x^n = x^{n-1} - \alpha^n g(x^{n-1},\omega^n)$$

where g() is a stochastic gradient, given by:

$$g(x^{n-1},\omega^n) = \nabla_x F(x^{n-1},\omega^n)$$

Important variation: gradient smoothing:

$$\bar{g}^n = (1-\alpha^{g,n})\bar{g}^{n-1} + \alpha^{g,n} g(x^{n-1},\omega^n)$$

Then use:

$$x^n = x^{n-1} - \alpha^n \bar{g}^n$$

Stochastic gradient algorithm

- Applying this to the newsvendor problem:

We start by writing:

$$F(x,\omega) \quad = c^o[x - D(\omega)]^+ + c^u[D(\omega) - x]^+$$

Note that in our problem, an iteration n corresponds to a time period t, which we indicate as a subscript.

This means that our stochastic gradient is:

$$g(x_{t-1},\omega_t) = \begin{cases} c^o & \text{if } x_{t-1} > D_t(\omega) \\ -c^u & \text{if } x_{t-1} \le D_t(\omega) \end{cases}$$

(Note that we have a tie-breaking problem if $x_{t-1} = D_t(\omega)$. We can break this tie arbitrarily - we only require that $g()$ be a valid subgradient - however, in the presence of censored demands ...)

Stochastic gradient algorithm

- Sample problem:

$$D_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \sim U(0,20) \qquad \text{(Uniform distribution)}$$

Set the underage cost $c^u = 1$ and the overage cost $c^o = 2$.

The optimal order quantity must satisfy $1/(1+2)$ of the demand. Since the demand distribution is uniform, that means the optimal order quantity is:

$$x_t^* = \mu_t + 20 \frac{c^u}{c^o + c^u}$$

If the distribution of D_t is not known, the stochastic gradient algorithm is:

$$g(x_{t-1}, \omega_t) = \begin{cases} 2 & \text{if } x_{t-1} > D_t(\omega) \\ -1 & \text{if } x_{t-1} \le D_t(\omega) \end{cases}$$

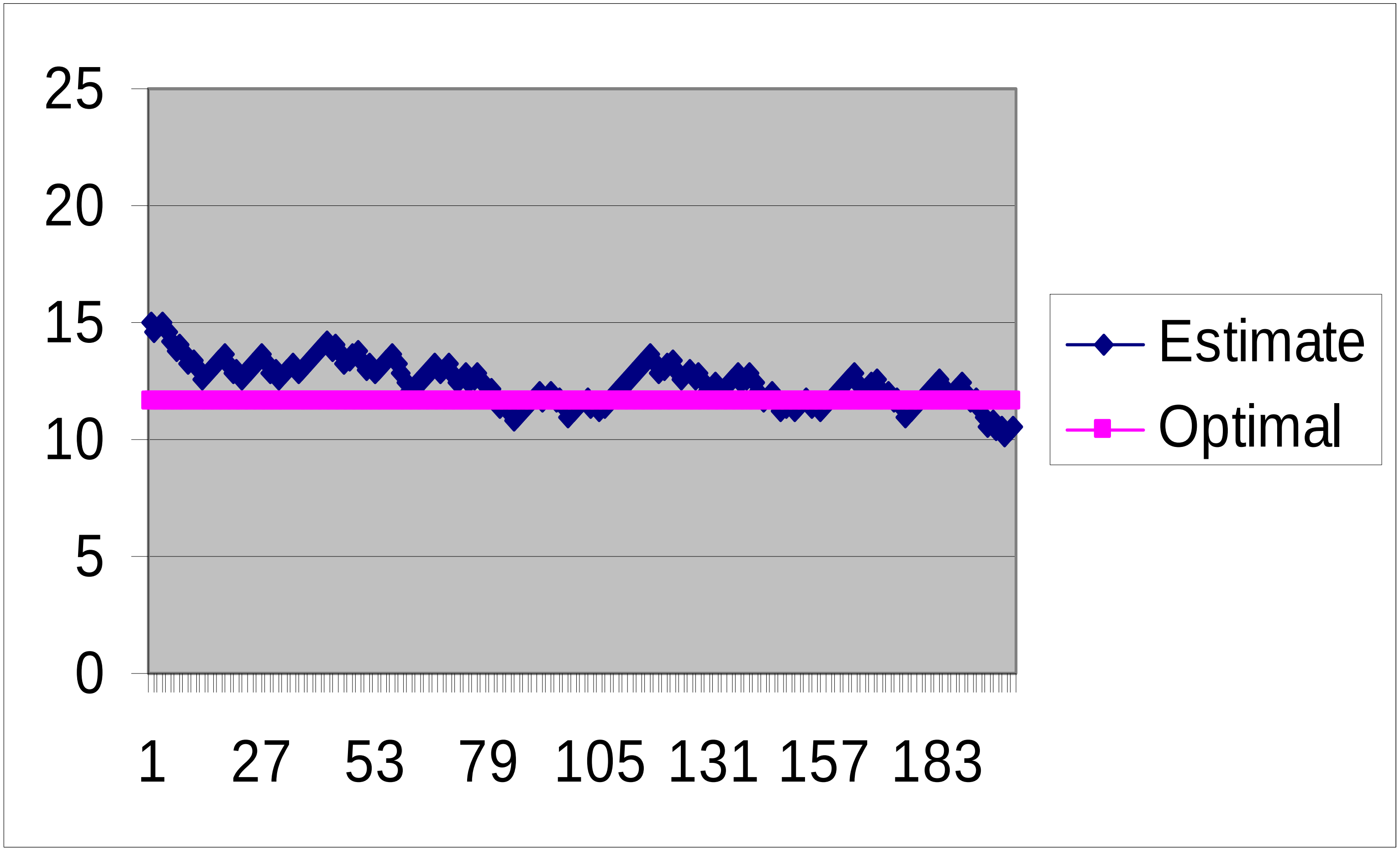
$$x_t = x_{t-1} - \alpha_t g(x_{t-1}, \omega_t)$$

Stochastic gradient algorithm

- Stepsize issues:
  - » As with forecasting, we have to choose stepsizes that handle:
    - Initialization problems
    - Transient data
    - Volatility
  - » But we also have to handle scaling problems:
    - The unit of the gradient is not the same as the unit of the order quantity.
    - Scale the stepsize so that initially, the size of the adjustments being made are roughly 20 – 50 percent of the order quantity, and then let the stepsize decline from there.
    - Use bigger initial stepsizes to increase adjustment from initial conditions; use smaller stepsizes to maintain stability.

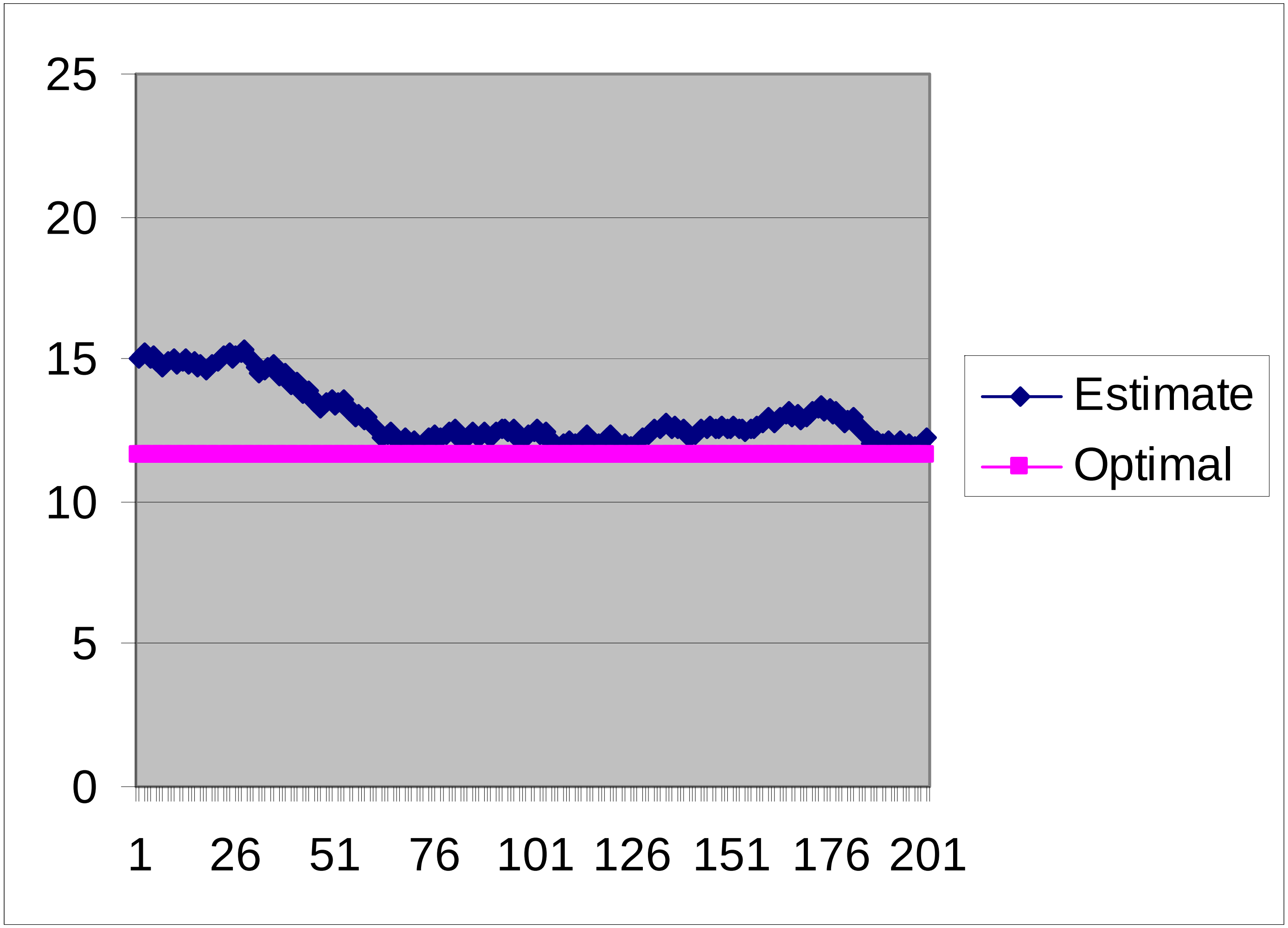
Stochastic gradient algorithm

- Constant stepsize: .2



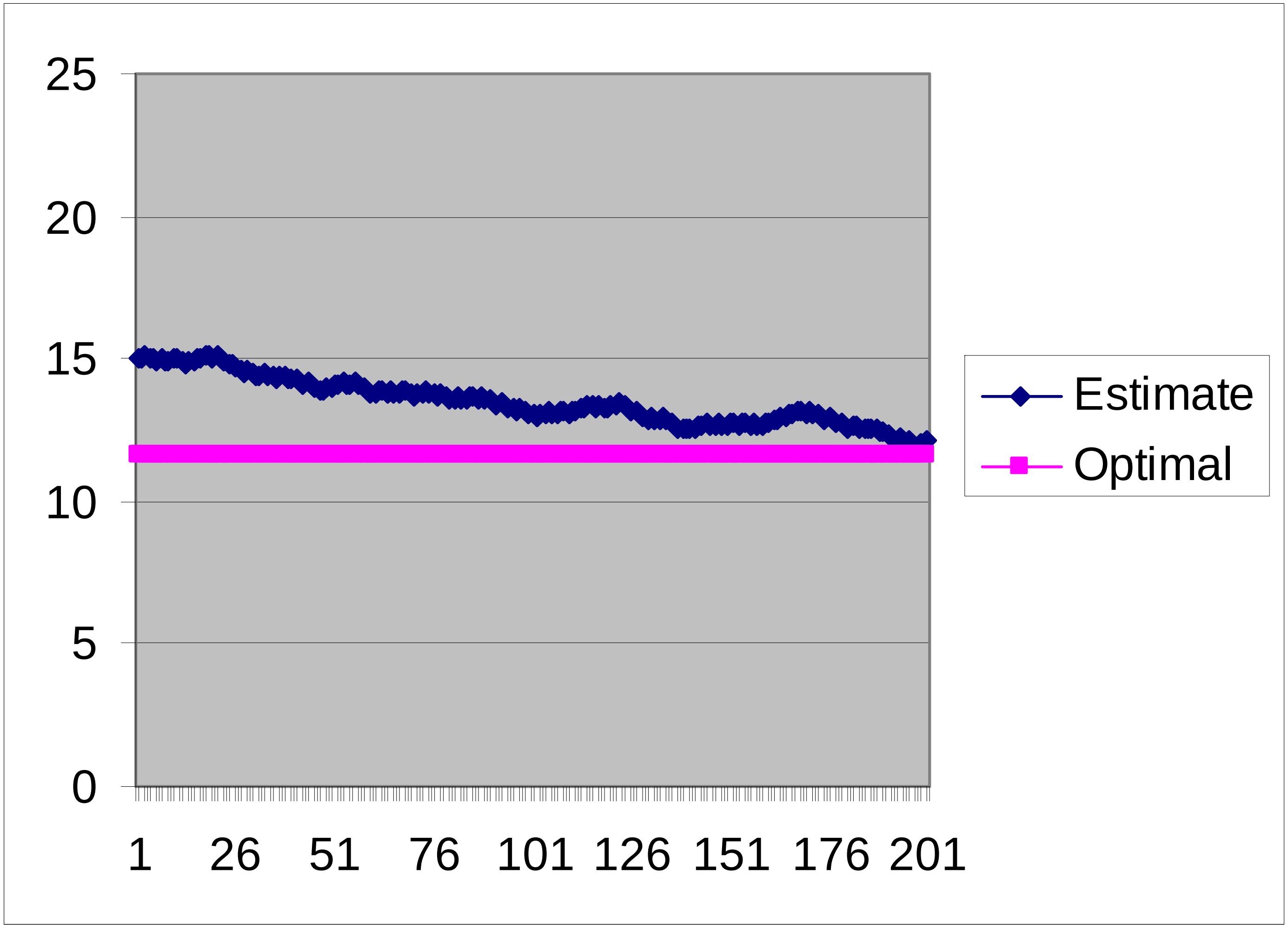
Stochastic gradient algorithm

- Constant stepsize: .1



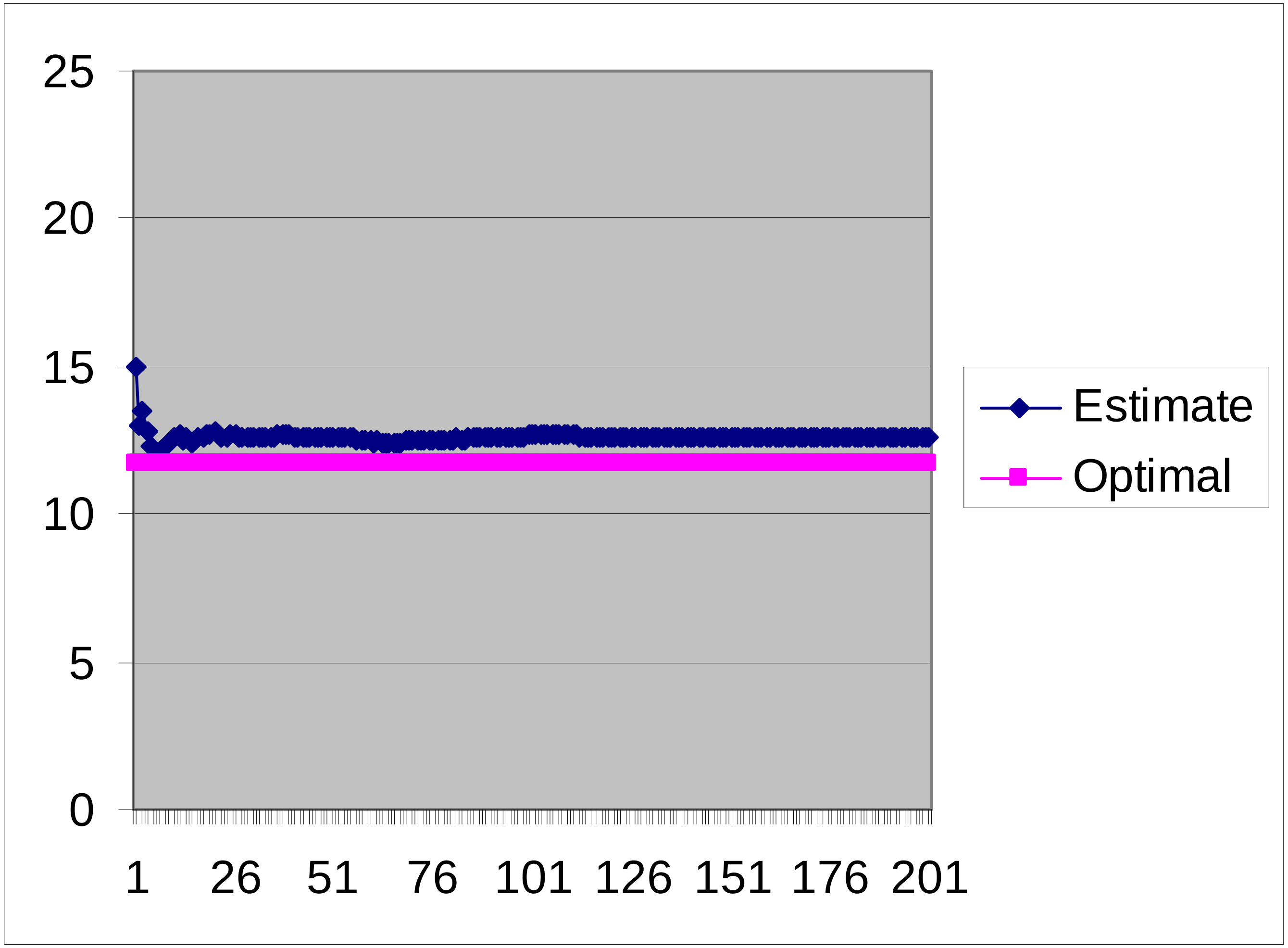
Stochastic gradient algorithm

- Constant stepsize: .05



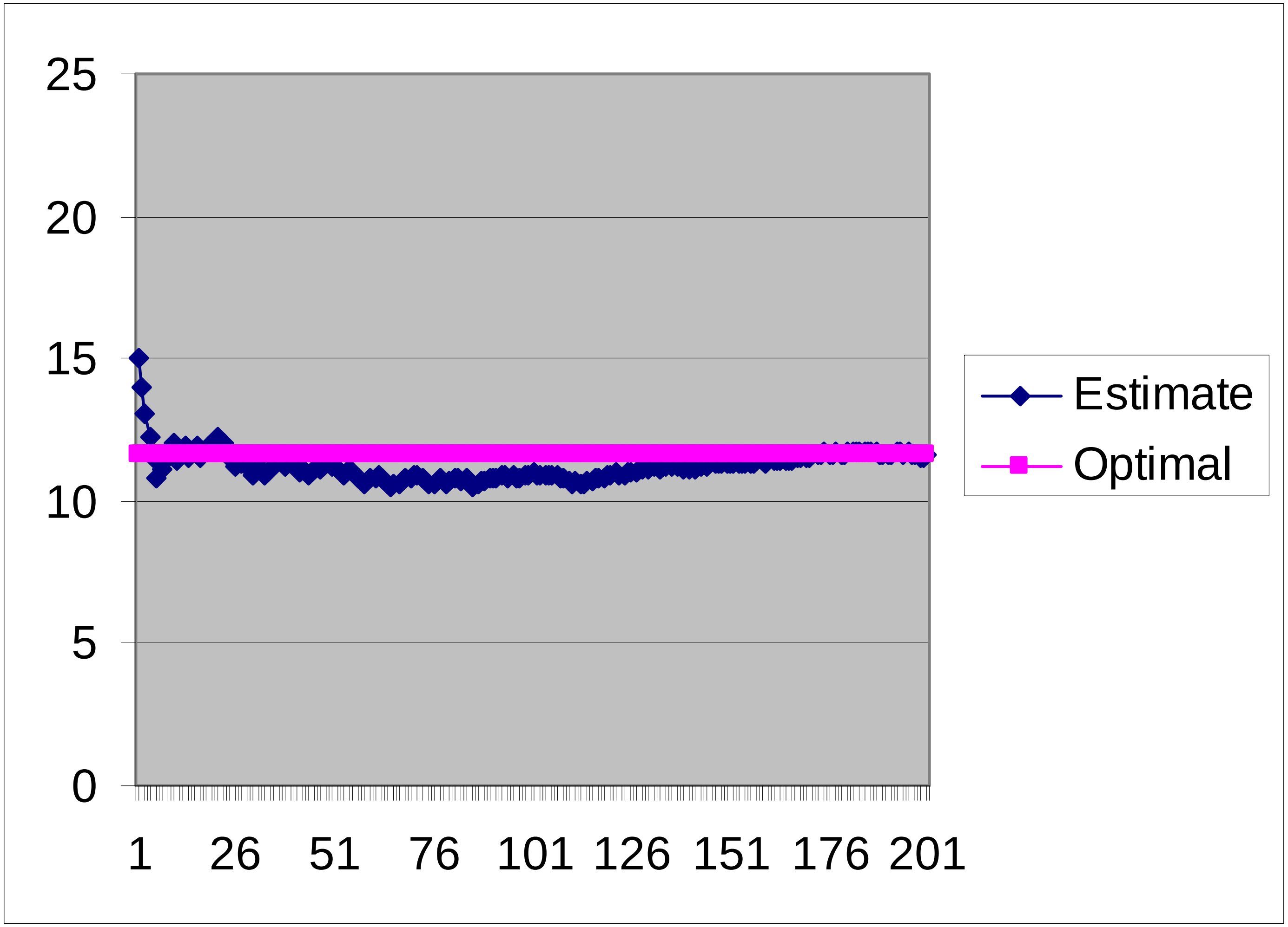
Stochastic gradient algorithm

- Declining stepsize: 1/n (high initial estimate)



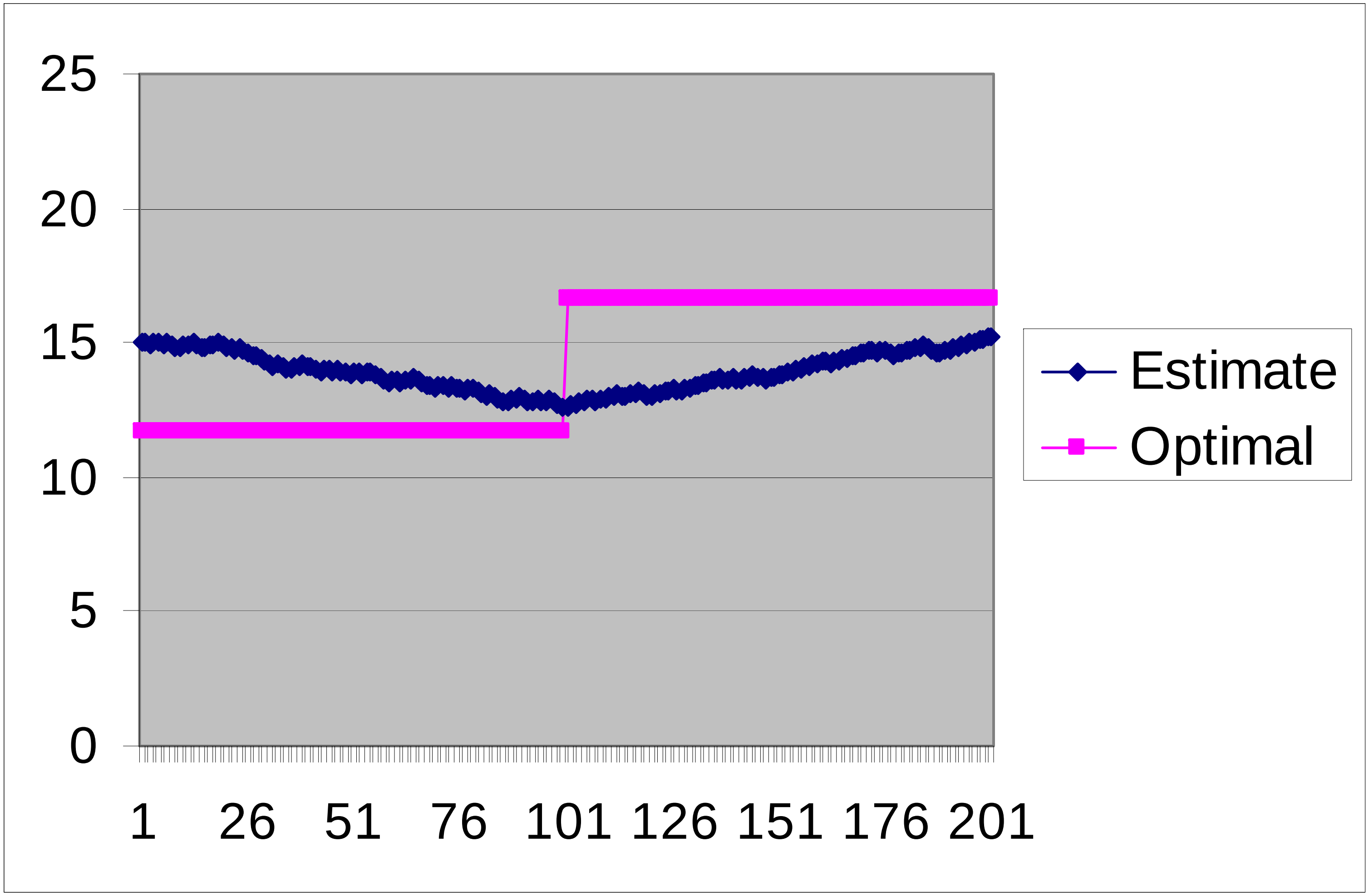
Stochastic gradient algorithm

- Declining stepsize: 5/(10+n)



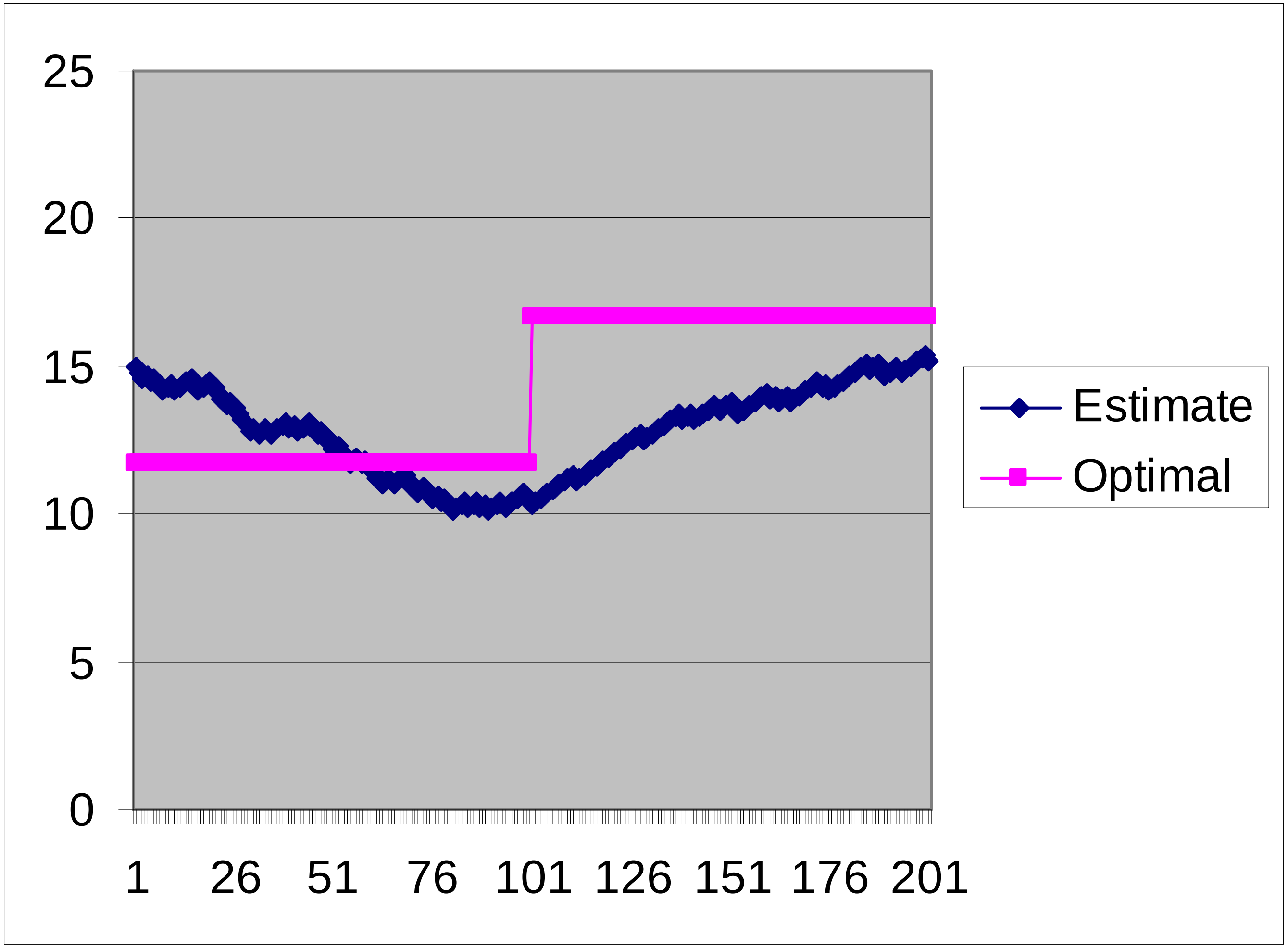
Stochastic gradient algorithm

- The problem with declining stepsizes is when there are shifts: Stepsize = .05



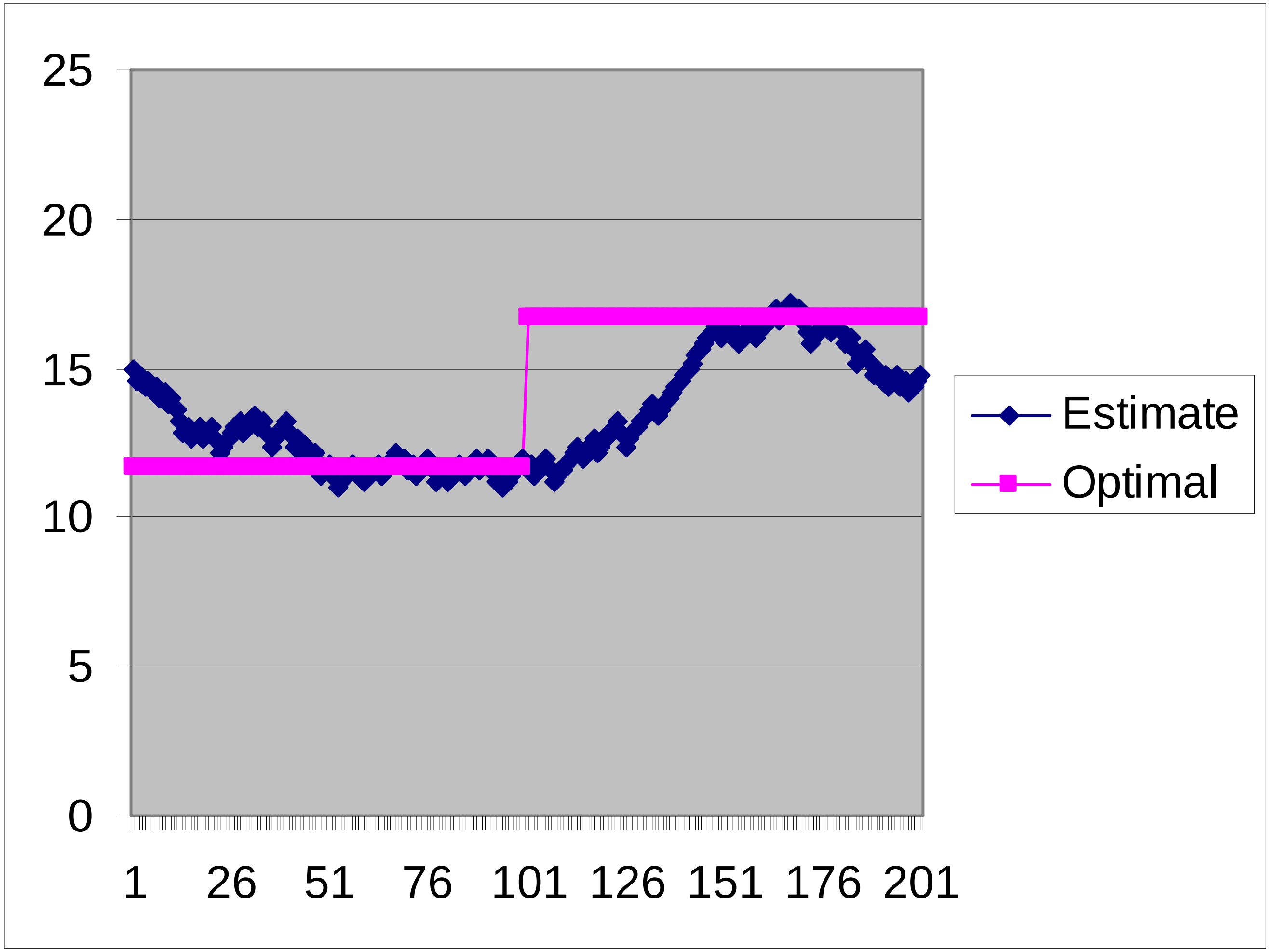
Stochastic gradient algorithm

- The problem with declining stepsizes is when there are shifts: Stepsize = .10



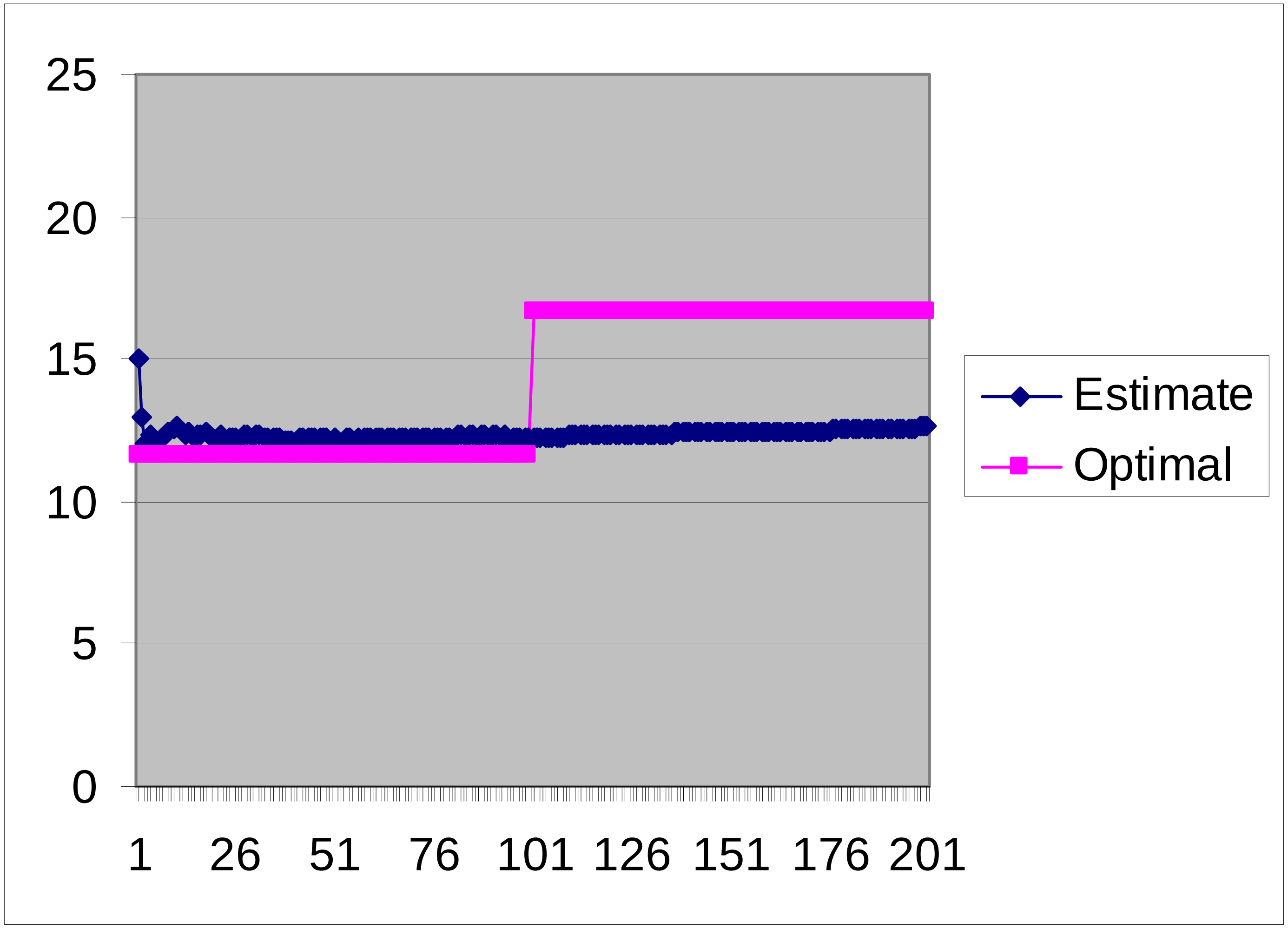
Stochastic gradient algorithm

- The problem with declining stepsizes is when there are shifts: Stepsize = .20



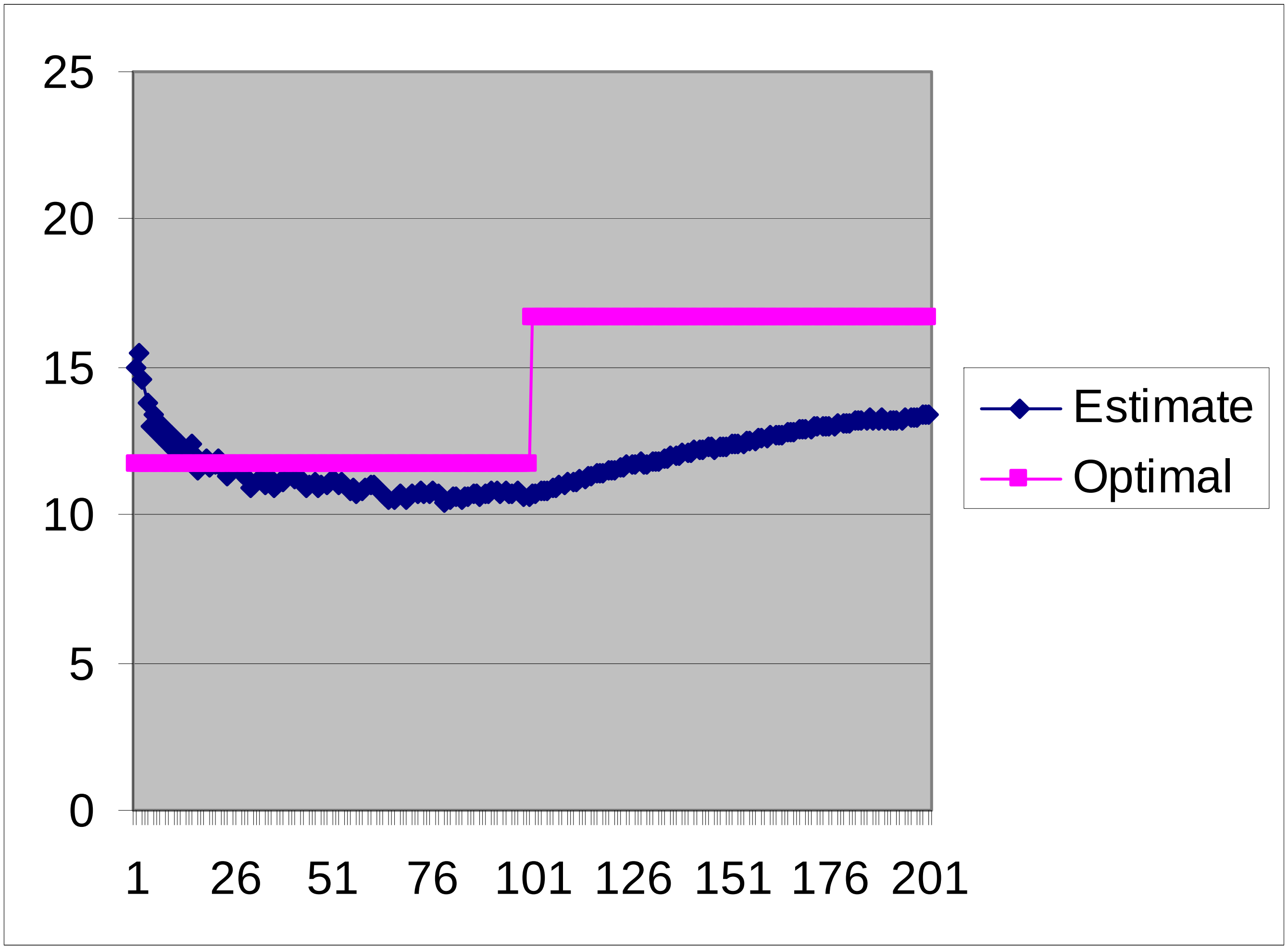
Stochastic gradient algorithm

- The problem with declining stepsizes is when there are shifts: Stepsize = 1/n



Stochastic gradient algorithm

- The problem with declining stepsizes is when there are shifts: Stepsize = 5/(10+n)



Stochastic gradient algorithm

- Some observations:
  - » We never assumed we knew anything about the demand distribution.
  - » If we compute total profits, and compare them to what we would have obtained if we knew the demand distribution, we are almost always within 1.5 percent of the best we could achieve (without the benefit of hindsight).

Stochastic gradient algorithm