Process Management (OPIM 631)

I. Basic Performance Measures

- process – transforms inputs (RM, unserved customers) into outputs (FG, customers served) with resources (machines / workers / capital)

- flow unit – unit of analysis
  - types
    - physical –
    - monetary – facilitates aggregating across different inventory types
  - challenges
    - flow unit breaks up in process
    - multiple types of flow units present (~ product mix)

- capacity – # workers / activity time
  - limit = minimum (bottleneck capacity, demand)
  - maximum rate at which process can generate supply
  - max rate at which process produces during given time
  - maximum amount a resource can produce in a given time
  - depends on – product mix

  - consider – number of resources

- product mix -

- Basic performance measures

  - WIP / Inventory – flow units contained within a process
    - represents a liability for operations
    - to evaluate – look at process over time (rather than at specific time); watch unit through whole process
    - measuring – can measure by tracking inputs and outputs alone
    - ***why hold?*** –
      - i) pipeline ~ due to in process time
      - ii) seasonal ~ build up inventory in advance
      - iii) cycle (EOS in processing batches) – more common if set-up required
      - iv) decoupling / buffer – allows mgmt to run steps independently; reduces effect of variations in flow
      - v) safety – addresses stochastic demand
        - > stochastic demand ~ variation in demand about predicted demand
          - ¬ problem in retailing / finished goods
    - ***why important?*** –
      - costly – obsolescence + perishable + shrinkage + overhead storage costs
        - > individual product's inventory cost = total inventory cost per yr / inventory turns

- ♦ ops sees as a liability
- ♦ hinders process improvement (multiplies the effects of any defects

- Flow time = time in queue + activity time = $T_q + p$
  - time takes flow unit to get through project
  - consider if flow units pass each other during process

- Flow rate (though-put rate) = flow units / time
  - rate at which process delivers output
  - limit = minimum (input rate, demand, bottleneck capacity)
  - most important
  - if inventory is in$ → look at COGS (not sales)

- inventory turns – frequency that inventory turns over during yr
  - = 365 days / flow time (in days)

- **Little's Law** – ***Inv = Avg Flow Rate * Avg Flow Time***

  - always holds BUT deals with avg
  - caveat – not impacted by order

- Consider
  - higher flow rate ~ higher revenues
  - shorter flow time ~ more sales / higher P
  - lower inv ~ lower WC + lower flow time

II. Supply Process

- bottleneck – resource with the smallest capacity
  - limits the capacity of the overall process
  - true bottleneck - has highest ***implied process utilization*** (flow rate / capacity)

- demand constrained – when demand < supply

- supply constrained – when demand > supply (limit is either bottleneck or inputs)

- steps for analyzing a process
  - ***i)*** define process boundaries – part of business will focus on
  - ***ii)*** identify appropriate level of detail
  - ***iii)*** define flow unit
    - consider if focus amounts = input or output
    - convert amounts for consistency by following unit volume / weight throughout process (front to back or vice versa)
    - input maybe better if process itself consumes resources (i.e., more goes in than comes out)
  - ***iv)*** draw process flow chart – consider from flow unit perspective
    - symbols
      - ♦ box – process step
      - ♦ triangle – buffer holding inventory
      - ♦ arrows – route of flow
    - ROT – include only processes / steps that affect process flow or economics

  - ***v)*** determine capacity of each step (flow units / period of time

  - ***vi)*** find Bottleneck – using Little's Law → calculate capacity of each step (~ flow rate per unit of time)
    - consider how long takes to process x units = x / Flow Rate
    - true-bottleneck –
      - ♦ a) find requested capacity – trace capacity from beginning to current step
        - > requested capacity ~ total capacity needed t/b processed per time period
      - ♦ b) determine implied utilization ~ (***capacity requested / available capacity***)

- ♦ c) true bottleneck – has highest implied utilization

    - *if product mix exists* – again compare requested capacity (across products) to available capacity (resources * x / flow rate)

- ➢ ***vi)*** calculate utilization
    - process utilization = Flow Rate / Process Capacity
        - ♦ ~ how much actually produced relative to how much could be produced
        - ♦ underutilization occurs if – Demand < Supply <u>or</u> insufficient Supply <u>or</u> process step has periods of limited availability <u>or</u> non-bottleneck
        - ♦ Bottleneck has highest process utilization

- ➢ ***vii)*** Calculate performance measures

- Remember
    - ➢ always fully utilize the bottleneck step
    - ➢ corp objective = max π <> max utilization

III. Balancing Process Steps

- cycle time = 1 / Flow Rate
    - ➢ time between the process spitting out finished product
    - ➢ ~ every x units of time, a flow unit enters / exits process (w/ x determined by the bottleneck)

- idle time (step i) = cycle time – the activity time at a given step (i)

- span of control = summed activity times of individual activities at a process step

- specialization – reduces span of control
    - ➢ bene = lower activity times
    - ➢ con = lower utilization (but we care about π / not utilization)

- work cell –
    - ➢ will have 100% utilization b/c must also be the bottleneck
    - ➢ but – reduces specialization (worker must master longer span of control)
    - ➢ may lead to longer activity times (due to less specialization)

- goal = max π <> max utilization

- line balancing = avoid mismatch b/w supply and demand of process steps
    - ➢ <u>effect</u>
        - = increase efficiency of process (better utilize resources) + increase capacity of process (reallocating b/w over / under used resources)
        - = reduce direct labor cost
    - ➢ <u>consider</u> – can re-sequence tasks
    - ➢ <u>harder</u> – with increase in specialization (easier with decrease in specialization)

- sequential process steps -
    - ➢ worker paced line – inventory exists b/w process steps
    - ➢ machine paced line – no inventory between process steps

    - ➢ <u>time to process x units</u> =

        - **a) *starting cold*** = time to finish first unit + [(x-1) / Flow Rate]

            - ♦ time to finish first unit – depends
                - > worker paced line – sum of activity times for each step
                - > machine paced line – number of steps * activity time of Bottleneck step

- b) ***continuous flow*** – time to finish first unit + [x / Flow Rate] **why different (p. 176)**

- labor -
  - direct labor content = sum of activity times of labor process steps
    - ignores how process operated (i.e., labor cost w/b different)

  - direct labor cost – s/b higher than labor content
    - = (total wages/ unit of time) / (# product / unit of time)
    - = (labor cost) / process flow rate
  - difference = from under utilization
  - remember –
    - adding more people <> change labor content (may change labor cost if added to non-bottleneck)
    - assumption = labor cost is fixed (but can day be shortened?)

  - average labor utilization -
    - **a)** labor content / (labor content + idle time)
    - **b)** average utilization across all workers = 1/m * (u_1 + . . . + u_m)

- Scaling up to higher volume (p. 183-)

  - balance line – increases utilization (reduces direct labor content)
  - add labor –
    - **i)** replicate balanced line (to new line) – but current line may still not be best way after more workers added
      - avg utilization remains unchanged
    - **ii)** add people to individual process steps
      - capacity = # workers / activity time (with more workers → capacity rises)
    - **iii)** divide work further (increase specialization)
      - **a)** determine targeted throughput / cycle time
      - **b)** how many workers does it take to met target (a)?
        - **1)** break each process steps into individual activities (w/ activity times)
        - **2)** disperse activities evenly across all workers
        - **3)** span of control = summed activity times of individual activities at a process step
      - **c)** consider – greater specialization (less training / quicker learning) BUT labor utilization may go down due to bottleneck

IV. Batching / Flow Interruptions

- batch processing – as opposed to mass production systems

  - generally - uses general purpose technology to produce larger variety of products

  - set up –
    - leads to lower capacity (b/c no production occurs during set-up)
    - ~ interruption in process flow (stealing capacity)
    - only one goal = eliminate / reduce time required for each one
    - capacity w/ set up ~ ***Batch size / [setup time + (batch size * activity time per unit)]*** ~ a function of batch size
      - as increase batch size (spread set up over more units) and approach max = 1 / activity time

    - Single Minute Exchange of Die (SMED) Method –
      - divide all set up tasks
        - internal setup tasks ~ only executed while machine is stopped
        - external set-up tasks ~ c/b done before machine stops
      - bene = if applied to bottleneck (not helpful if improving flow of non-bottleneck)

  - production batch – collection of flow units that are processed before a resource needs to go through another set-up

  - transfer batch – collection of flow units transferred as group from one process step to another

    - may be beneficial to reduce below size of production batch (e.g., to reduce inventories & flow time)

- batch size
    - increasing size → increases capacity (output / time period)
- problems –
    - larger batch size = more inventory
    - more waiting (from flow unit perspective) throughout process
    - customer buying patterns may not coincide with batches
        - solution = Mixed Model production = batch size of 1 (eliminates cycle inventory + perfect align S & D)
- ROT –
    - if set-up occurs at bottleneck → increase batch size (to increase process capacity)
    - if set up at non-bottleneck → decrease batch size (to decrease inventory)
- Choosing a batch size – key = focus on set-up times (not set-up costs)~to capture effect of set-up on process capacity
    - **i)** calculate capacity of each step
    - **ii)** determine optimal batch size – equate Batch step capacity with next lowest capacity step and solve for B
        - optimal batch size – smallest batch size that does not adversely affect process capacity
- Economic Order Quantity (EOQ) Model –
    - always check to see if Bottleneck shifted

- Other Flow interruption issues
    - production run – quantity between two flow interruptions
        - capacity ~ (amt processed b/w stops) / [duration of stop + (time to produce one flow unit * amt processed b/w stops)]
    - machine down-time – step may be limited by down time required in other steps if no inventory buffers exist
        - capacity ~ B / S + B*p (where S is machine downtime)
        - i.e., constraint may be different while running vs. during shutdown
    - Buffers – increase process capacity in face of flow interruptions

- remember
    - capacity at bottleneck = extremely valuable
    - capacity at non-bottleneck = free
    - investment in machinery = sunk (focus on time / not cost in choosing batch size)
    - focus on time of set-up (not cost) b/c this captures the effect on the overall process capacity

V. Variability – Simple Process

- Key point – variability causes flow units to incur wait time even if utilization < 100%

- inter-arrival time ~ time between two consecutive arrivals in a process
  - follows exponential distribution → Prob (A > t) = exp (-1/a)
    - A – random inter arrival time
    - a – avg inter-arrival time

- exponential distribution –
  - includes - customers arriving independently

  - ROT –
    - σ = avg inter-arrival time (a) → $CV_a = 1$
    - memoryless property – number of arrivals in next time slot is independent of when last arrival occurred
  - limits to ROT = if
    - arrivals are scheduled
    - arrivals are batched
    - arrivals are negatively correlated (go to different line b/c see someone just entered)
    - arrivals negatively correlated due to limited population of customers

- Variability
  - effect –
    - ***only relevant in presence of dependence***
    - even with under-utilization (or match of S & D)→ flow units wait
    - averages out over time ***IFF observations are independent of each other*** (i.e., independence b/w S & D)
    - inventory – introduces independence

  - sequential dependence – not possible to start step until prior step finishes
    - critical chain effect – mgrs adjust task durations to include contingency buffer (due to high cost of moving scheduled resources when flow unit not finished in earlier step)

  - source –
    - inflow of units –
      - biggest source in service organizations
    - activity times
      - human operators – better at some things than others
      - customer involvement – level of knowledge / attention varoes
    - resources – random breakdowns
    - routing – randomness in path of flow unit (e.g., hospital)

  - measurement –
    - σ - no good b/c is an absolute number (same unit as mean)
    - coefficient of variation ~ ***σ / mean***
      - measures variability relative
      - CV_p = CV activity time
      - CV_a = CV of inter-arrival time
    - problem ~ seasonality ~ predictable pattern
      - has nothing to do with variability
    - Solution – break process into step / periods during which seasonality does not occur

- waiting time

  - simple process – one buffer (unlimited space) + single activity
    - **i)** test reasonableness of model (HO 7)
      - **a)** presence of dependence
      - **b)** stationary process – any seasonality w/i process step? test by
        - **1)** sort all arrival so increasing in time (label - a1 . . an)

- > **2)** plot graph (y = number customers arriving; x = time)
- > **3)** diagonal line from first to last arrival (if stationary → little deviation b/w graph & line)
- > ***solution to seasonality*** = divide up process into smaller intervals w/ mean for each

♦ **c)** exponential distribution of inter-arrival times?

– **ii)** compute CV_a
  - ♦ a – avg inter-arrival time (~ mean)

– **ii)** calculate capacity ~ ***# resource / p***
  - ♦ p – time to process flow unit

– **iii)** calculate utilization ~ ***flow rate / capacity***

– **iv)** calculate wait time – w/ either
  - ♦ **a)** time flow unit waits in buffer (flow unit perspective) ~ T_q (**measures LR avg wait time**)
    - > T_q = ***Activity time * (u / 1-u) *[(CV_a² + CV_p²)/ 2]*** – i.e. $T_q = \text{Activity time} \cdot (u / 1-u) \cdot [(CV_a^2 + CV_p^2)/2]$
      - ¬ is per resource
      - ¬ is exact so long as inter-arrival times follow exponential distribution
    - > utilization (u) m/b < 100% → if utilization > 100%, then real issue is capacity (not variability)
    - > waiting time grows linearly with waiting time in system
  - ♦ **b)** number units waiting in buffer (system perspective) = ???
    - > Total inventory (I_T) = $I_q + I_p$
    - > I_p – b/w 0 & 1 (flow unit is either in or not)
    - > probability that process is busy working on unit = utilization

– **v)** calculate total process time = $T_q + p$
  - ♦ w/b LR avg – first customers might be less than others

– **vi)** calculate inventory – $I = R * T = 1/a(T_q + p)$
  - ♦ buffered inventory – $I_q = 1/a*T_q$
  - ♦ processing inventory – $I_p = (1-u) * 0 + u*1 = u$
    - > I_p = prob (0 processed) * 0 + prob (1 processed) * 1

– **vii)** account for variance in waiting times –
  - ♦ service level = prob (waiting time ≤ x)
  - ♦ target wait time = $-\ln(1-\text{service level}) * T_q$

- remember
  - ➢ cannot just match avg D to avg S – variability causes wait times to arise
  - ➢ seasonality ~ predictable pattern
    - – has nothing to do with variability
  - ➢ if utilization > 100% → we do not care about variability b/c inventory builds up (real issue is capacity)
  - ➢ increasing utilization → increases wait time
  - ➢ variability impacts quality of service (in addition to wait times)

VI Variability – Multiple Processes

- focus – only utilizations less than 100% (otherwise is a capacity issue)

- utilization = flow rate / capacity
  - ➢ = 1 / inter-arrival time / (number of resources * activity time)
  - ➢ = activity time / (inter arrival time * number of resources)
  - ➢ = p / am

- avg waiting time = $(\text{activity time} / m) * (\text{utilization}^{(2(m+1))^{.5} - 1} / 1-u) * (C_a^2 + C_p^2) / 2$ **is an approximation**
  - ➢ once compute → can compute (with Little's Law) flow units in buffer (I_q)→ compute I_p and I_T

- Inventory in process = m * u (~ number of resources * utilization) – is cumulative across resources

. . . .

VII Matching Supply & Demand (OPIM 632) – Newsvendor Model (One Production Opportunity = Make to Stock)

- key steps for profit maximizing order Q –

- i) gather financial inputs (data)
- ii) generate demand model / forecast (choose distribution function)
  - source – history, opinions, etc
- iii) choose objective (~ maximize Π)
- iv) choose order Q to achieve objective

- Process for Π maximizing order Q
  - if NOT normal distribution
    - 1) find & rank A/F ratios in ascending order
    - 2) calculate empirical distribution [F(A/F) = % = rank / total # AF ratios]
    - 3) match critical ratio to A/F distribution %
    - 4) Q = Forecast * matched A/F ratio

  - if normal distribution
    - 1) Expected A/F ratio = avg of all A/F ratios
    - 2) σ of A/F ratio = σ of all ratios
    - 3) Expected Actual Demand = E(A/F) * Forecast
    - 4) Expected σ = σ of A/F ration * Forecast
    - 5) find Z based on critical ratio
    - 6) Q = E(Actual Demand) + [E(σ) * Z]

- Process
  - **1)** Gather Inputs - Forecast Demand (not sales)
    - **a)** need distribution for every possible demand
      - why? reveals accuracy / variance
      - solution – choose a defined distribution to describe your forecast
      - F(Q) ~ distribution function ~ probability outcome will be value or smaller
      - f(Q) ~ density function ~ probability outcome will be exactly a value

    - **b)** make demand forecast ~ opinions, past data, etc.
      - **i)** assume normal distribution for simplicity
      - **ii)** A/F = actual demand / forecast - focus on ABS errors (not relative errors)
        - Expected Actual Demand = Expected A/F ratio * forecast
        - σ of Actual Demand = σ of A/F Ratio * Forecast
      - **iii)** rank, sort A/F, calc rank / total # different products
      - **iv)** find $\mu = E(Demand_{New}) = avg(A/F) * Forecast_{New}$
      - **v)** find $\sigma = E(\sigma_{New}) = E(\sigma_{A/F}) * Forecast$

    - **c)** goodness of fit test - plot observed distribution function of A/F ratio

  - **2)** Find Order Q -
    - **a)** Profit maximizing Order Q
      - A/F ratio - Observe Empirical Distribution
        - **i)** find critical ratio
        - **ii)** find distribution – observe historical relative forecast errors → A/F ratios for numerous products
        - **iii)** rank products (low to hi) based on AF ratio
        - **iv)** evaluate empirical distribution function F(A/F) for each item
          - F(A/F) = each item's rank / number of items
        - **v)** find F (A/F) ~ critical ratio and identify the appropriate A/F ratio (choose the larger)
        - **vi)** Q = selected A/F ratio * Forecast

      - A/F ratio - Fit Normal Distribution
        - **i)** assume normal distribution for simplicity
        - **ii)** A/F = actual demand / forecast - focus on ABS errors (not relative errors)
        - **iii)** rank, sort A/F, calc rank / total # different products
        - **iv)** find $\mu = E(Demand_{New}) = avg(A/F) * Forecast_{New}$
        - **v)** find $\sigma = E(\sigma_{New}) = E(\sigma_{A/F}) * Forecast$
        - **vi)** find Z stat for critical ratio
        - **vii)** $Q = \mu + \sigma * Z$

- Use Std Normal Distribution
  - **i)** find critical ratio → C_O*F(Q) = C_U*(1-F(Q)) → F(Q) = C_U / (C_O + C_U) [focus is on C_U b/c using distrib funct.]
  - **ii)** find Q that satisfies → **F(Q)** = C_U / (C_O + C_U)
    - **A)** graph
    - **B)** equation for F(Q) ~ we have for Standard Normal distribution (and some others)
      - **1)** find standard normal order Q assuming Z = critical ratio (always pick larger Z)
      - **2)** convert Z from standard normal to normal distribution ($\sigma_{new\ order} = Z * \sigma_{historical}$)
      - **3)** Q = μ + Z σ

- **b)** Minimize Losses -
  - key = order less than π max Q

- **c)** Maximize customer service
  - consider - which measure to use depends on industry (fill rate = avg; in stock = conservative)

  - **i)** Target Fill Rate → Q = μ + σ (Z)
    - **A)** Recall →
      - Fill Rate = Expected Sales / μ
      - Expected Sales = μ – Expected Lost Sales = μ – σ * L(Z)
    - **B)** Combining equations → L(Z) = (μ / σ) (1-Fill Rate)
    - **C)** find Z – by solving for L(Z) and examining L(Z) table
    - **D)** solve for Q

    - **remember – choose lower Z stat b/c fill rate increases if expected loss decreases (190)**
      - this ensures that will achieve at least the target fill rate

  - **ii)** Target In Stock Probability → Q = μ + σ (Z)
    - **A)** find Z where F(Z) = desired in-stock probability
    - **B)** solve for Q

  - **iii)** Minimize Out of Stock Probability -

- remember
  - expected gain from selling 1^{st} unit is much higher than for selling last unit
  - E(Gain) decreases relative to E(Loss) as we order more
  - order more until E(Gain) = E(Loss) of ordering next unit → C_O* F(Q) = C_U * (1-F(Q))

3) Other Performance Measures

- **a)** Expected Lost Sales = σ * L(Z)
  - **i)** find Z (for a given Q) → Z = (Q-μ) / σ (just reversing what did earlier)
    - remember - link b/w Q & Z ~ probability that demand < Q is same as probability that standard normal distribution is less than Z
  - **ii)** find L(Z) → look up Z in Standard Normal Loss Function Table
  - **iii)** multiply - σ * L(Z)

- **b)** Expected Sales = μ – Expected Lost Sales
  - remember –
    - each unit of Demand = either sale or lost sale
    - Expected demand = μ
    - Expected Sales ~ always less than Expected Demand (μ) b/c Expected lost sales will always be positive unless you order an unlimited Q (very expensive)
- **c)** Expected Left Over Inventory = Q – Expected Sales
  - remember
    - E(Left Over Inventory) and E(Lost Sales) can both be positive (b/c dealing w/ expectation)
    - In reality → can only be either left over inventory or lost sales

- **d)** Expected Π = (C_U * Expected Sales) – (C_O * Expected Leftover Inventory)

- remember
  - > C_U = opportunity cost of a lost sale = per sale margin

- **e)** Expected GM % = 1 – (Cost / Expected Revenue) = 1 –Cost / (P*Expected Sales + V * Expected L/O Inv)
  - GM% = (Revenue – Cost) / Revenue = 1 – C / R
    - > never higher than ~ 1 – C / P
  - Revenue = (P * Sales) + (V * left over inventory) (~ where V = salvage value)
  - Cost = c * Q

- **f)** Expected Fill Rate = Expected Sales / Expected Demand (μ)
  - Fill Rate = Sales / Demand (~ % of demand that is satisfied = sales)
  - measures - avg customer service (treats all customers equal)

- **g)** In-Stock Prob = F(Q) (~ distribution function)
  - ~ probability store ends season satisfying all demand
  - ~ probability that Q_D < F(Q)
  - ~ fill rate at end of season (i.e., prob that last customer demand c/b satisfied)
  - measures – customer service conservatively (i.e., b/c evaluates when at its worst)
  - process –
    - > **i)** find Z stat for given order Q → $Z = (Q - \mu) / \sigma$
    - > **ii)** find F(Q) ~ the % for the Z stat

- **h)** Out of Stock Probability = 1-F(Q) (~ 1- in stock probability)

- Remember
  - Newsvendor model requires – p + c + v + demand forecast w/ probability for every possible demand
  - Actual Demand = only historic sales made at regular P (~ discounted sales are not real demand)
  - Prior data (A/F) → provides distribution (standard deviation)
  - Profit Maximizing Order Q <> Expected Demand (μ) → if $C_U > C_O$ → Q > Expected Demand
  - separate forecasting from order Q decision (means two products w/ same mean forecast can have different π max Q)
  - Opportunity Costs ~ just as important as explicit costs

VIII Quick Response w/ Reactive Capacity (Make to Order)

- Mismatch Cost
  - what is it ~ measures cost of supply / demand mismatches ~ the most you are willing to pay for perfect info

  - calculation -
    - **a)** = (C_O * Expected Leftover Inventory) + (C_U * Expected Lost Sales) = Tangible + Intangible Costs
    - **b)** = E(π) with a perfect forecast – E(π) with actual forecast
      - E(π) with a perfect forecast = (p-c) * Expected Demand (μ) = C_U * μ (~ b/c Expected Demand = Q)
      - E(π) with actual forecast = using Newsvendor Model
    - **c)** = Maximum Π – Expected Π (Newsvendor)
      - Maximum Π = (p-c) * μ = C_U * μ
    - **d)** = mismatch tax rate = mismatch cost / unit of Demand = $C_U * [f(Z) / F(Z)] * [\sigma / \mu]$
      - tells the per unit % cost of not having perfect information (also maximum benefit from perfect info)
      - f(Z) / F(Z)
        - > ~ decreases as Q increases (if we graph the ratio for Z)
        - > ~ decreases as critical ratio increases (b/c critical ratio drives Q)
        - > **Key Point – product with lower critical ratio has higher mismatch tax rate**
      - σ / μ = coefficient of variation
        - > ~ good measure of variability of demand
        - > needed b/c σ alone does not capture the relative variability in demand (b/c relative size of μ is not captured)
        - > ROT – low < 0.25; high > 0.75
        - > reflects quality of firms forecasting ability
        - > recall – coefficient of variation w/ exponential distribution = 1 (always) (differs from normal distribution)

- ♦ key points – mismatch cost increases if either
  - > **a)** critical ratio decreases
  - > **b)** demand variability increases (as captured in CoV)

- ➢ evaluation – compare to Expected Demand (μ) (b/c everything else depends on μ)
  - – assume – Q = Π max Q (this minimizes the mismatch cost)
  - – mismatch cost / unit of Demand = C_U * [f(Z) / F(Z)] * [σ / μ]
  - – ordering the π max Q minimizes the mismatch cost
  - – E(Π Newsvendor) = Maximum Π – Mismatch Cost
  - – CU = maximum Π / unit of expected demand
  - – Expected Π / Unit of demand = C_U – C_U * Mismatch Tax Rate

- ➢ why – understand the per unit supply / demand mismatch cost

- Make to order (reduces Mismatch Costs)
  - ➢ pros – reduces cost from lost sales / excess inventory
  - ➢ cons – expensive + time lags are unavoidable + idle capacity from queuing theory (must have idle capacity to eliminate wait time, but isle capacity is supply / demand mismatch)

- Reactive Capacity
  - ➢ what is it – capacity allowing a firm to make one added replenishment during the season
  - ➢ why bother – is intermediate solution less costly than make to order but better than newsvendor
  - ➢ Types
    - – **A)** Unlimited Reactive Capacity –
      - ♦ cons ~ expensive (to mfr who passes cost to retailer - designer) + complex (much more difficult to choose order Q's with multiple replenishments) + does not help if over order
      - ♦ simplifying assumptions – don't stock out before second replenishment + after observe initial sales → can predict Demand perfectly (may be close to true depending on historical results)
      - ♦ Key – use Newsvendor to make initial order
      - ♦ C_O = same as newsvendor
      - ♦ C_U = premium must pay for under ordering in a prior round
      - ♦ Maximum Π = same = (P-C) * μ
      - ♦ Expected Π = Maximum Π – (C_O * Expected Left Over Inventory) – (C_U * Expected Lost Sales)
      - ♦ Expected Replenishment Q = Newsvendor's Expected Lost Sales
      - ♦ Π max initial order Q –
        - > **i)** find C_O & C_U (see above)
        - > **ii)** calculate critical ratio
        - > **iii)** find Z value for critical ratio (Standard Normal Distribution Function F(Q))
        - > **iv)** convert Z value into order quantity for actual demand distribution function → Q = μ + σ(Z)
      - ♦ metrics (measuring value of replenishment) ~ % Π increases + reduction in mismatch cost + Fill Rate increase + reduction in leftover inventory
    - – **B)** Limited Reactive Capacity –
      - ♦ key – postpone Supply until after Demand is known
      - ♦ cons – capacity limit on replenishment + minimum order Q on replenishment
      - ♦ simplifying assumptions – don't stock out before second replenishment + after observe initial sales → can predict Demand perfectly (may be close to true depending on historical results) + order once constraint (to ensure can get above - and can ignore - the minimum order Q in replenishment)
      - ♦ remember
        - > anything in initial order = Newsvendor Optimal Order Q ~ generates Newsvendor Expected Π
        - > anything in replenishment ~ generates max Π
        - > key – choose initial order - use either
          - ¬ **a)** mismatch quantity ratio
            - ▪ = mismatch cost / Q
            - ▪ initial order s/b those items with the lowest ratio

- ¬ **b)** enumerate all possible initial order permutations and compare for max Π
  - > ideal candidate for initial order = low P, low GM, low CoV, high critical ratio

- Remember
  - E(Π Newsvendor) < E(Π reactive capacity) < E (Π perfect info)
    - – consider this if reactive capacity cost difficult to calculate
  - information is valuable only if you are willing and able to react to it
  - products w/ low gross margins tend to have low mismatch quantity ratios
  - products w/ low CoV and high critical ratios have low mismatch tax rates → have low mismatch –quantity ratios

IX. Service Levels & Lead Times (Order Up to Model)

- Situation – obsolescence is not problem + not limited to single production run

- Goal – balance too much inventory w/ too little

- Solution = Order up to Model (~ 1-for-1 ordering policy)

  - Defn
    - – period ~ order interval (assume all are same)
    - – lead time (£) ~ number of periods required to receive orders (always measured in periods ~ matches frequency in which orders can be made / orders received)
      - >
    - – backorders = (negative on-hand inventory)
      - ♦ ~ when demand arises and no units in current inventory
      - ♦ caused by – excess demand over short time (temp supply / demand imbalance)
      - ♦ causes a stock out (see below)
    - – inventory position = on-order inventory + on-hand inventory
      - ♦ on-order inventory ~ ordered in prior periods but not received (may be 0, negative only if returning inv)
      - ♦ on-hand inventory = S – [μ (£+1))
        - > ~ [order up to level - expected avg demand over (lead time + 1)]
        - > ~ [periods units in inventory - backorders]
    - – order up to level (S) ~ max # units willing to have (on-hand + on order) = base stock level = inventory BoP
    - – Poisson distribution – defined by mean (not σ); BUT $\sigma = (\mu)^{0.5}$
    - – stock out –
      - ♦ arises if - (a) out of stock and (b) demand arrives
      - ♦ causes a backorder
    - – out of stock – when have no inventory on-hand
    - – in stock – if inventory available to meet all demand

  - Assumes – demand is random + same demand distribution applies in all periods + replenishment is at beginning of period (before any demand) + order always received in £ periods + no limit to Q + no obsolescence in inventory + all demand is filled eventually + all unmet Demand is backordered + no seasonality in Demand [expected demand (μ) for any period is same as any other]

  - Choosing demand distribution
    - – **a)** covering the relevant period (day, week, mo, etc.) – convert from long to short if necessary (or vice versa w/ reciprocal)
      - ♦ Expected period demand (μ) = expected (monthly) demand / periods per (month)
      - ♦ σ of period demand = σ of (monthly) demand / $[\text{periods per (month)}]^{0.5}$
    - – **b)** with right distribution
      - ♦ Poisson – discrete + only positive
      - ♦ Normal – continuous (more of issue w/ low avg demand) + 50% is negative
      - ♦ if inter-arrival time is exponential distribution → # arrivals in fixed interval is a Poisson distribution

  - Quantity per period (Q) = **S – Inventory position** = last period's demand where
    - ♦ inventory position = on-order inventory + on-hand inventory
    - ♦ S - depends on the target metric . . .

- minimize inventory, maintain target in-stock probability → simply eval in-stock prob for increasing S until > than desired probability

- minimize inventory, maintain target F/R → eval E(F/R) at increasing S until > than desired probability
  - = E(F/R) ≤ E(demand in one period) * (1 – F/R)
  - **BUT** → E(B/O) table depends on (£+1) periods

- minimize inventory holding costs (h)
  -

- maintain minimum service level (w/ minimum F/R)
  -

➢ Performance Measures (Poisson Distribution) -

- In-stock Probability = **Prob [demand over (£+1) periods ≤ S]** = 1- Stock-out Probability
  - Expected demand over (£+1) periods **= λ = (£+1) * μ demand per period**
  - Probability = Poisson (S, expected demand over (£+1) periods, 1) → see table

- Stock-Out Probability = **1-Prob [**λ **≤ S]** = Prob [λ > S]
  - works better if service level target is high (~ stock outs are rare)

- Expected Backorder (at end of any period) = **[λ – S] + sum_{S-1} [prob (λ ≤ (S-1)]**
  - = [λ – S] + sum_{S-1} [Possion($S_{(0 \text{ to } S-1)}$, λ, 1] (~ expected lost sales in NV)

- Expected Fill Rate **= 1 – [E (B/O) / E(Demand in one period)]**
  - ignore S = 0 → F/R of 100% is unreasonable
  - theory ~ # of 1 period customers = μ in one period → # customers not served in period = E(B/O) → [E(B/O) / E(Demand)] = fraction of customers not served → 1-[E(B/O) / E(Demand)] = fraction of customers that are served
  - remember – lead time influences number of customers not served (numerator) / does not influence demand in one period

- Expected Inventory **= S – λ + E(B/O)**
  - cf pg. 31 – fill this in
  - $\text{Expected Inventory}_{BoP}$ = E(Demand in one period) + E(inventory)
  - $\text{Expected Inventory}_{\text{random } t}$ = [E(Demand in one period) / 2] + E(inventory)

- Expected on-order inventory = pipeline inventory **= E (Demand for one period) * £**
  - cf Little's Law ~ Inventory = FR * FT

➢ Performance Measures (Normal Distribution) –

-

- Lead Time Issues
  - ➢ reducing lead time → reduces expected inventory (especially as service requirement increases) + reduces pipeline inventory

- Remember
  - ➢ Order up to is pull system (NV is push system) = better when avg demand is steady
  - ➢ Poisson Distribution – is different for each mean
  - ➢ after place order, inventory position → increases to S