Vorlesung 04: Bestandsmanagement

Wann bestellen und wie viel auf Lager halten bei ungewisser Nachfrage?

Tobias Vlćek

1. Ungewisse Nachfrage

Stellen wir uns vor, wir betreiben einen Online-Shop für ein beliebtes Produkt, zum Beispiel hochwertige Kaffeebohnen. Jeden Tag bestellen Kunden bei uns, aber wir wissen nie genau, wie viele es sein werden. Mal sind es 10 Beutel, mal 30.

Das bringt uns in ein klassisches Dilemma:

- **Zu viel auf Lager?** Wir binden Kapital, das wir anderswo nutzen könnten, und riskieren, dass die Bohnen an Aroma verlieren. Das sind **Lagerhaltungskosten**.
- **Zu wenig auf Lager?** Kunden wollen bestellen, aber das Produkt ist ausverkauft. Wir sind unzufrieden, gehen zur Konkurrenz und kommen vielleicht nie wieder. Das sind **Fehlmengenkosten**.

Im Gegensatz zu den bisherigen Vorlesungen, haben wir dieses mal eine ungewisse Nachfrage! Die zwei zentralen Fragen sind:

- **Wann** sollen wir eine neue Lieferung bei unserem Kaffeeröster bestellen?
- **Wie viel** sollen wir dann bestellen?

2. Kennzahlen im Lager

Um unser Lager gut zu steuern, müssen wir die Bestände präzise erfassen. Es gibt verschiedene Arten von Beständen, die wir unterscheiden müssen.

- **Physischer Bestand (I^P):** Das ist die Menge, die wir tatsächlich in unserem Regal zählen können.
- **Bestellbestand (I^O):** Das ist die Menge, die wir bereits bestellt haben, die aber noch nicht geliefert wurde. Man nennt dies auch "offene Bestellungen".
- **Fehlbestand (I^B):** Die Menge, die von Kunden nachgefragt wurde, die wir aber nicht liefern konnten, weil unser Lager leer war. Diese Kunden warten noch auf ihre Lieferung.

Die wichtigste Kennzahl für unsere Bestellentscheidung ist der **Disponibler Bestand (I^D)**, auch "Bestandsposition" genannt. Er gibt an, wie viel Ware uns *insgesamt* zur Verfügung steht, um zukünftige Nachfragen zu decken.

$$I^D = \text{Physischer Bestand} + \text{Bestellbestand} - \text{Fehlbestand}$$

$$I^D = I^P + I^O - I^B$$

Warum ist diese Zahl so wichtig? Weil sie die *gesamte* Versorgungslage berücksichtigt. Eine Bestellung auszulösen, nur weil der physische Bestand niedrig ist, wäre ein Fehler, wenn bereits eine große Lieferung unterwegs ist.

Beispiel: Eine Woche im Lager

Schauen wir uns an, wie sich diese Zahlen im Zeitverlauf ändern. Wir verwenden eine **(s, q)-Politik**.

- **Bestellpunkt (s):** Wenn der *disponible Bestand* auf **s** oder darunter fällt, wird eine neue Bestellung ausgelöst.
- **Bestellmenge (q):** Es wird immer eine feste Menge von **q** Stück bestellt.
- **Wiederbeschaffungszeit (L):** Es dauert **L** Zeiteinheiten (z.B. Wochen), bis eine Bestellung ankommt.

Szenario:

- $s = 100$, $q = 250$, $L = 3$ Wochen.
- Start in Woche 0 mit $I_0^P = 120$, $I_0^O = 0$, $I_0^B = 0$.
- Daraus folgt: $I_0^D = 120 + 0 - 0 = 120$.

Ablauf innerhalb einer Woche:

1. **Wareneingang:** Zu Beginn der Woche treffen eventuell alte Bestellungen ein.
2. **Nachfrage:** Kunden bestellen, was den Bestand reduziert.
3. **Bestellentscheidung:** Am Ende der Woche prüfen wir den disponiblen Bestand und bestellen bei Bedarf.

Woche	Ereignisse	d_t	I^P	I^O	I^B	I^D	Kommentar
0	**Start**	-	120	0	0	**120**	Anfangszustand. $I^D > s$, keine Bestellung.
1	Nachfrage	30	90	0	0	**90**	I^P sinkt. $I^D = 120 - 30 = 90$. $I^D < s$, also Bestellung!
	Bestellung		90	250	0	**340**	Bestellung über 250 Stück wird aus-

Woche	Ereignisse	d_t	I^P	I^O	I^B	I^D	Kommentar
							gelöst. I^O steigt. I^D steigt *sofort* um 250.
2	Nachfrage	40	50	250	0	**300**	I^P sinkt. $I^D = 340 - 40 = 300$. $I^D > s$, keine neue Bestellung.
3	Nachfrage	50	0	250	0	**250**	I^P sinkt auf 0. $I^D = 300 - 50 = 250$. $I^D > s$, keine neue Bestellung.
4	**Wareneingang**		250	0	0	**250**	Die Bestellung aus Woche 1 (vor $L = 3$ Wochen) trifft ein!
	Nachfrage	45	205	0	0	**205**	I^P sinkt. $I^D = 250 - 45 = 205$. $I^D > s$, keine neue Bestellung.

3. Servicegrade

Der α-Servicegrad (Zyklus-Servicegrad)

> **Frage:** “Wie hoch ist die Wahrscheinlichkeit, dass wir in einem Bestellzyklus *keinen* Fehlbestand haben?”

Ein α-Servicegrad von 99% bedeutet, dass in 99 von 100 Bestellzyklen (die Zeit zwischen zwei Bestellungen) die Nachfrage vollständig aus dem Lager bedient werden kann. Er fokussiert auf das *Ereignis* eines Fehlbestands.

Formelhaft ausgedrückt ist der α-Servicegrad[1] die Wahrscheinlichkeit, dass die Nachfrage im Risikozeitraum (Y_L) den Bestellpunkt (s) nicht übersteigt:

$$\alpha = P(Y_L \leq s)$$

Der β-Servicegrad (Mengen-Servicegrad oder “Fill Rate”)

> **Frage:** “Welchen prozentualen Anteil der gesamten Nachfrage können wir direkt aus dem Lager bedienen?”

Ein β-Servicegrad von 99% bedeutet, dass 99 von 100 nachgefragten Einheiten sofort geliefert werden können. Er fokussiert auf die *Menge* der nicht gelieferten Einheiten.

Er wird berechnet, indem man die erwartete Fehlmenge pro Zyklus ($E(B)$) ins Verhältnis zur Bestellmenge (q) setzt:

$$\beta = 1 - \frac{E(B)}{q}$$

Der γ-Servicegrad (Zeitbezogener Servicegrad)

> **Frage:** “Wie viel Prozent der Zeit ist unser Produkt tatsächlich auf Lager und verfügbar?”

Dieser Servicegrad betrachtet das Problem aus einer zeitlichen Perspektive. Ein γ-Servicegrad von 99% bedeutet, dass an 99 von 100 Tagen ein Kunde, der unseren Shop besucht, das Produkt in den Warenkorb legen könnte. Er fokussiert sich auf die *permanente Verfügbarkeit.*

Konzeptionell lässt er sich so darstellen:

$$\gamma \approx \frac{\text{Zeit mit positivem Lagerbestand}}{\text{Gesamtzeit}}$$

Die exakte Berechnung ist aufwändiger, aber das Ziel ist einfach die Minimierung der Zeitfenster, in denen Kunden ein leeres Regal vorfinden.

Zusammenfassend:

- $\boldsymbol{\alpha}$: Vermeiden des *Ereignisses* “Fehlbestand”.

[1] Bei einer Normalverteilung entspricht dies genau dem Wert, den wir in der z-Tabelle suchen: $\alpha = \Phi(z)$.

- β: Minimieren der *Menge* nicht gelieferter Einheiten.
- γ: Maximieren der *Zeitspanne* der Lieferfähigkeit.

4. Die Nachfrage im Risikozeitraum

Der kritischste Moment für jedes Lager ist die **Wiederbeschaffungszeit (WBZ)**, auch **Risikozeitraum** genannt. Das ist die Zeitspanne L zwischen dem Absenden einer Bestellung und dem Eintreffen der Ware. In dieser Zeit können wir nicht auf unvorhergesehen hohe Nachfrage reagieren – unser Lager ist auf sich allein gestellt.

Um unseren Bestellpunkt s festzulegen, müssen wir die Nachfrage während dieses Risikozeitraums (Y_L) modellieren.

Fall 1: Nachfrage ist normalverteilt

Sehr oft kann die Summe vieler kleiner, zufälliger Nachfragen durch eine Normalverteilung (Glockenkurve) angenähert werden.

Angenommen, die wöchentliche Nachfrage hat einen Mittelwert μ_d und eine Standardabweichung σ_d. Für eine WBZ von L Wochen gilt dann für die Gesamtnachfrage Y_L:

- **Erwartete Nachfrage während der WBZ:** $\mu_L = L \cdot \mu_d$
- **Standardabweichung der Nachfrage während der WBZ:** $\sigma_L = \sqrt{L} \cdot \sigma_d$

Beispiel:

- Wöchentliche Nachfrage $\mu_d = 50$, $\sigma_d = 15$.
- WBZ $L = 4$ Wochen.

Daraus folgt:

- $\mu_L = 4 \cdot 50 = 200$ Stück
- $\sigma_L = \sqrt{4} \cdot 15 = 2 \cdot 15 = 30$ Stück

Wir erwarten also, in den 4 Wochen 200 Stück zu verkaufen, mit einer Unsicherheit von $\sigma_L = 30$.

Fall 2: Nachfrage ist diskret (Faltung)

Manchmal ist die Nachfrage klein und diskret. Zum Beispiel verkauft ein Kiosk pro Tag 0, 1 oder 2 Flaschen einer speziellen Limonade. Wie ermittelt man die Nachfrageverteilung für eine WBZ von 2 Tagen? Wir müssen alle Möglichkeiten kombinieren. Diesen Prozess nennt man **Faltung**.

Beispiel:

- Tägliche Nachfrage D: $P(D = 0) = 0.2$, $P(D = 1) = 0.6$, $P(D = 2) = 0.2$.
- WBZ $L = 2$ Tage.

Wir wollen die Verteilung für $Y_2 = D_1 + D_2$ finden, wobei D_1 die Nachfrage an Tag 1 und D_2 die an Tag 2 ist.

- $\boldsymbol{P(Y_2 = 0)}$**:** Geht nur mit (0, 0) $\Longrightarrow 0.2 \cdot 0.2 = 0.04$
- $\boldsymbol{P(Y_2 = 1)}$**:** Geht mit (0, 1) und (1, 0) $\Longrightarrow (0.2 \cdot 0.6) + (0.6 \cdot 0.2) = 0.24$
- $\boldsymbol{P(Y_2 = 2)}$**:** Geht mit (0, 2), (1, 1), (2, 0) $\Longrightarrow (0.2 \cdot 0.2) + (0.6 \cdot 0.6) + (0.2 \cdot 0.2) = 0.44$
- $\boldsymbol{P(Y_2 = 3)}$**:** Geht mit (1, 2) und (2, 1) $\Longrightarrow (0.6 \cdot 0.2) + (0.2 \cdot 0.6) = 0.24$

- $\boldsymbol{P(Y_2 = 4)}$**:** Geht nur mit (2, 2) $\Longrightarrow 0.2 \cdot 0.2 = 0.04$

Die Verteilung für die Gesamtnachfrage über 2 Tage ist also:

Y_2	0	1	2	3	4
P(Y_2)	0.04	0.24	0.44	0.24	0.04

5. Den Bestellpunkt s berechnen

Der Bestellpunkt s muss so hoch sein, dass er die Nachfrage während der Wiederbeschaffungszeit mit hoher Wahrscheinlichkeit decken kann. Er besteht aus zwei Teilen:

$$s = (\text{Erwartete Nachfrage während WBZ}) + (\text{Sicherheitspuffer})$$

$$s = \mu_L + \text{Sicherheitsbestand (SS)}$$

Der **Sicherheitsbestand (SS)** ist der zusätzliche Vorrat, den wir nur für den Fall halten, dass die Nachfrage höher als erwartet ausfällt. Seine Höhe hängt direkt von unserem gewünschten Servicegrad ab.

Sicherheitsbestand bei Normalverteilung

Bei normalverteilter Nachfrage lässt sich der Bestellpunkt direkt mit einer Formel berechnen, die den Sicherheitsbestand bereits enthält:

$$s = \mu_L + z \cdot \sigma_L$$

Hierbei ist:

- μ_L der Erwartungswert und σ_L die Standardabweichung der Nachfrage während der WBZ.
- **z** ist der **Sicherheitsfaktor**. Er ist ein Wert aus der Standardnormalverteilung, der direkt von unserem gewünschten α-Servicegrad abhängt.
- Der Teil $z \cdot \sigma_L$ ist der **Sicherheitsbestand (SS)**.

Je höher der Servicegrad, desto größer der z-Wert und desto mehr Sicherheitsbestand halten wir.[2]

α-Servicegrad	z-Wert	Bedeutung
50%	0.000	Exakt der Erwartungswert (kein Puffer).
55%	0.126	0.126 Standardabweichungen als Puffer.
60%	0.253	0.253 Standardabweichungen als Puffer.
65%	0.385	0.385 Standardabweichungen als Puffer.
70%	0.524	0.524 Standardabweichungen als Puffer.
75%	0.674	0.674 Standardabweichungen als Puffer.
80%	0.842	0.842 Standardabweichungen als Puffer.
85%	1.036	1.036 Standardabweichungen als Puffer.

[2]Den z-Wert finden wir in einer z-Tabelle oder berechnen ihn mit einer Funktion wie `NORM.S.INV()` in Excel oder `norm.ppf()` in Python.

α-Servicegrad	z-Wert	Bedeutung
90%	1.282	1.282 Standardabweichungen als Puffer.
95%	1.645	1.645 Standardabweichungen als Puffer.
99%	2.326	2.326 Standardabweichungen als Puffer.

Beispiel (Fortsetzung von Fall 1):

- $\mu_L = 200$, $\sigma_L = 30$.
- Wir streben einen α-Servicegrad von 99% an.

1. **z-Wert finden:** Für $\alpha = 99\%$ ist $z \approx 2.326$.
2. **Sicherheitsbestand berechnen:** $SS = 2.326 \cdot 30 \approx 69.78 \approx 70$ Stück.
3. **Bestellpunkt berechnen:** $s = \mu_L + SS = 200 + 70 = 270$ Stück.

Wir würden also immer dann eine neue Bestellung auslösen, wenn unser verfügbarer Bestand auf 270 Stück sinkt.

6. Wie groß ist die Fehlmenge?

Wir wissen nun, *wie wahrscheinlich* ein Fehlbestand ist (α-Servicegrad).

- Aber *wie viele Einheiten fehlen uns im Durchschnitt*, wenn es zu einem Fehlbestand kommt?
- Das ist die **erwartete Fehlmenge pro Zyklus**, $E(B)$.

Erwartete Fehlmenge bei Normalverteilung

Hier können wir die sogenannte **Einheiten-Verlustfunktion** $G_u(z)$ verwenden:

$$E(B) = \sigma_L \cdot G_u(z)$$

Der Wert für $G_u(z)$ hängt nur vom Sicherheitsfaktor z ab.[3]

z-Wert	$G_u(z)$
1.282	0.047
1.645	0.021
2.326	0.003

Beispiel (Fortsetzung von Fall 1): $z = 2.326$, $\sigma_L = 30$.

- $E(B) = 30 \cdot G_u(2.326) = 30 \cdot 0.003 = 0.09$ Stück.
- Da unser Servicegrad hier sehr hoch ist, erwarten wir im Durchschnitt (über viele Zyklen) nur eine kleine Fehlmenge von 0.09 Stück pro Zyklus.

Erwartete Fehlmenge bei diskreter Verteilung

Hier berechnen wir den Erwartungswert direkt, indem wir alle möglichen Fehlmengen mit ihrer Wahrscheinlichkeit gewichten:

$$E(B) = \sum_{\text{alle } y} \max(0, y - s) \cdot P(Y_L = y)$$

[3] Der Wert für $G_u(z)$ kann aus Tabellen abgelesen oder berechnet werden ($\phi(z) - z(1 - \Phi(z))$).

Beispiel (Limonaden-Kiosk, unser Beispiel aus Fall 2):

- Bestellpunkt $s = 3$.
- Die Nachfrageverteilung $P(Y_2)$ kennen wir von oben.

Das bedeutet, dass wir im Durchschnitt 0.04 Flaschen pro Zyklus verlieren.

- Eine Fehlmenge tritt nur auf, wenn die Nachfrage $Y_2 > 3$ ist. Das ist nur bei $Y_2 = 4$ der Fall.
- Fehlmenge bei $Y_2 = 4$ ist $4 - s = 4 - 3 = 1$ Flasche.[4]
- $P(Y_2 = 4) = 0.04$.
- $E(B) = (1 \quad \text{Flasche}) \cdot 0.04 = 0.04$ Flaschen.

Den β-Servicegrad berechnen

Mit unserem berechneten Wert für die erwartete Fehlmenge $E(B)$ und der in Abschnitt 3 vorgestellten Formel können wir nun den β-Servicegrad (Fill Rate) für eine gegebene Bestellmenge q ausrechnen.

Beispiel (Smartphone): Wir haben $E(B) = 0.09$ berechnet, die feste Bestellmenge sei $q = 500$.

$\beta = 1 - \frac{E(B)}{q} = 1 - \frac{0.09}{500} = 1 - 0.00018 = 0.99982$

- Der β-Servicegrad ist 99.982%.
- Das ist extrem hoch und zeigt, dass bei einer so großen Bestellmenge die wenigen erwarteten Fehlmengen kaum ins Gewicht fallen.

7. Erweiterte Service-Perspektiven

Manchmal ist es nicht nur wichtig, *ob* ein Fehlbestand auftritt, sondern auch, *wen* er betrifft.

Kundenpriorisierung

Stellen wir uns vor, wir beliefern nicht nur Privatkunden, sondern auch ein Luxus-Café, das ein wichtiger Stammkunde ist. Ein Fehlbestand bei diesem Kunden wäre weitaus schlimmer als bei einem einmaligen Online-Besteller. Hier kann man über **Kundenklassen** nachdenken.

Man kann den Bestand so steuern, dass wichtige Kunden bevorzugt behandelt werden. Zwei einfache Strategien sind:

1. **Getrennte Bestände:** Wir halten einen kleinen, separaten "Notfall-Bestand" nur für das Café. Das ist sicher, aber ineffizient, weil dieser Bestand die meiste Zeit nur herumliegt.
2. **Rationierung:** Eine intelligentere Methode. Wir haben einen gemeinsamen Bestand. Fällt dieser aber unter ein kritisches Niveau (z.B. 20 Stück), werden ab diesem Moment nur noch Bestellungen vom Luxus-Café angenommen. Alle anderen Kunden müssen warten. So schützt man seine wichtigsten Partner, ohne permanent Ware blockieren zu müssen.

Die Kundenwartezeit: Reichweite und Lieferunfähigkeit

Neben den Servicegraden können wir auch fragen: Was bedeutet ein bestimmter Sicherheitsbestand für die Wartezeit der Kunden?

[4] Wir haben hier die Nachfrageverteilung $P(Y_2)$ verwendet, um die Wahrscheinlichkeit für eine Fehlmenge zu berechnen.

- **Reichweite:** Stellt die Frage: "Wie viele Perioden (z.B. Wochen) reicht unser Lagerbestand noch, ab dem Moment, in dem wir nachbestellen?" Eine einfache Näherungsformel für die *erwartete Reichweite* lautet:

  $$\text{Erwartete Reichweite} \approx \frac{s}{\mu_d}$$

  wobei s der Bestellpunkt und μ_d die durchschnittliche Nachfrage *pro Periode* ist. Bei einem Bestellpunkt von $s = 270$ und einer wöchentlichen Nachfrage von $\mu_d = 50$ wäre die erwartete Reichweite also $270/50 = 5.4$ Wochen.[5]

- **Lieferunfähigkeitsdauer:** Dies ist die durchschnittliche Zeit, in der wir *keine Ware* haben. Eine Lieferunfähigkeit entsteht, wenn die Wiederbeschaffungszeit länger ist als die Reichweite des Bestands.

  $$\text{Erwartete Lieferunfähigkeitsdauer} \approx \max(0, L - \text{Erwartete Reichweite})$$

  wobei L die Wiederbeschaffungszeit ist. Nehmen wir an, wir würden einen niedrigeren Servicegrad von 85% anstreben ($\alpha = 0.85$, $z = 1.036$). Der neue Bestellpunkt wäre $s = 200 + 1.036 \cdot 30 \approx 231$. Die erwartete Reichweite wäre dann $231/50 \approx 4.62$ Wochen. Mit einer Wiederbeschaffungszeit von $L = 4$ Wochen ergibt sich:

  $$\text{Erwartete Lieferunfähigkeitsdauer} \approx \max(0, 4 - 4.62) = 0$$

  Selbst bei einem 85%-Servicegrad erwarten wir nach dieser Näherung keine Lieferunfähigkeit. Würde unsere Wiederbeschaffungszeit jedoch auf $L = 5$ Wochen steigen, wäre die Situation anders:

  $$\text{Erwartete Lieferunfähigkeitsdauer} \approx \max(0, 5 - 4.62) = 0.38 \quad \text{Wochen}$$

  Wir müssten also mit einer durchschnittlichen Lieferunfähigkeit von fast einer halben Woche rechnen.[6]

Beide Kennzahlen hängen direkt von unserem Sicherheitsbestand ab. Ein höherer Sicherheitsbestand (und damit ein höherer Bestellpunkt s) erhöht die Reichweite und senkt die Dauer der Lieferunfähigkeit. Diese Kennzahlen helfen uns, die Auswirkungen unserer Bestandsentscheidungen direkt in "Kundenwartezeit" zu übersetzen.

[5] Dies ist eine **Näherung**, weil die tatsächliche Reichweite von der zufälligen Abfolge der Nachfragen abhängt. Die Formel gibt uns aber einen sehr guten und schnellen Schätzwert.

[6] Diese Formel ist ebenfalls eine **Näherung**, da sie auf der angenäherten Reichweite basiert.