Produkt og marked - matematiske og statistiske metoder

Rasmus Waagepetersen
Institut for Matematiske Fag Aalborg Universitet

February 10, 2022

Kursusindhold:

- Sandsynlighedsregning og lagerstyring
- Normalfordelingen og Monte Carlo-metoder

Smagsprøve på fundamentale matematiske metoder af relevans for jeres videre uddannelse og virke

Hver undervisningsgang tager udgangspunkt i et konkret eksempel på en anvendelse af de gennemgåede matematiske metoder

Lagerstyring

Hver måned leveres et antal møtrikker L.

Forbrug den $i$te dag i måneden: X_i (tilfældigt antal)

Beholdning ved månedens indgang: $B_0 = L$

Beholdning efter t dage:
$B_t = B_0 - X_1 - X_2 - \cdots - X_t = B_0 - \sum_{i=1}^{t} X_i$.

Forventet værdi af B_t ? Sandsynlighed for at $B_t < 0$? Hvor stor skal L være ? $\Rightarrow$ brug for at regne på tilfældige/stokastiske variable.

Metoder:

1. teoretiske resultater
2. Monte Carlo metoder (computer-simulation)

Forventet værdi/middelværdi

Gentager et eksperiment igen og igen (f.eks. kast med terning)

X_i: tilfældig værdi af $i$te eksperiment

Gennemsnit (repræsentativ værdi) af n gentagelser:

$$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Antag X_i kun antager værdierne $1, 2, \ldots, M$.

$$\bar{X} = \frac{1}{n}\sum_{m=1}^{M} mA_m = \sum_{m=1}^{M} m\frac{A_m}{n}$$

hvor A_m er antal gange værdien m $(1, 2, \ldots, M-1$ eller $M)$ kom ud.

Hvis $n \to \infty$: A_m/n vil tilnærme sig sandsynligheden $P(X_i = m)$ for værdien m

Forventet værdi - fortsat

Definition den forventede værdi (middelværdien) for en stokastisk variabel X med mulige værdier $1, \ldots, M$ er

$$\mathbb{E}X = \sum_{m=1}^{M} m p_m \text{ hvor } p_m = P(X = m)$$

$P(X = m)$ er sandsynligheden for, at X tager værdien m

Eksempel: X antal øjne ved kast med ærlig terning, $p_m = 1/6$.

$$\mathbb{E}X = 1\frac{1}{6} + 2\frac{1}{6} + \cdots + 6\frac{1}{6} = 3.5$$

Hvis observationerne ikke er positive heltal, bruges samme opskrift:

$\mathbb{E}X$ er det vægtede gennemsnit af de mulige værdier for X, med vægte givet ved de tilknyttede sandsynligheder.

Eksempel: Antag X kan antage værdier -2, 0 og 3 med sandsynlighederne 0.4, 0.5 og 0.1. Da er

$$\mathbb{E}X = -2 \cdot 0.4 + 0 \cdot 0.5 + 3 \cdot 0.1 = -0.8 + 0 + 0.3 = -0.5$$

Regneregler

Antag at X og Y er to stokastiske variable

$$\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$$

$$\mathbb{E}aX = a\mathbb{E}X$$

Tilbage til lageret

Forventet værdi af beholdning:

$$\mathbb{E}B_t = L - t\mathbb{E}X_1$$

(antager alle X_i har samme middelværdi)

For at beregne $\mathbb{E}X_1$ skal vi kende sandsynlighedsfunktionen $p(m)$ for X_1.

Definition En funktion $p(m)$ er en sandsynlighedsfunktion hvis

$$p(m) \geq 0 \text{ og } \sum_{m=1}^{M} p(m) = 1$$

X har sandsynlighedsfunktion $p(m)$ hvis $P(X = m) = p(m)$.

Fundamental sandsynlighedsregning: hændelser

Vi vil bruge A og B som notation for hændelser, som vi tillægger sandsynligheder $P(A)$ og $P(B)$.

Eksempler på hændelser:

- AAB vinder over FCK i næste kamp
- Det regner i morgen
- Kursen på Novo-aktie falder i næste uge
- Mælk bliver udsolgt i REMA 1000
- Terningkast giver 1

Eksempel: $P(\text{Terningkast giver 1}) = 1/6$.

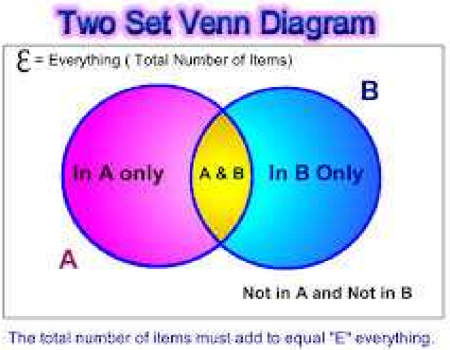
Fundamental sandsynlighedsregning

Lad A og B være to hændelser. Da gælder

$$P(A \text{ eller } B) = P(A) + P(B) - P(A \text{ og } B)$$

hvor $P(A \text{ eller } B)$ er sandsynligheden for at *mindst en* af hændelserne indtræffer og $P(A \text{ og } B)$ er sandsynligheden for at *begge* hændelser indtræffer.

Illustration vha. Venn diagram:



Eksempler

Eksempel: kast med terning A: kast med terning giver 1, B kast med terning giver 2. $P(A) = P(B) = 1/6$. $P(A \text{ og } B) = 0$ da vi ikke kan få både 1 og 2 i et terningkast. Dvs.

$$P(A \text{ eller } B) = P(\text{ 1 eller 2}) = 1/6 + 1/6 - 0 = 2/6 = 1/3$$

A: kast giver 4 eller 5 og B: kast giver 5 eller 6.
$P(A) = P(B) = 1/3$ og $P(A \text{ og } B) = P(\text{kast giver 5}) = 1/6$.
Dermed $P(A \text{ eller } B) = 1/3 + 1/3 - 1/6 = 3/6$.

Fundamental sandsynlighedsregning

Hvis A og B er *uafhængige* hændelser gælder

$$P(A \text{ og } B) = P(A)P(B)$$

A og B uafhængige: viden om A er indtruffet eller ej påvirker ikke sandsynligheden for B.

Eksempel (to kast med terning) A: 6'er i første kast. B: 6'er i andet kast.

$$P(A \text{ og } B) = P(A)P(B) = 1/36$$

Eksempel: A: det regner søndag $P(A) = 0.8$. B: AAB vinder over FCK søndag $P(B) = 0.8$. Hvis uafhængighed,

$$P(\text{det regner søndag og AAB vinder}) = 0.8 \cdot 0.8 = 0.64$$

Binomialfordelingen

Lad S_i være fejlstatus for $i$te komponent ($i = 1, \ldots, n$):

$$S_i = \begin{cases} 1 & i\text{te komponent defekt} \\ 0 & \text{ellers} \end{cases}$$

Antager S_i'er er uafhængige.

Total antal defekte: $X = (S_1 + \cdots + S_n) = \sum_{i=1}^{n} S_i$

Eksempel: lad $n = 3$ og p være sandsynligheden for defekt komponent. Sekvensen 010 angiver $S_1 = 0, S_2 = 1, S_3 = 0$. Dermed

$$P(X = 1) = P(100 \text{ eller } 010 \text{ eller } 001) =$$
$$P(100) + P(010) + P(001) =$$
$$p(1-p)(1-p) + (1-p)p(1-p) + (1-p)(1-p)p = 3p(1-p)^2$$

Bemærk, at f.eks. $A = 100$ og $B = 010$ ikke begge kan indtræffe. Antallet af måder en defekt ud af tre kan fremkomme=3.

På foregående slide antog vi at S_i'erne *uafhængige* (udfaldet af S_1 giver ikke viden om udfaldet af S_2) hvorved f.eks.

$$P(S_1 = 1 \text{ og } S_2 = 0) = P(S_1 = 1)P(S_2 = 0) = p(1-p)$$

Generelt n:

$$P(X = m) = p(m) = \binom{n}{m} p^m (1-p)^{n-m}$$

hvor

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

er antal måder vi kan få m 1'ere i n forsøg.

Ovenstående er sandsynlighedsfunktionen for en *binomial*-fordeling.

Fakultet: f.eks. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ og $0! = 1$.

Middelværdi for en binomialfordeling

Antag X er binomialfordelt $b(n, p)$ med antalsparameter n og sandsynlighedsparameter p. Så er

$$\mathbb{E}X = np$$

Kan udregnes vha. sandsynlighedsfunktionen, men nemmere:

$$\mathbb{E}X = \mathbb{E}[S_1 + \cdots + S_n] = \mathbb{E}S_1 + \cdots + \mathbb{E}S_n = np$$

Sum af to uafhængige binomialfordelinger med *samme* sandsynlighedsparameter:

$$X_1 \sim b(n_1, p) \text{ og } X_2 \sim b(n_2, p) \Rightarrow X = X_1 + X_2 \sim b(n_1 + n_2, p)$$

Eksempel (lager): der bruges mellem 0 og 10 møtrikker pr. dag og forventet værdi er 2 $\Rightarrow X_i \sim b(10, 0.2)$.

Total forbrug for en måned:

$$X = \sum_{i=1}^{30} X_i \sim b(300, 0.2)$$

Forventet månedsforbrug: $\mathbb{E}X = 60$. Sæt $L = \mathbb{E}X = 60$.

Sandsynlighed for at lageret tømmes :

$$P(B_{30} < 0) = P(L - X < 0) = P(X > L) = 1 - P(X \leq 60) = 0.4655$$

Udregnes i `R` vha. `pbinom()`-funktionen.

Fraktiler

Større lager nødvendigt. F.eks. vælge L så $P(X > L) \leq 0.01\%$.

Fraktil: tallet x er en q fraktil for X hvis $P(X \leq x) = q$. F.eks. er 2 en 7% fraktil for X hvis $P(X \leq 2) = 0.07$.

Dvs. i lager-eksemplet skal vi bestemme L så L mindst er en 99.99% fraktil for X.

Gøres f.eks. i `R` vha. `qbinom()`-funktionen.

Løsning: $L > 87$.

Er binomialfordeling rimelig model ?

Poisson-processen og Poisson-fordelingen

Antag at hændelser (f.eks. opkald til help-desk) indtræffer over tid i henhold til følgende:

1. antal hændelser i to disjunkte tidsintervaller er uafhængige
2. i meget kort tidsrum højst en hændelse
3. sandsynligheden for en hændelse i meget kort tidsrum $[t, t + \mathrm{d}t[$ er $\lambda \mathrm{d}t$

Da kaldes følgen af tidspunkter for hændelser en *Poisson*-proces.

Betragt det totale antal hændelser i et tidsrum $[0, T]$. Inddel tidsrummet i n små intervaller hver af længde T/n. Lad S_i være 1 hvis hændelse i i'te interval og nul ellers. Da har $X = \sum_i S_i$ tilnærmelsesvist en binomialfordeling $b(n, \lambda T/n)$ med middelværdi $\mu = \lambda T$.

$\mu = \lambda T$: forventet antal hændelser over tidsrum af længde T. λ: forventet antal hændelser pr. tidsenhed.

Man kan vise, at når $n \to \infty$ så vil sandsynlighedsfunktionen for $b(n, \lambda T/n) = b(n, \mu/n)$ gå mod

$$P(X = m) = \exp(-\mu)\frac{\mu^m}{m!}$$

Dette er sandsynlighedsfunktionen for en *Poisson-fordeling*.

NB: en Poisson-fordelt stokastisk variabel kan antage alle heltalsværdier.

Udledning af Poisson fordeling udfra binomial-fordelinger giver

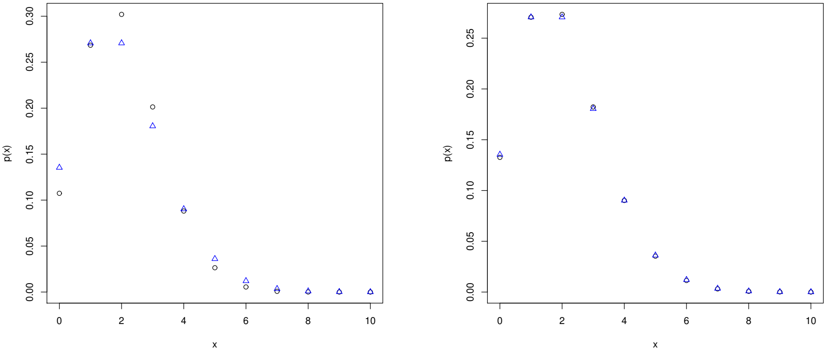
$$\mathbb{E}X = \mu$$

og (sum af uafhængige Poisson-variable)

$X_1 \sim Poisson(\mu_1)$ og $X_2 \sim Poisson(\mu_2) \Rightarrow X_1 + X_2 \sim Poisson(\mu_1 + \mu_2)$

Binomial og Poisson-fordeling

$b(10, 0.2)$ og Poisson med $\mu = 2$ henholdsvis $b(100, 0.02)$ og Poisson $\mu = 2$:



I hvert plot samme middelværdi for binomial og Poisson fordeling.

Tilbage til lageret

Antag dagligt forbrug $Poisson(2)$. Dermed månedsforbrug $Poisson(60)$.

Sandsynlighed for at lageret tømmes :

$$P(B_{30} < 0) = P(L - X < 0) = P(X > L) = 1 - P(X \leq 60) = 0.4657$$

Stort set samme tal som med binomialfordeling da $b(300, 0.2) \approx Poisson(60)$ ($n = 300$ stor !)

Med den simple lager-model kan vi nemt eksplicit beregne de fleste middelværdier og sandsynligheder.

Men hvad hvis vi f.eks. spørger om noget mere kompliceret - f.eks. forventede ventetid V til lageret går i nul:

$$V = \max\{t | B_t \geq 0\}$$

Da har V ikke en simpel sandsynlighedsfordeling.

Løsning kan være Monte Carlo beregning $=$ "empirisk beregning af middelværdier vha. computer-simulationer"

Opgaver

1. (spil med uærlig mønt) lad $X = -1$ hvis plat og 1 hvis krone. Antag sandsynlighed for plat 0.4. Hvad er $\mathbb{E}X$?
2. (kast med uærlig terning) Lad $P(X = 5) = P(X = 6) = 0.25$ og $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4)$.
   2.1 Hvad er $P(X = 1)$?
   2.2 hvad er $P(X \leq 3)$?
   2.3 Hvad er $\mathbb{E}X$?
3. Lad $\mathbb{E}X = 4$ og $\mathbb{E}Y = 5$.
   3.1 Hvad er $\mathbb{E}(2X + 3Y)$?
   3.2 Hvad er $\mathbb{E}[X - Y]$?
4. Antag to kort udtages tilfældigt (med tilbagelægning) fra et sæt spillekort. Lad hændelserne A og B være at henholdsvis det første og det andet spillekort er et es.
   4.1 Hvad er sandsynligheden for, at begge kort er esser ($P(A \text{ og } B)$) ?
   4.2 Hvad er sandsynligheden for, at første kort er et es og andet kort ikke er et es ?
   4.3 Hvad er sandsynligheden for, at ingen af kortene er et es ?
   4.4 Hvad er sandsynligheden for, at mindst et kort er et es ($P(A \text{ eller } B)$) ? På hvor mange måder kan du udregne dette ?

5. Antag $P(A) = 1/3$, $P(B) = 1/4$ og $P(A \text{ eller } B) = 1/2$ hvor A er hændelsen, at det regner på vej til arbejde, og B er hændelsen, at bussen er forsinket. Udregn $P(A \text{ og } B)$. Er A og B uafhængige hændelser ?
6. Lad $X \sim b(3, 0.25)$. Hvad er $P(X \leq 1)$? Hvad er $\mathbb{E}X$?
7. Antag $X_1 \sim b(10, 0.10)$ og $X_2 \sim b(30, 0.10)$ er uafhængige. Argumenter for at summen X er $b(40, 0.10)$. Hvad er $\mathbb{E}X$?
8. (trafikuheld i et vejkryds) Antag at sandsynligheden for en ulykke indenfor et tidsrum af 1 minut er 0.00007, at der ikke sker 2 ulykker på en gang og at ulykker indtræffer uafhængigt af hinanden.
    8.1 Hvad er det forventede antal ulykker indenfor en måned (31 dage) ?
    8.2 Hvad er sandsynligheden for, at der ikke forekommer ulykker indenfor en måned ?
    8.3 Er opgavens antagelser realistiske ?
9. Antag antal mål i første halvleg er Poisson-fordelt med middelværdi 1.5 og Poisson-fordelt med middelværdi 1 i anden halvleg og at antallene af mål er uafhængige. Hvad er sandsynligheden for at der i alt scores mindst en gang ?

Facit

1: 0.2, 2.1: $1/8$, 2.2: $3/8$, 2.3: 4, 3.1: 23, 3.2: -1, 4.1: $(1/13)^2$ 4.2: $(1/13)(12/13)$ 4.3: $(12/13)^2$ 4.4: $2/13-(1/13)^2 = 1-(12/13)^2 = 2(1/13)(12/13)+(1/13)^2$ 5: $1/12$ ja ($P(A \text{ og } B) = 1/12$ og $P(A)P(B) = 1/12$), 6: 0.84375, 0.75, 7: 4, 8.1: 3.12 8.2: 0.04, 9: 0.92