1

Utfallsrum och händelser

Def: Utfallsrumet till ett experiment är mängden av samtliga möjliga utfall (kallas Ω i fortsättningen).

Exempel: Om vi singlar slant så är utfallsrummet $\Omega = \{K, G\} = \{\text{Krona, Gubbe}\}$.

Exempel: Antalet kunder i en bank en viss dag, $\Omega = \{1, 2, 3, \ldots\}$

Exempel: Längden hos en slumpvis vald fisk, $\Omega = \{x : x > 0\}$

Vi skall senare titta på mer generella utfallsrum.

2

Def: En händelse är en mängd av olika utfall i Ω. En händelse är alltså en delmängd av Ω. En händelse kallas enkel om den består av endast ett utfall ur Ω.

Exempel: Om vi singlar slant två gånger så är $\Omega = \{KK, KG, GK, GG\}$. $A = \{KK\}$ är den enkla händelsen att vi får två kronor. $B = \{KK, KG\}$ är den sammansatta händelsen att första kastet är en krona.

Exempel: Om vi tittar på fiskexemplet kan vi definiera en händelse A som $A = \{x : 1 \leq x \leq 2\} = [1, 2]$. A är nu händelsen att fisken är mellan en och två meter lång.

Notera: Även $\emptyset$ är en händelse, den tomma händelsen. På samma sätt är Ω en händelse.

3

Mängdlära

Att händelser är mängder gör att vi kan använda begrepp från mängdlära.

Def:

1 Unionen av två händelser $A \cup B$ ges då av att minst en av händelserna inträffar.

2 snittet av två händelser $A \cap B$ ges av att båda händelserna inträffar.

3 Komplementet till en händelse A ges av A^C.

Def: Om A och B ej har några gemensamma händelser sägs de vara disjunkta, $A \cap B = \emptyset$.

4

Sannolikhetsmått

Def: Ett sannolikhetsmått är en funktion P från delmängder av Ω som för varje händelse A i Ω ger ett reellt tal sådant att

1 $0 \leq P(A) \leq 1$
alla händelser har en sannolikhet mellan 0 och 1.

2 $P(\Omega) = 1$, Utfallsrummet är en säker händelse

3 Om $A \cap B = \emptyset$ så har vi att $P(A \cup B) = P(A) + P(B)$. Sannolikheten för en union av skilda händelser är lika med summan av sannolikheterna för händelserna. Annars har vi att $P(A \cup B) \leq P(A) + P(B)$, måttet är subadditivt.

5

Nästan säkert

Def: En händelse A sägs inträffa nästan säkert om den inträffar med sannolikhet 1, d.v.s $P(A) = 1$

Notera att detta ej betyder att $A = \Omega$. Som exempel kan vi ha en likafördelad variabel på $[0, 1]$. Denna antar nästan säkert inte värdet 0.5.

Def: Något som kan inträffa men som inträffar med sannolikhet 0 kallas för en nollhändelse.

6

Betingad sannolikhet

Alla sannolikheter som vi omnämnt hittills är baserade på att vi inte vet något om utfallet innan vi gör vårt experiment (Kastar vår krona eller fångar vår fisk). Om vi vet något innan vi gör experimentet så har vi extra information vilket återspeglar sig på sannolikheten

Def: Om vi har två hanädelser A, B med $P(B) > 0$ så ges sannolikheten för A *betingat* på B av

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Notera att $P(A|B)$ uppfyller kraven på ett sannolikhetsmått.

7

Oberoende

Betingad sannolikhet handlar om vilken information vi får om en händelse A om vi vet att en annan händelse B inträffat. Om vi inte får någon information sägs händelserna vara oberoende.

Def: Två händelser sägs vara *oberoende* om

$$P(A \cap B) = P(A)P(B)$$

Detta medför att

$$P(A|B) = P(A)$$

8

Slumpvariabler

Def: En reellvärd funktion definierad på ett utfallsrum sägs vara en *slumpvariabel* En slumpvariabel är alltså en regel som associerar ett tal med varje utfall i ett utfallsrum.

Exempel: Om vi singlar slant så är $\Omega = \{K, G\}$. Vi kan definiera en slumpvarialbel som

$$X = \begin{cases} 1 & \text{om K inträffar} \\ 0 & \text{om G inträffar} \end{cases}$$

Exempel: Om vi fiskar (och utfallsrummet Ω därmed är samtliga fiskar i sjön) kan vi definiera två slumpvariabler Y, Z som vikten respektive längden hos den fisk vi får upp.

Diskreta slumpvariabler

Def: En slumpvariabel sägs vara *diskret* om den bara kan anta ett uppräknerligt antal värden

Def: *Täthetsfunktionen* eller *frekvensfunktionen* till en diskret slumpvariabel ges av

$$f_X(x) = P(X = x) = P(\{\omega \in \Omega : X(\omega) = x\})$$

Def: *Fördelningsfunktionen* till en diskret variabel ges av

$$F_X(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\})$$

Kontinuerliga slumpvariabler

Def: X sägs vara en kontinuerlig slumpvariabel med täthetsfunktion f_X om

$$P(a \leq X \leq b) = P(x \in [a, b]) = \int_a^b f_X(u)du \quad \forall a, b \in \Re, a \leq b$$

Def: *Fördelningsfunktionen* till en kontinuerlig variabel ges av

$$F_X(x) = \int_{-\infty}^x f_X(u)du = P(X \leq x) = \int_{\Omega : X(\omega) \leq x} 1 dP(\omega)$$

Notera: Om X är en kontinuerlig variabel har vi för alla tal c att $P(X = c) = 0$, d.v.s varje tänkbart enskilt utfall av variabeln är en nollhändelse.

Väntevärde

Def: *Väntevärde* till en stokastisk variabel ges av

$$\mu_X = \mathsf{E}[X] = \sum_{i=1}^{\infty} x_i f_X(x_i)$$

$$\mu_X = \mathsf{E}[X] = \int_{-\infty}^{\infty} u f_X(u)du = \int_{\Omega} X(\omega)dP(\omega)$$

för diskreta respektive kontinuerliga variabler

Varians

Variansen för diskreta respektive kontinuerliga slupvariabler ges av

$$= \mathrm{Var}[X] = \sum_{i=1}^{\infty}(x - \mu_X)^2 f_X(x)$$

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty}(u - \mu_x)^2 f_X(u)du = \int_{\Omega}(X(\omega) - \mu_X)^2 dP(\omega)$$

Standardavvikelsen ges av roten ur variansen

$$\sigma_X = \sqrt{\mathrm{Var}[X]}$$

och representerar genomsnittliga avvikelsen från väntevärdet

Skewness och Kurtosis

Generellt ges högre ordningens moment μ_k av

$$\mu_k = \sum_{i=1}^{\infty}(x - \mu_X)^k f_X(x)$$

$$\mu_k = \int_{-\infty}^{\infty}(u - \mu_x)^k f_X(u)du = \int_{\Omega}(X(\omega) - \mu_X)^k dP(\omega)$$

och vi har skewness (skevhet) som

$$\frac{\mu_3}{\sigma^3}$$

och kurtosis (toppighet) som

$$\frac{\mu_4}{\sigma^4}$$

Exempel på några fördelningar

De fördelningar vi mest kommer att jobba med i denna kursen är

Likafördelning $\quad f_X(x) = \begin{cases} (b-a)^{-1} & x \in [a, b] \\ 0 & \text{annars} \end{cases}$

Normalfördelning $\quad X$ är normalfördelad med väntevärde μ och varians σ^2 om den har fördelningsfunktion

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(\frac{(x-\mu)^2}{2\sigma^2})$$

Log-normal fördelning $\quad X$ är log-normal fördelad om $log(X)$ är normalfördelad.

Räkneregler

med a, b konstanter och X, Y slumpvariabler har vi att

$$\mathsf{E}[aX + b] = a * \mathsf{E}[X] + b$$

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

$$\mathrm{Var}[aX + b] = a^2 \mathrm{Var}[X]$$

$$\mathrm{Var}[X] = \mathsf{E}[X^2] - \mathsf{E}[X]^2$$

Multivariata slumpvariabler

När vi har definierat slumpvariabler på ett slumprum (Ω, P) som reellvärda funktioner från Utfallsrumet blir det lätt att hantera multivariata slumptal. Multivariata slumptal blir helt enkelt en uppsättning av funktioner från *samma* utfallsrum till de reella talen. Vi kan nu definiera den gemensamma frekvensfunktionen för två diskreta variabler som

$$f_{X,Y} = P(X = x, Y = y) = P(\{s : X(s) = x, Y(s) = y\})$$

Oberoende slumpvariabler

Två slumpvariabler sägs vara oberoende om varje par av händelse definierad enbart med variablerna var och en för sig är oberoende

$$P(A(X) \cap B(Y)) = P(A(X))P(B(Y))$$

Exempel: En händelse definierad för X kan vara $1 \le X \le 2$

Kovarians

Vi definierar kovariansen mellan två variabler som

$$\mathrm{Cov}[X,Y] = \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} (x-\mu_X)(x-\mu_Y) f_{X,Y}(x,y)\,dxdy$$

vilket även kan skrivas som

$$\mathrm{Cov}[X,Y] = \int_{\omega\in\Omega} (X(\omega)-\mu_X)(Y(\omega)-\mu_Y)\,dP(\omega)$$

Def: Två variabler sägs vara okorrelerade om $\mathrm{Cov}[X,Y]=0$.

OBS: att variabler är okorrelerade behöver ej betyda att de är oberoende.

Mer om normalfördelning

1 Summan av två normalfördelade variabler är alltid en normalfördelad variabel, oavsett om variablerna är korrelerade eller inte

2 Om vi vill simulera en multivariat normalfördelad variabel med kovariansmatris Σ och medelvärde μ kan vi göra detta genom att bilda

$$X = A\epsilon + \mu$$

där A är Cholesky-faktoriseringen av Σ ($AA^{\mathrm{T}} = \Sigma$), μ är de önskade medelvärdena och ϵ är en vektor av oberoende $N(0,1)$ variabler.

Betingad fördelningsfunktion

Om vi vet utfallet hos den ena slumpvariabeln kan vi ta fram den betingade fördelningsfunktonen. Låt oss anta att slumpvariabeln Y har antagit utfallet $\bar{y}$

$$f_{X|Y=\bar{y}}(x) = \frac{f_{X,Y}(x,\bar{y})}{\int_{-\infty}^{\infty} f_{X,Y}(x,\bar{y})dx}$$

Stokastiska optimeringsproblem—Optimering under osäkerhet

Först ett deterministiskt optimeringsproblem

$$\begin{array}{rrcl} \min & g_0(x) & & \\ \text{då} & g_i(x) & \le & 0, \quad i=1,\dots,m \\ & x & \in & X \subset \Re^n \end{array}$$

Ett stokastiskt dito

$$\begin{array}{rrcl} \text{“min”} & g_0(x,\xi) & & \\ \text{då} & g_i(x,\xi) & \le & 0, \quad i=1,\dots,m \\ & x & \in & X \subset \Re^n \end{array}$$

Ett stokastiskt optimeringsproblem

$$\begin{array}{rrcl} \text{“min”} & g_0(x,\xi) & & \\ \text{då} & g_i(x,\xi) & \le & 0, \quad i=1,\dots,m \\ & x & \in & X \subset \Re^n \end{array}$$

- Antag att vi har fattat ett beslut $\tilde{x}$ och att en realisering $\tilde{\xi}$ av ξ har inträffat.
- Om “resurserna” inte räcker till, d.v.s, om $g_i(\tilde{x},\tilde{\xi}) > 0$ för ngt i?
- Låt $g_i^+(x,\xi) = \begin{cases} 0, & \text{om } g_i(x,\xi) \le 0, \\ g_i(x,\xi), & \text{annars.} \end{cases}$

Recourse (tillflykt, utväg)

- Inför ett straff q_i per enhet otillåtenhet:

$$Q(x,\xi) = \min_y \left\{ \sum_{i=1}^m q_i y_i(\xi) \;\middle|\; y_i(\xi) \ge g_i^+(x,\xi),\ i=1,\dots,m \right\}$$

- Funktionen $Q(x,\xi)$ kallas recourse-funktion
- Total steg 1- och recourse-kostnad: $f_0(x,\xi) = g_0(x,\xi) + Q(x,\xi)$
- Deterministiskt ekvivalent problem:

$$\min_{x\in X} E_\xi\{f_0(x,\xi)\} = \min_{x\in X} \overbrace{\{E_\xi g_0(x,\xi) + Q(x,\xi)\}}^{h(x)}$$

Exempel LP under osäkerhet

$$\begin{array}{ll} \text{“min”} & c^{\mathrm{T}}x \\ \text{Då} & T(\xi)x = h(\xi) \\ & \left.\begin{array}{l} Ax = d \\ x \ge 0 \end{array}\right\} = X \end{array}$$

- Deterministiskt ekvivalent problem:

$$\min_{x\in X} E_\xi\{c^T x + Q(x,\xi)\}$$
$$Q(x,\xi) = \min_y \{ q^T y \mid Wy = h(\xi) - T(\xi)x,\ y \ge 0 \}$$

- W oberoende av ξ — SLP med “fixed recourse”

- q = straff för att bryta mot bivillkoren (*)
- W = definierar begreppet "bryta mot"
- T.ex. $W = (I, -I)$ (enhetsmatriser) betyder straff för att bryta mot vart och ett av bivillkoren, åt båda håll
- "Simple recourse:" $W = (I, -I)$
- "Complete fixed recourse:" $\{z \mid z = Wy,\ y \geq 0\} = \Re^{m_1}$, där $W : m_1 \times n$

Deterministisk ekvivalent

Vi kan nu skriva upp hela problemet som

$$\begin{array}{lll} \min & c^T x + E_\xi[q(\xi)^T y(\xi)] & \\ \text{då} & Ax & = b \\ & T(\xi)x + W(\xi)y(\xi) & = h(\xi) \quad N.S. \\ & x, y(\xi) & \geq 0 \end{array}$$

Deterministisk ekvivalent

Om vi antar att ξ bara har ett ändligt antal utfall kan vi skriva

$$\begin{array}{lll} \min & c^T x + \sum_{i=1}^n p_i q_i^T y_i & \\ \text{då} & Ax & = b \\ & T_i x + W_i y_i & = h_i, \quad i = 1 \ldots N \\ & x, y_i & \geq 0 \end{array}$$

Ett annat sätt att modellera ett deterministiskt ekvivalent problem

$$\begin{array}{ll} \min\limits_{x \in X} & E_\xi\{g_0(x,\xi)\} \\ \text{då} & P(g_i(x,\xi) \leq 0,\ i = 1,\ldots,m) \geq \alpha \\ \text{alt.} & P(g_i(x,\xi) \leq 0) \geq \alpha_i, \quad i = 1,\ldots,m \end{array}$$

där $\alpha \in (0,1)$ alt. $\alpha_i \in (0,1)$, $\forall i$

- Probabilistiska bivillkor ("(joint) chance constraints")
- Minimera väntevärdet då risken att bryta mot något bivillkor måste vara $< 1 - \alpha$
- De första stokastiska optimeringsmodellerna som formulerades (1959) var av denna typ

Beslutsvariabler som slumpvariabler

Om vi tittar på de optimala recourse-variablerna

$$y^*(x,\xi) = \underset{y}{\operatorname{argmin}} \left\{ q^T y \;\middle|\; Wy = h(\xi) - T(\xi)x,\ y \geq 0 \right\}$$

så ser vi att för ett givet x är dessa en funktion från Ω till $\Re^n$. Alltså är det optimala värdet hos våra beslutsvariabler också slumpvariabler.

diskretisering

När vi skall optimera våra problem med hjälp av datorer är det lättare att hantera diskreta stokastiska variabler än kontinuerliga. Vi kan därför approximera kontinuerliga variabler med diskreta Detta görs vanligen via

sampling Vi drar sampel och låter dessa representera fördelningen.

anpassning Vi väljer utfall och sannolikheter så att dessa stämmer så väl med fördelningen som möjligt i något mått

Diskreta stokastiska processer

Def: Vi kan definiera en *diskret stokastisk process* som en familj av slumpvariabler indexerade med ett tidsindex.

Exempel: Dygnsmedeltempraturen i Sverige under det närmaste året är en process.

Notera: Diskret i diskret stokastisk process står för tidsdimensionen. Utfallen av slumptalen kan fortfarande ha en kontinuerlig fördelning.

Scenarioträd

Om vi diskretiserar utfallsrummet för varje steg i en stokastisk process så får vi ett scenarioträd. Vi gör som följer

1 Diskretisera det första utfallet på vanligt sätt.

2 För vart och ett av de valda diskreta utfallen skapar vi den betingade fördelningsfunktionen för den andra slumpvariabeln i processen.

3 Diskretisera den betingade fördelningen

4 Fortsätt...

Scenarioträd

Nod Ett utfall i trädet

Förälder Utfallet närmast innan en nod

Barn Utfallen efter en nod

Rot Det sanna utfallet i början av trädet, saknar förälder

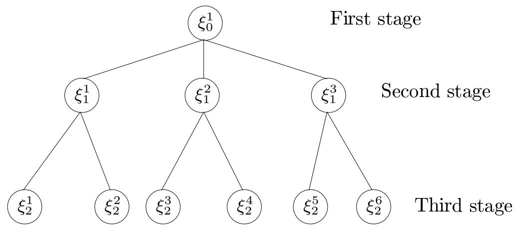
Löv Det sista steget i diskretiseringen, noder utan barn.

Flerstegsproblem

Sekvens av beslut, utfall, beslut, utfall, beslut. Vi tittar på ett fall i tre steg (endast lp)

$$\begin{array}{lcll}
\min & q_0^T x_0 + \mathsf{E}_{\xi_1}[q_1(\xi_1)_1^T x_1(\xi_1) + \mathsf{E}_{\xi_2|\xi_1}[q_2(\xi_2)_2^T x_2(\xi_2)]] & & \\
\text{då} & W_0 x_0 & = b & \\
 & A_{1,0}(\xi_1)x_0 + W_1(\xi_1)x_1(\xi_1) & = h_1(\xi_1) & N.S. \\
 & A_{2,0}(\xi_2)x_0 + A_{2,1}(\xi_2)x_1(\xi_1) + W_2(\xi_2)x_2(\xi_2) & = h_2(\xi_2) & N.S. \\
 & x_0, x_1(\xi_1), x_2(\xi_2) & \geq 0 N.S. &
\end{array}$$

Scenario tree



Multistage slp

$$\begin{aligned}
\min.\quad & c_0^{\mathrm{T}} x_0 + \mathsf{E}_{\xi_1}[c_1(\xi_1)^{\mathrm{T}} x_1(\xi_1) + \mathsf{E}_{\xi_2|\xi_1}[c_2(\xi_2)^{\mathrm{T}} x_2(\xi_2) + \\
& \ldots + \mathsf{E}_{\xi_T|\xi_{T-1}}[c_T(\xi_T)^{\mathrm{T}} x_T(\xi_T)]]], \\
\text{s.t.}\quad & W_0 x_0 = h_0, \\
& \sum_{k=0}^{t-1} A_{t,k}(\xi_t) x_k(\xi_k) + W_t(\xi_t) x_t(\xi_t) = h_t(\xi_t), \quad t = 1, \ldots, T, \\
& x_t(\xi_t) \geq 0, \quad t = 0, \ldots, T,
\end{aligned}$$

Extensive form

$$\hat{x}^{\mathrm{T}} = [\underbrace{x_1^{1\,\mathrm{T}}}_{\text{Stg. 1.}}, \underbrace{x_2^{1\,\mathrm{T}}, x_2^{2\,\mathrm{T}}, x_2^{3\,\mathrm{T}}}_{\text{Stg. 2.}}, \underbrace{x_3^{1\,\mathrm{T}}, x_3^{2\,\mathrm{T}}, x_3^{3\,\mathrm{T}}, x_3^{4\,\mathrm{T}}, x_3^{5\,\mathrm{T}}, x_3^{6\,\mathrm{T}}}_{\text{Stg. 3.}}]$$

$$\hat{c}^{\mathrm{T}} = [c_1^{1\,\mathrm{T}}, c_2^{1\,\mathrm{T}}, c_2^{2\,\mathrm{T}}, c_2^{3\,\mathrm{T}}, c_3^{1\,\mathrm{T}}, c_3^{2\,\mathrm{T}}, c_3^{3\,\mathrm{T}}, c_3^{4\,\mathrm{T}}, c_3^{5\,\mathrm{T}}, c_3^{6\,\mathrm{T}}]$$

$$\hat{A} = \begin{bmatrix}
W_1^1 & & & & & & & & & \\
A_{2,1}^1 & W_2^1 & & & & & & & & \\
A_{2,1}^2 & & W_2^2 & & & & & & & \\
A_{2,1}^3 & & & W_2^3 & & & & & & \\
A_{3,1}^1 & A_{3,2}^1 & & & W_3^1 & & & & & \\
A_{3,1}^2 & A_{3,2}^2 & & & & W_3^2 & & & & \\
A_{3,1}^3 & & A_{3,2}^3 & & & & W_3^3 & & & \\
A_{3,1}^4 & & A_{3,2}^4 & & & & & W_3^4 & & \\
A_{3,1}^5 & & & A_{3,2}^5 & & & & & W_3^5 & \\
A_{3,1}^6 & & & A_{3,2}^6 & & & & & & W_3^6
\end{bmatrix}
\quad
\hat{h} = \begin{bmatrix} h_1^1 \\ h_2^1 \\ h_2^2 \\ h_2^3 \\ h_3^1 \\ h_3^2 \\ h_3^3 \\ h_3^4 \\ h_3^5 \\ h_3^6 \end{bmatrix}$$

Slumprum

Vi har tidigare pratat om ett utfallsrum S och ett mått på samtliga händelser på S, P. Om vi formellt betecknar mängden av samtliga delmängder av S med $\mathscr{F}$ så har vi slumprummet

$$(S, \mathscr{F}, P)$$

En händelse A är nu mätbar m.a.p $\mathscr{F}$ om $A \in \mathscr{F}$ (Kom i håg att händelser är delmängder av Ω. Vi säger nu att mängden av samtliga händelser vi kan definiera för en viss slumpvariabel X är den mängd av händelser som genereras av X, $\mathscr{F}_X$.

Vi får då

$$\mathscr{F}_x \subset \mathscr{F}$$

Mätbarhet

Vi säger att en slumpvariabel är mätbar m.a.p en familj av händelser om den antas samma värde över varje enskild händelse (minsta delmängd) i sagda familj av händelser.

Filtration

Om vi låter de olika utfallen av en stokastisk process generera en sekvens av familjer av händelser kommer vi att få

$$\mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{F}_n$$

Denna sekvens av familjer av händelser kallar vi för en filtration. En annan process Y_t sägs vara mätbar m.a.p filtrationen om varje utfall Y_i är mätbar m.a.p. $\mathscr{F}_i$. I korthet betyder detta att om X_t representerar vad vi vet vid tid t så utnyttjar Y ingen information från framtiden.