Övning 1 Sannolikhetsteorins grunder

Två händelser A och B är disjunkta om $\{A \cap B\} = \emptyset$, det vill säga att snittet inte innehåller några element. Om vi har en mängd händelser $A_1, A_2, A_3, \ldots, A_n$, vilka är parvis oföreneliga, innebär detta att $\{A_i \cap A_j\} = \emptyset$ för $j \neq i$.

Kolmogorovs axiomsystem

1. *Axiom 1*: För en händelse $A \subset \Omega$, gäller $0 \leq P(A) \leq 1$.
2. *Axiom 2*: För utfallsrummet Ω gäller, $P(\Omega) = 1$.
3. *Axiom 3*: Additionsformeln: om $A_1, A_2, A_3, \ldots, A_n$ är en ändlig, eller uppräkneligt oändlig, följd av disjunkta händelser så är $P(A_1 \cup A_2 \cup A_3 \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$. Därmed fås för två disjunkta händelser som bekant att: $P(A \cup B) = P(A) + P(B)$, ty deras snitt är tomma mängden.

Additionssatsen för två godtyckliga mängder $A \subset \Omega$ and $B \subset \Omega$ gäller

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ur detta följer **Booles olikhet**

$$P(A \cup B) \leq P(A) + P(B).$$

Ty om A och B disjunkta fås $P(A \cap B) = 0$, och om de är icke-disjunkta fås $P(A \cap B) > 0$.

De Morgans regler är hjälpsamma då man ska hantera komplement eller snitt/unioner av ett flertal händelser. De är

$$(A \cap B)^* = A^* \cup B^*, \quad (A \cup B)^* = A^* \cap B^*$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Den **klassiska sannolikhetsdefinitionen** lyder: Vid likformigt sannolikhetsmått är sannolikheten för en händelse lika med kvoten mellan antalet för händelsen gynnsamma utfall och antal möjliga utfall.

Där **likformigt sannolikhetsmått** innebär att $P(\omega_i) = 1/m$, $i = 1, 2, \ldots, m$. Ett exempel är ett tärningskast, här fås utfallsrummet $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\} = \{1, 2, 3, 4, 5, 6\}$, där alla utfall har samma sannolikhet att inträffa och då fås att $P(\omega_i) = 1/6$, då vi har $m = 6$.

För teori kring kombinatorik vänligen se kursboken eller föreläsningsanteckningar.

Övning 2 och 3 Betingade sannolikheter, oberoende händelser

Betingade sannolikheter, låt A och B vara två händelser, den betingade sannolikheten för B, givet A motsvarar

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

Om $P(A) = 0$ låter vi ovanstående sannolikhet vara obestämd. Idén bakom betingade sannolikheter är att om vi vet att händelsen A inträffat så befinner vi oss i ett mindre och mer begränsat utfallsrum, således "normerar" vi alla sannolikheter med $1/P(A)$ för att summan av alla sannolikheter i A ska motsvara 1, i enlighet med Kolmogorovs andra axiom. Man kan tänka att $P(A) = P(A|\Omega) = P(A \cap \Omega)/P(\Omega) = P(A \cap \Omega) = P(A)$, d.v.s. att vi implicit betingar på hela utfallsrummet då bara sannolikheten för A söks.

Betingade sannolikheter uppfyller axiomsystemet: $P(B^*|A) = 1 - P(B|A)$, då axiomsystemet säger att

$$P(\Omega) = P(B \cup B^*) = \{B \text{ och } B^* \text{ disjunkta}\} = P(B) + P(B^*) = 1$$

Lagen om total sannolikhet lyder: Om händelserna $H_1, H_2, \ldots, H_n$ är parvis oföreneliga och $\cup_{i=1}^n H_i = \Omega$, alltså att händelserna täcker hela utfallsrummet, gäller för varje händelse $A \subset \Omega$ att

$$P(A) = P(A \cap H_1) + P(A \cap H_2) + \cdots + P(A \cap H_n) = \sum_{i=1}^{n} P(A \cap H_i) =$$

$$= \{\text{använd formeln för betingad sannolikhet}\} = \sum_{i=1}^{n} P(A|H_i)P(H_i)$$

Tolkning av detta är att vi delar upp försöket då A kan inträffa i n delfall, motsvarande $H_1, H_2, \ldots, H_n$, där $P(A|H_i)$ är sannolikheten att A inträffar i varje delfall. Då fås $P(A)$ som summan av snitten mellan A och vardera H_i, vilket kan uttryckas i termer av betingade sannolikheter.

Utifrån formeln för betingade sannolikheter och LOTS, lagen om total sannolikhet, kan vi formulera **Bayes sats**

$$P(H_i|A) = \frac{P(A \cap H_i)}{P(A)} = \frac{P(A|H_i)P(H_i)}{\sum_{i=1}^{n} P(A|H_i)P(H_i)}.$$

Övning 4 Diskreta stokastiska variabler

En **stokastisk variabel** betecknas ofta med en stor bokstav, ex. X eller Y. Den stokastiska variabeln X är en funktion från $\Omega \to \mathbb{R}$, det vill säga från utfallsrummet till reella linjen.

Det mest allmänna sättet att beskriva hur X varierar är genom dess **fördelningsfunktion**, vilken ges utav $P(X \leq x)$; som ofta förkortas till $F_X(x) = P(X \leq x)$. Tre viktiga egenskaper hos fördelningsfunktioner är att

1. De är icke-avtagande, d.v.s. om $x \leq y \Rightarrow F_X(x) \leq F_X(y)$.
2. Högerkontinuerlig, vilket kan tolkas som att det inte finns några "hopp" om man närmar sig gränsvärdet från höger (om det finns ett sådant).
3. Samt att $F_x(x) \to 1$ när $x \to \infty$, och $F_x(x) \to 0$ när $x \to -\infty$.

Diskreta stokastiska variabler antar värden i diskreta utfallsrum. Ett exempel på en diskret stokastisk variabel är: Vi satsar pengar då vi singlar slant. Säg att vi satsar 100 kronor på att få *klave*, så att vi därmed får betala 100 kronor om det blir *krona*. Vi ser alltså att vårt utfallsrum ges utav

$$\Omega = \{\omega_1, \omega_2\} = \{\text{krona}, \text{klave}\}$$

Om vi låter den stokastiska variabeln X motsvara den vinst vi gör fås att

$$X(\omega) = \begin{cases} -100, & \text{om } \omega = \omega_1 \\ 100, & \text{om } \omega = \omega_2. \end{cases}$$

Detta är ett konkret exempel på hur en stokastisk variabel mappar mellan utfallsrummet och den reella linjen.

Vidare har vi att

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a).$$

Notera dock att för en diskret stokatiskt variabel har vi att

$$P(a < X \leq b) \neq P(a \leq X \leq b) = P(a < X \leq b) + P(X = a),$$

där

$$P(X \leq x) = \sum_{k \leq x} P(X = k).$$

Ofta förkortas $P(X = k)$ till $p_X(k)$, som kallas **sannolikhetsfunktionen** för X. Från Kolmogorvs axiom har vi därefter att

- $p_X(k) \geq 0, \forall k.$
- $\sum_{k\in\Omega} p_X(k) = 1.$
- $P(X > b) = 1 - P(x \leq b) = 1 - F_X(b),\ b \in \mathbb{R}.$

Det finns ytterligare två mått som beskriver egenskaperna för vår stokastiska variabel X, och det är dess **väntevärde** och **varians**. Väntevärdet kan löst tolkas som tyngdpunkten för sannolikhetsfunktionen, och variansen som ett mått på hur mycket X sprider ut sig kring sitt väntevärde. Vi har att väntevärdet för en diskret s.v. X motsvarar

$$\mu = E[X] = \sum_{k\in\Omega} k \cdot p_X(k).$$

Väntevärdet har linjära egenskaper och således fås att

$$E[aX + b] = aE[X] + b.$$

Variansen för en diskret s.v X motsvarar

$$\sigma^2 = V(X) = E[(X - \mu)^2] = \{\text{efter utveckling av uttrycket och förenkling fås}\} = E[X^2] - (E[X])^2$$

. Vidare har vi att

$$E[g(X)] = \sum_{k\in\Omega} g(k)p_X(k) \quad \rightarrow \quad E[X^2] = \sum_{k\in\Omega_X} k^2 p_X(k).$$

Ytterligare ett mått på spridning är **standardavvikelsen**: $\sigma = D(X) = \sqrt{V(X)}$.

Variansen har, till skillnad från väntevärdet, inga linjära egenskaper utan: $V(aX + b) = a^2V(X)$. Konstanten b försvinner från uttryycket då den endast förskjuter fördelningen i sidled, och inte påverkar dess spridning.

Några exempel på diskreteta fördelningar är:

- **Poissonfördelning:** Används ofta för att beskriva antalet händelser som inträffar, oberoende av varandra, under ett givet tidsintervall. En Poissonfördelad variabel antar positiva diskreta värden. Man betecknar att $X \in \text{Po}(\lambda)$, där λ är fördelningens parameter, tillika väntevärde. Ofta kallas λ intensitet.
- **Binomialfördelning:** Låt oss beteckna antal lyckade försök i ett experiment med den stokastiska variabeln X. Vid n oberoende upprepningar av ett och samma försök, där vardera försök har en sannolikhet p att lyckas, sägs X var binomialfördelad. Detta betecknas med $X \in \text{Bin}(n, p)$.
- **För första gångs-fördelning:** Betrakta ett oberoende försök, sådant att ett visst resultat uppträder med en sannolikhet p vardera gång försöket utförs. Låt X vara antalet gånger man upprepar försöket innan resultatet uppträder. Givet detta, så vet vi att X är för första gångs-fördelad, $X \in \text{ffg}(p)$.

Övning 5 Kontinuerliga stokastiska variabler

Kontinuerliga stokastiska variabler antar värden i ett kontinuerligt utfallsrum. Vi kan, i likhet med diskreta s.v., karakterisera en kontinuerlig stokastisk variabel med hjälp av dess **fördelningsfunktion**

$$F_x(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt,$$

och utifrån denna kan vi definiera variabelns **täthetsfunktion** $f_x(x) = \frac{d}{dx}(F_X(x))$. För kontinuerliga s.v så kan man tänka att vi fördelar ut sannolikhetsmassan 1 på reella linjen i Ω, enligt täthetsfunktionen $f_X(x)$. Och sannolikheten att X antar värden inom ett intervall A motsvarar:

$$P(X \in A) = \int_A f_X(x)dx \quad \rightarrow \quad P(a < X \leq b) = \int_a^b f_X(x)dx.$$

Här ser vi alltså sannolikheten som integralen av vår sannolikhetsmassa, över intervallet av intresse. Sannolikheten att X antar ett distinkt värde motsvarar alltså 0 eftersom det blir arean under en punkt och inte under ett intervall. Vidare har vi, från Kolmogorov, att $f_X(x) \geq 0$ och att $\int_{\mathbb{R}} f_X(x) = 1$.

Såsom för diskreta s.v. kan vi beräkna **väntevärde**, **standardavvikelse** och **varians** av vår kontinuerliga s.v X

$$\mu = E[X] = \int_{\Omega_X} x f_x(x)dx$$

och standardavvikelsen ges av $\sigma = D(X) = \sqrt{V(X)}$ där variansen $V(x)$ ges av

$$\sigma^2 = E[X^2] - (E[X])^2 = \left\{E[g(X)] = \int_{\Omega_X} g(X)f_X(x)dx\right\} = \int_{\Omega_X} x^2 f_x(x)dx - \left(\int_{\Omega_X} x f_x(x)dx\right)^2$$

Övning 6 Flerdimensionella stokastiska variabler

En **tvådimensionell stokastisk variabel**, (X, Y) karakteriseras också av dess fördelningsfunktion $F_{XY}(x, y) = P(X \leq x, Y \leq y) = P(X \leq x \cap Y \leq y)$. För **diskreta tvådimensionella stokastiska variabel** fås följande sannolikhetsfunktion, då X och Y antas anta diskreta heltalsvärden:

$$p_{XY}(j, k) = P(X = j, Y = k) = P(X = j \cap Y = k), \quad j, k = 0, 1, 2, 3, \ldots$$

Således kan vi formulera fördelningsfunktionen såsom:

$$F_{XY}(x, y) = \sum_{j \leq x, k \leq y} p_{XY}(j, k)$$

Om vi vill beräkna sannolikheten att (X, Y) antar värden i en delmängd $A \subset \Omega_{XY}$ beräknar vi detta på följande vis

$$P((X, Y) \in A) = \sum_{(j,k) \in A} p_{XY}(j, k)$$

Som innan har vi att $\sum_{j \in \Omega_X} \sum_{k \in \Omega_Y} p_{XY}(j, k) = 1$ och att $0 \leq p_{XY}(j, k) \leq 1$, $\forall j, k \in \Omega_{XY}$.

För en **kontinuerlig tvådimensionell stokastisk variabel** har vi att fördelningsfuntionen motsvarar

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x, y)dydx,$$

där $f_{XY}(x, y)$ är täthetsfunktionen för (X, Y). Vi har att $f(x, y) \geq 0$, $\forall x, y \in \Omega_{XY}$ och $\iint_{\Omega_{XY}} f(x, y)dxdy = 1$, dvs att täthet är en positiv storhet och vi kan tolka att den totala sannolikhetsvolymen under $f_{XY}(x, y)$ motsvarar 1.

Om vi har fördelningsfuntionen och vill erhålla täthetsfuntionen gör vi det på följande vis:

$$f_{XY}(x,y) = \frac{\partial^2(F_{XY}(x,y))}{\partial x \partial y},$$

vi deriverar alltså med avseende på både X och Y.

Om den simultana fördelningen $f_{XY}(x,y)$ är given, kan man återfå $f_X(x)$ respektive $f_Y(y)$ genom att "marginalisera", vilket innebär att man integrerar bort en av variablerna, på följande vis

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy, \quad \text{respektive} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

För oberoende stokastiska variabler kan man erhålla den simultana täthetsfunktionen enligt:

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

och om X och Y är oberoende fås att

$$P(X \in C, Y \in D) = P(X \in C \cap Y \in D) = P(X \in C)P(Y \in D), \quad \forall D \subset \Omega_X, \forall C \subset \Omega_Y.$$

Utifrån ovanstående resomenang fås redskap då man ska behandla fördelningen för max och min av oberoende stokastiska variabler. Säg att X_1 och X_2 är två oberoende stokastiska variabler, vars fördelningar är kända. Låt oss sedan bilda en ny stokastiska variabel

$$Y = \max(X_1, X_2).$$

Vilken fördelning har Y? Detta kan vi svara på genom

$$F_Y(y) = P(Y \leq y) = P(\text{max}(X_1, X_2) \leq y) = \{\text{Om den största} \leq y, \text{ är båda} \leq y\} =$$

$$= P(\{X_1 \leq y\} \cap \{X_2 \leq y\}) = \{X_1 \text{ och } X_2 \text{ är oberoende}\} = P(X_1 \leq y)P(X_2 \leq y) = F_{X_1}(y) \cdot F_{X_2}(y)$$

Då vi känner till fördelningarna för X_1 och X_2, kan denna sannolikhet enkelt beräknas.

Om vi istället har $U = \min(X_1, X_2)$ kan vi bestämma fördelningen för U genom

$F_U(u) = P(U \leq u) = P(\text{min}(X_1, X_2) \leq u) = 1 - P(\text{min}(X_1, X_2) > u) =$

$\{\text{Om den minsta är större än u, är båda större än u}\} = 1 - P(\{X_1 > u\} \cap \{X_2 > u\}) =$

$\{X_1 \text{ och } X_2 \text{ är oberoende}\} = 1 - P(X_1 > u)P(X_2 > u) = 1 - (1 - F_{X_1}(u))((1 - F_{X_2}(u)).$

Sammanställt av: Hanna Fredenklo Jansson