Kapitel 2

Betingede fordelinger og uafhængighed

2.1. Betinget sandsynlighed.

Lad P være en sandsynlighedsfordeling på E, og lad $A, B \subseteq E$ være hændelser. Antag at $P(A) > 0$.

Definition. Ved den *betingede sandsynlighed* for B, *givet* A, forstås størrelsen

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

Fortolkningen af en betinget sandsynlighed er, at $P(B|A)$ er sandsynligheden for, at hændelsen B vil indtræffe, når man ved – eller antager, eller *betinger med* – at A vil indtræffe.

Eksempel 2.1.1. Hvad er den betingede sandsynlighed for at få mindst 11 øjne i et slag med to terninger, givet at de to terninger viser det samme? Svaret er naturligvis 1/6, fordi de seks mulige udfald (1,1), (2,2), ..., (6,6) i den hændelse, vi betinger med, er lige sandsynlige, og netop ét af dem giver en sum ≥ 11. Udregning efter ovenstående definition giver også dette resultat:

$$P(X_1 + X_2 \geq 11 \mid X_1 = X_2) = \frac{P(X_1 + X_2 \geq 11 \text{ og } X_1 = X_2)}{P(X_1 = X_2)} = \frac{1/36}{6/36} = \frac{1}{6}.$$

Betingede sandsynligheder kan også tillægges en frekvensfortolkning. Et forsøg, der går ud på at slå et slag med to terninger under den betingelse, at de to terninger skal vise det samme, kan man helt konkret udføre ved at slå om, indtil terningerne viser det samme. Den relative hyppighed af hændelsen $\{X_1 + X_2 \geq 11\}$ i en lang række af sådanne forsøg bliver en brøk, hvor nævneren er antallet af slag med $X_1 = X_2$ og tælleren er antallet af slag blandt disse, der giver $X_1 + X_2 \geq 11$. Eller, med en letforståelig (lidt sjusket) notation: Den betingede relative hyppighed af hændelsen $\{X_1 + X_2 \geq 11\}$, givet $\{X_1 = X_2\}$, er

$$\frac{\#\{X_1 + X_2 \geq 11 \text{ og } X_1 = X_2\}}{\#\{X_1 = X_2\}}$$

Hvis vi her i tæller og nævner dividerer med det samlede antal slag N (altså inklusive de "mislykkede") fås en brøk, hvor både tæller og nævner er sædvanlige relative hyppigheder. Hvis vi erstatter disse hyppigheder med de tilsvarende sandsynligheder (svarende til at lade $N \to \infty$) får vi netop den betingede sandsynlighed, som defineret ovenfor. I denne forstand passer definitionen af betinget sandsynlighed med den sædvanlige frekvensfortolkning.

OPGAVE 2.1.1. Følgende spørgsmål refererer til et slag med to terninger. Vi anvender notation som i eksempel 1.4.1.

(a) Hvad er den betingede sandsynlighed for at få sum 12, givet at summen er mindst 11?

(b) Hvad er sandsynligheden for at terningerne viser det samme, givet at summen er 7?

(c) Udregn $P(X_1 = 3 \mid X_2 = 5)$.

(d) Udregn $P(Z_1 = 2 \mid Z_2 < 6)$.

OPGAVE 2.1.2. Lad $X \in \{1, 2, \ldots, 8\}$ følge den i opgave 1.4.1 angivne fordeling P, og sæt $Y = t(X) \in \{1, 2, 3, 4\}$. Udregn

(a) $P(X = 1 \mid Y = 1)$.

(b) $P(X < 4 \mid Y < 3)$.

OPGAVE 2.1.3. En mønt kastes 10 gange.

(a) Hvad er sandsynligheden for at få "krone" den tiende gang, givet at de ni første kast giver "plat"?

(b) Hvad er sandsynligheden for at få "krone" den tiende gang, givet at netop ni af de ti kast giver "plat"?

OPGAVE 2.1.4. I eksemplet med de 25 børn og deres fødselsdage (jvf. §1.1 og §1.2), hvad er sandsynligheden for, at der ikke er to børn med samme fødselsdag, givet at ingen af børnene er født i januar?

2.2. Betinget fordeling.

SÆTNING 2.2.1. *Lad P være en sandsynlighedsfordeling på E, og lad A være en hændelse med $P(A) > 0$. Afbildningen*

$$P(\,\cdot\,|A)\colon \mathcal{D}(E) \to \mathbf{R}$$

givet ved $P(B|A) = \frac{P(A\cap B)}{P(A)}$ er da igen en sandsynlighedsfordeling. Den tilsvarende sandsynlighedsfunktion er givet ved

$$p(x|A) = \begin{cases} p(x)/P(A) & \textit{for } x \in A \\ 0 & \textit{ellers.} \end{cases}$$

BEVIS. Det er klart at $P(B|A) \geq 0$, og at $P(E|A) = 1$. For disjunkte $B, C \subseteq E$ gælder (idet $A \cap B$ og $A \cap C$ jo også er disjunkte)

$$\begin{aligned} P(B \cup C \mid A) &= \frac{P(A \cap (B \cup C))}{P(A)} = \frac{P((A \cap B) \cup (A \cap C))}{P(A)} \\ &= \frac{P(A \cap B) + P(A \cap C)}{P(A)} = P(B \mid A) + P(C \mid A). \end{aligned}$$

Hermed har vi vist, at $P(\cdot|A)$ er en sandsynlighedsfordeling. Den tilsvarende sandsynlighedsfunktion bliver

$$p(x \mid A) = P(\{x\} \mid A) = \frac{P(A \cap \{x\})}{P(A)},$$

hvilket let omskrives til sætningens sidste påstand.

Hvis X er en stokastisk variabel med fordeling P kaldes den således definerede fordeling $P(\cdot \mid A)$ den *betingede fordeling af* X*, givet* $X \in A$.

Eksempel 2.2.1. Den betingede fordeling af udfaldet af et slag med to terninger, givet at de to terninger viser det samme, har punktsandsynlighederne

$$p(x_1, x_2 \mid X_1 = X_2) = \begin{cases} 1/6 & \text{for } x_1 = x_2 \\ 0 & \text{ellers} \end{cases}$$

Bemærk, at denne fordeling naturligt kan fortolkes som ligefordelingen på den delmængde af udfaldsrummet vi betinger med, bortset fra at den altså er "udvidet" til en fordeling på det (større) oprindelige udfaldsrum, ved at punktsandsynlighederne er sat til 0 uden for den betingende hændelse. Generelt gælder åbenbart (med dette forbehold), at betingning i en ligefordeling fører til ligefordelingen på den mængde, der betinges med.

Betinget fordeling af afledt variabel. Hvis $Y = t(X)$ er en stokastisk variabel, der er fremkommet ved transformation med $t\colon E \to F$, har det umiddelbart mening at tale om den *betingede fordeling af* Y*, givet* $X \in A$. Denne fordeling, som til $B \subseteq F$ tilordner sandsynligheden

$$P(Y \in B \mid X \in A) = P\left(t^{-1}(B)|A\right) = \frac{P\left(A \cap t^{-1}(B)\right)}{P(A)}$$

fremkommer ved transformation af den betingede fordeling $P(\,\cdot\,|A)$ med t (jvf. sætning 1.4.1)

Eksempel 2.2.2. I tilfældet med de to terninger kan vi se på $Y = X_1 + X_2 \in \{2, 3, \ldots, 12\}$. Den betingede fordeling af Y, givet $X_1 = X_2$, får åbenbart punktsandsynlighederne

$$P(Y = y \mid X_1 = X_2) = \begin{cases} 1/6 & \text{for } y \text{ lige} \\ 0 & \text{for } y \text{ ulige.} \end{cases}$$

Opgave 2.2.1. Vi betragter igen et slag med to terninger, med notation som i eksempel 1.4.1.

(a) Hvad er den betingede fordeling af X_1, givet $Y = 7$?

(b) Hvad er den betingede fordeling af Z_1, givet $Y = 7$?

(c) Hvad er den betingede fordeling af Z_2, givet $X_1 = 3$?

(d) Hvad er den betingede fordeling af Y, givet $Y \geq 4$? Tegn sandsynlighedsfunktionen og sammenlign med tegningen i eksempel 1.2.1.

OPGAVE 2.2.2. En mønt kastes 10 gange. Som udfaldsrum tager vi $\{0,1\}^{10}$, med "plat" og "krone" kodet som henholdsvis 0 og 1. Den stokastiske variable betegnes $(X_1,\dots,X_{10})$.

(a) Hvad er den betingede fordeling af X_1, givet $X_1+\dots+X_{10}=7$?

(b) Hvad er den betingede fordeling af $X_1+\dots+X_{10}$, givet $X_5=X_6=X_7=X_8=X_9=X_{10}=0$?

2.3. Nogle regneregler for betingede sandsynligheder.

I nogle situationer er visse betingede sandsynligheder givet ved problemstillingen, og man må så finde de ubetingede sandsynligheder ved at "regne baglæns". Resultaterne i denne paragraf kan blandt andet bruges til dette. Vi underforstår i det følgende, at et sandsynlighedsmål P på E er givet.

SÆTNING 2.3.1 *(kædereglen). Lad $A_1,\dots,A_n$ være hændelser, således at $P(A_1\cap\dots\cap A_{n-1})>0$. Da er*

$$\begin{aligned}&P(A_1\cap\dots\cap A_n)\\&\qquad=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1\cap A_2)\dots P(A_n\mid A_1\cap\dots\cap A_{n-1})\end{aligned}$$

BEVIS. Omskriv samtlige faktorer på højre side efter definitionen af betinget sandsynlighed. Herved bliver højre side til et produkt af brøker, hvor hver tæller (bortset fra den sidste) forkorter ud mod næste brøks nævner.

EKSEMPEL 2.3.1. Betragt situationen med de 25 børns fødselsdage, altså $E=\{1,\dots,365\}^{25}$, $P=$ ligefordelingen. Lad A_k betegne den hændelse, at det k'te barns fødselsdag er forskellig fra de $k-1$ foregående, altså

$$A_k=\{X_k\neq X_1, X_k\neq X_2,\dots,X_k\neq X_{k-1}\}.$$

Så er $A_2\cap A_3\cap\dots\cap A_{25}$ netop den hændelse, at de 25 børn har 25 forskellige fødselsdage. Ifølge sætningen kan vi udregne dens sandsynlighed som produktet af $P(A_2)$ og de 23 betingede sandsynligheder

$$P(A_k\mid A_2\cap\dots\cap A_{k-1}),\qquad k=3,4,\dots,25.$$

I "lodtrækningsfortolkningen" (se §1.2) er dette sandsynligheden for, at den kugle, som barn nr. k trækker fra sin kasse, har et nummer forskelligt fra de $k-1$ numre, der allerede er brugt. Denne betingede sandsynlighed er naturligvis $\frac{365-(k-1)}{365}$, idet netop $365-(k-1)$ af de 365 mulige numre giver det ønskede resultat. Altså er

$$P(A_2\cap\dots\cap A_{25})=\frac{364}{365}\times\frac{363}{365}\times\dots\times\frac{341}{365}=0.4313.$$

Ved denne udregning har vi underforstået, at fordelingen af barn nr. k's fødselsdag, når de første $k-1$ fødselsdage er givne, er lig med den ubetingede fordeling af barnets fødselsdag (nemlig ligefordelingen på $\{1, 2, \ldots, 365\}$). Dette er faktisk en antagelse om, at de 25 fødselsdage er *stokastisk uafhængige*, men vi venter lidt endnu med den præcise definition af dette begreb.

Eksempel 2.3.2. Lad X betegne levealderen for en tilfældigt udvalgt dansk dreng, født i 1964. Som udfaldsrum er det her naturligt at vælge $E = \mathbf{N}_0 = \{0, 1, 2, \ldots\}$. Men eftersom vi i øjeblikket arbejder under den begrænsning, at udfaldsrummet skal være endeligt, bør vi egentlig vælge $E = \{0, 1, 2, \ldots, N\}$ for et passende stort tal N, der så må fortolkes som en øvre grænse for levealderen. I praksis kan det være ligegyldigt, fordi N kan vælges vilkårligt stor (f.eks. $N = 120$, eller tusind milliarder). Vi skal senere se, at man udmærket kan tale om sandsynlighedsfordelinger på udfaldsrum, som ikke er endelige, og de følgende betragtninger er gyldige uanset om vi vælger $E = \{0, 1, 2, \ldots, N\}$, eller vælger $E = \mathbf{N}_0$ med reference til definitioner og resultater som vil komme senere.

Den i §1.1 omtalte sandsynlighed for, at en mand, som fylder 20 i 1984, vil opleve sin 50 års fødselsdag, kan (i den ramme vi nu betragter, hvor også muligheden for død før det 20. år tages i betragtning) fortolkes som den betingede sandsynlighed

$$P(X \geq 50 | X \geq 20) = \frac{P(X \geq 50 \text{ og } X \geq 20)}{P(X \geq 20)} = \frac{P(X \geq 50)}{P(X \geq 20)}.$$

Ved anvendelse af sætning 2.3.1 på hændelserne $A_k = \{X \geq k\}$ fås let (idet man bemærker at $E = A_0 \supseteq A_1 \supseteq A_2 \supseteq \ldots$, således at $A_0 \cap A_1 \cap \cdots \cap A_k = A_k$),

$$\begin{aligned} &P(X \geq 50 \mid X \geq 20) \\ &\quad = P(X \geq 21 \mid X \geq 20) P(X \geq 22 \mid X \geq 21) \ldots P(X \geq 50 \mid X \geq 49). \end{aligned}$$

Det er præcis disse betingede sandsynligheder af typen "sandsynligheden for at blive $x+1$ år, givet at man er blevet x", som vi i §1.1 antog konstante og lig med de tilsvarende relative hyppigheder, observeret i 1984–85. En dødelighedstavle (se Statistisk Årbog) er netop konstrueret ved hjælp af denne succesive multiplikationsregel.

Det skulle fremgå af de to eksempler, at sætning 2.3.1 især er anvendelig i situationer, hvor man kan tænke på udfaldet som et forløb over tiden, og den hændelse man er interesseret i kan beskrives som en fællesmængde af hændelser, der vedrører hvert sit tidspunkt. Den følgende sætning handler om en helt anden type af problemer. Den er nyttig i situationer, der kan simplificeres ved "opdeling i tilfælde":

Sætning 2.3.2. *Lad* $A_1, A_2, \ldots, A_n$ *være parvis disjunkte hændelser med positive sandsynligheder, således at* $A_1 \cup A_2 \cup \cdots \cup A_n = E$. *For en* *vilkårlig hændelse B gælder da*

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n).$$

Bevis. Ved hjælp af definitionen af betinget sandsynlighed omskrives højre side umiddelbart til

$$P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n)$$

som (iflg. (3), §1.3) er lig med $P(B)$.

Eksempel 2.3.3. På et bord står to kasser. Den ene indeholder 50 røde og 50 hvide kugler. Den anden indeholder 20 røde, 80 hvide og 50 sorte kugler. Ved lodtrækning (f.eks. med en mønt) vælges en kasse, og en kugle fra denne udtages tilfældigt. Hvad er sandsynligheden for at få en rød kugle? Som udfaldsrum kan vi tage $E = \{1, 2\} \times \{\text{rød}, \text{hvid}, \text{sort}\}$. Den stokastiske variable betegnes (X, Y), med X = kassens nummer og Y = kuglens farve. De størrelser, vi kender, er $P(X = 1) = P(X = 2) = \frac{1}{2}$ og punktsandsynlighederne i de betingede fordelinger af kuglens farve, givet kassens nummer, som er

$$\begin{aligned} P(Y = \text{rød} \mid X = 1) &= \frac{50}{100}, & P(Y = \text{rød} \mid X = 2) &= \frac{20}{150} \\ P(Y = \text{hvid} \mid X = 1) &= \frac{50}{100}, & P(Y = \text{hvid} \mid X = 2) &= \frac{80}{150} \\ P(Y = \text{sort} \mid X = 1) &= \frac{0}{100}, & P(Y = \text{sort} \mid X = 2) &= \frac{50}{150} \end{aligned}$$

Af sætningen følger nu at

$$\begin{aligned} &P(Y = \text{rød}) \\ &\quad = P(X = 1)P(Y = \text{rød} \mid X = 1) + P(X = 2)P(Y = \text{rød} \mid X = 2) \\ &\quad = \frac{1}{2} \times \frac{50}{100} + \frac{1}{2} \times \frac{20}{150} = \frac{19}{60} = 31.67\%. \end{aligned}$$

Følgende "omvendingsformel" for betingede sandsynligheder kan også være nyttig.

Sætning 2.3.3. *For $P(A) > 0$ og $P(B) > 0$ gælder*

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

Bevis. Helt trivielt, ud fra definitionen af betinget sandsynlighed.

Eksempel 2.3.4. Betragt igen situationen fra eksempel 2.3.3. Hvad er den betingede sandsynlighed for, at kasse nr. 1 blev valgt, givet at kuglen er rød? For at kunne tillægge dette spørgsmål en konkret mening, må man forestille sig en situation, hvor kuglens farve kan observeres, uden at man ved hvilken kasse, den stammer fra. Situationen kunne f.eks. bestå i, at en anden person udfører lodtrækningen, trækker kuglen og viser den frem, uden at man får lejlighed til at se, hvilken kasse den blev taget fra. Sætning 2.3.3 giver

$$\begin{aligned} P(X = 1 \mid Y = \text{rød}) &= \frac{P(X = 1)}{P(Y = \text{rød})} P(Y = \text{rød} \mid X = 1) \\ &= \frac{1/2}{19/60} \times \frac{50}{100} = \frac{15}{19} = 78.95\%. \end{aligned}$$

Dette eksempel illustrerer et væsentligt aspekt af begrebet betinget fordeling. Uden kendskab til forsøgets udfald ville vi naturligvis sige, at sandsynligheden for at kuglen kommer fra kasse nr. 1 er 1/2. Men når vi ser, at kuglen er rød, ændres denne vurdering, fordi en rød kugle er et mere sandsynligt resultat af en udtagning fra kasse 1 end fra kasse 2. Man kan sige, at kuglens farve indeholder *information* om, hvilken kasse den kommer fra. En betinget sandsynlighed er en sandsynlighed, der er udregnet under hensyn til den information, der ligger i, at den betingende hændelse er indtruffet. Denne fortolkning af betingede sandsynligheder er afgørende for forståelsen af begrebet stokastisk uafhængighed, som vi nu endelig kommer til.

OPGAVE 2.3.1. En mønt kastes 10 gange. Med $X \in \{1, 2, \ldots, 11\}$ betegnes nummeret på det første kast, som giver “krone”, idet vi sætter $X = 11$ hvis alle kast giver “plat”. Angiv fordelingen af X. (Vink: Udregn først sandsynlighederne for hændelserne $\{X \geq n + 1\}$. Ud fra disse beregner man let punktsandsynlighederne i fordelingen af X).

OPGAVE 2.3.2. En terning kastes indtil den viser seks, dog højst 100 gange. Hvad er fordelingen af antal kast?

OPGAVE 2.3.3. En kasse indeholder 10 røde og 2 hvide kugler. Kugler trækkes tilfældigt uden tilbagelægning, indtil man får en hvid kugle. Hvad er fordelingen af antallet af kugler, som er tilbage i kassen efter dette?

OPGAVE 2.3.4. En mønt kastes tre gange. Hver gang mønten viser “krone” kastes en terning. Hvad er sandsynligheden for at få mindst én sekser ved denne metode? (Vink: Lad A_n, $(n = 0, 1, 2, 3)$ være den hændelse, at mønten viser “krone” netop n gange. Benyt sætning 2.3.2).

2.4. Stokastisk uafhængighed.

I det følgende skal (X_1, X_2) betegne en stokastisk variabel på et udfaldsrum af formen $E_1 \times E_2$ med sandsynlighedsfordeling P. Typisk vil X_1

og X_2 være givet som funktioner af en “underliggende” variabel X på et andet udfaldsrum E, men det behøver vi ikke at tænke på her.

I forbindelse med fordelinger på produktrum benyttes følgende sprogbrug. Fordelingen af X_1, altså sandsynlighedsfordelingen P_1 på E_1 givet ved $P_1(A_1) = P(X_1 \in A_1) = P(A_1 \times E_2)$, kaldes den *marginale fordeling* af X_1. Tilsvarende defineres den marginale fordeling P_2 af X_2. Fordelingen af (X_1, X_2) kaldes i denne sammenhæng den *simultane fordeling* af X_1 og X_2. Sprogbrugen refererer naturligvis til en opstilling af punktsandsynlighederne $p(x_1, x_2)$ for P i en tosidet $E_1 \times E_2$ tabel, hvor række- og søjlesummerne så bliver punktsandsynlighederne i de to marginale fordelinger.

Definition. X_1 og X_2 siges at være *stokastisk uafhængige*, hvis det for alle $A_1 \subseteq E_1$, $A_2 \subseteq E_2$, gælder at

$$P(\{X_1 \in A_1\} \cap \{X_2 \in A_2\}) = P(X_1 \in A_1)P(X_2 \in A_2)$$

eller (uden anvendelse af notationskonventionerne i forbindelse med stokastiske variable)

$$P(A_1 \times A_2) = P_1(A_1)P_2(A_2).$$

Betingelsen siger altså, at det for hvilke som helst to hændelser, der vedrører henholdsvis X_1 og X_2, skal gælde, at sandsynligheden for, at begge hændelser indtræffer, kan beregnes som produktet af sandsynlighederne for, at hændelserne hver for sig indtræffer.

Eksempel 2.4.1. Vi ser endnu en gang på et slag med to terninger, altså $E = E_1 \times E_2$, $E_1 = E_2 = \{1, \ldots, 6\}$, $P =$ ligefordelingen på $E_1 \times E_2$. Her er X_1 og X_2 åbenbart stokastisk uafhængige, idet vi for vilkårlige $A_1 \subseteq E_1, A_2 \subseteq E_2$, har

$$\begin{aligned} P(X_1 \in A_1 \text{ og } X_2 \in A_2) = P(A_1 \times A_2) &= \frac{|A_1 \times A_2|}{|E_1 \times E_2|} = \frac{|A_1|}{|E_1|}\frac{|A_2|}{|E_2|} \\ &= P(X_1 \in A_1)P(X_2 \in A_2). \end{aligned}$$

Det gælder åbenbart generelt, at hvis (X_1, X_2) er ligefordelt på en mængde af formen $E_1 \times E_2$, så er X_1 og X_2 uafhængige (og de to marginale fordelinger vil i øvrigt også være ligefordelinger).

Vi giver to andre karakteriseringer af begrebet uafhængighed for to stokastiske variable.

Sætning 2.4.1. *Lad $p\colon E_1 \times E_2 \to \mathbf{R}$ betegne sandsynlighedsfunktionen for fordelingen af (X_1, X_2), og lad $p_1\colon E_1 \to \mathbf{R}$ og $p_2\colon E_2 \to \mathbf{R}$ være sandsynlighedsfunktionerne for de marginale fordelinger af X_1 og X_2. Da er X_1 og X_2 stokastisk uafhængige hvis og kun hvis der for alle $(x_1, x_2) \in E_1 \times E_2$ gælder*

$$p(x_1, x_2) = p_1(x_1)p_2(x_2).$$

BEVIS. Hvis X_1 og X_2 er stokastisk uafhængige følger det umiddelbart af definitionen (med $A_1 = \{x_1\}$, $A_2 = \{x_2\}$) at $p(x_1, x_2) = p_1(x_1)p_2(x_2)$. Omvendt, hvis dette gælder for alle (x_1, x_2) fås for vilkårlige $A_1 \subseteq E_1$, $A_2 \subseteq E_2$,

$$P(A_1 \times A_2) = \sum_{(x_1,x_2)\in A_1\times A_2} p(x_1, x_2) = \sum_{(x_1,x_2)\in A_1\times A_2} p_1(x_1)p_2(x_2)$$

$$= \sum_{x_1\in A_1} \Big[\sum_{x_2\in A_2} p_1(x_1)p_2(x_2) \Big] = \sum_{x_1\in A_1} p_1(x_1) \Big[\sum_{x_2\in A_2} p_2(x_2) \Big]$$

$$= \sum_{x_1\in A_1} p_1(x_1)P_2(A_2) = P_1(A_1)P_2(A_2).$$

Sætningen viser, at når X_1 og X_2 er uafhængige kan den simultane fordeling konstrueres ud fra de to marginale, simpelthen ved at man ganger punktsandsynlighederne sammen. Det herved fremkomne sandsynlighedsmål på $E_1 \times E_2$ kaldes undertiden for *produktet* af P_1 og P_2, og betegnes $P_1 \otimes P_2$.

SÆTNING 2.4.2. *X_1 og X_2 er stokastisk uafhængige hvis og kun hvis det for ethvert par af hændelser $A_1 \subseteq E_1$ og $A_2 \subseteq E_2$ med $P(X_1 \in A_1) > 0$ gælder, at*

$$P(X_2 \in A_2 | X_1 \in A_1) = P(X_2 \in A_2).$$

BEVIS. Dette ses at være en simpel omskrivning af definitionen af stokastisk uafhængighed, idet vi ved at skrive venstre side om efter definitionen af en betinget sandsynlighed og gange over med $P(X_1 \in A_1)$ får

$$(*) \qquad P(X_1 \in A_1 \text{ og } X_2 \in A_2) = P(X_1 \in A_1)P(X_2 \in A_2).$$

Blot skal vi huske, at definitionen af uafhængighed også kræver $(*)$ opfyldt for $P(X_1 \in A_1) = 0$, hvor sætningens betingelse ikke har mening, fordi den betingede sandsynlighed ikke er defineret. Men det er klart, at $(*)$ altid vil være opfyldt i dette tilfælde også, da begge sider af lighedstegnet giver 0.

I eksempel 2.3.4 så vi, hvordan en betinget fordeling kan opfattes som en fordeling, der er udregnet under hensyn til den information, der ligger i den betingende hændelse. Sætningen ovenfor viser, at stokastisk uafhængighed af to variable kan fortolkes som den egenskab, at den betingede fordeling af den anden, givet en vilkårlig betingelse på den første, er lig med den oprindelige (ubetingede) fordeling af den anden. Uafhængighed betyder altså, at *uanset hvad vi får at vide om X_1, vil dette ikke ændre på vore forventninger til X_2.*

Som kriterium for uafhængighed kan sætningens betingelse svækkes en smule, idet det er nok at antage den opfyldt i tilfældet $A_1 = \{x_1\}$:

SÆTNING 2.4.3. *En tilstrækkelig betingelse for uafhængighed af X_1 og X_2 er, at der for alle $x_1 \in E_1$ således at $P(X_1 = x_1) > 0$, og for alle $A_2 \subseteq E_2$ gælder*

$$P(X_2 \in A_2 \mid X_1 = x_1) = P(X_2 \in A_2),$$

dvs. den betingede fordeling af X_2, givet $X_1 = x_1$, er lig med den marginale fordeling af X_2.

BEVIS. Af sætningens betingelse følger umiddelbart, at relationen

$$p(x_1, x_2) = p_1(x_1)p_2(x_2)$$

vil være opfyldt for alle x_1 således at $P(X_1 = x_1) > 0$. Men for $P(X_1 = x_1) = 0$ er denne relation trivielt opfyldt. Det følger derfor af sætning 2.4.1, at X_1 og X_2 er stokastisk uafhængige.

I eksemplet vedrørende et slag med to terninger kan vi nu – lidt mere overbevisende end tidligere – begrunde valget af ligefordelingen som den relevante sandsynlighedsmodel. Det er på forhånd klart, at de to marginalfordelinger på $\{1, \ldots, 6\}$ skal være ligefordelinger. Dette er jo nærmest definitionen på en terning, idet vi ellers ville tale om "skæve" eller "falske" terninger. Endvidere er det rimeligt at forudsætte, at de to terninger opfører sig stokastisk uafhængigt. Det modsatte ville jo betyde, at en oplysning om den ene ternings antal øjne kunne ændre på vore forventninger til den anden. Argumentet er måske mere oplagt, hvis man tænker på to kast med én terning. Men det er vel klart, at det ikke gør nogen forskel at terningerne kastes samtidigt. Vi er altså tvunget til at antage X_1 og X_2 ligefordelte og stokastisk uafhængige, og heraf følger umiddelbart (ved brug af sætning 2.4.1), at fordelingen af (X_1, X_2) er ligefordelingen på $\{1, \ldots, 6\} \times \{1, \ldots, 6\}$.

I situationen med de to terningkast (og i tilsvarende situationer med flere terningkast, møntkast, børnefødselsdage osv.) optræder den stokastiske uafhængighed altså nærmest som en definitionsmæssig antagelse. Vi kan sammenfatte vore antagelser i terningkasteksemplet i vendingen "Lad X_1 og X_2 være stokastisk uafhængige, ligefordelte på $\{1, \ldots, 6\}$", og dette vil normalt være at foretrække frem for de tungere formuleringer vi tidligere har benyttet (f.eks. i §1.4). En af de mest benyttede vendinger i sandsynlighedsregningen er "Lad $X_1, \ldots, X_n$ være uafhængige med fordelinger givet ved ... ", og i denne vending indgår stokastisk uafhængighed netop som en definitionsmæssig antagelse, der fastlægger en simultan fordeling på simplest tænkelige måde ud fra givne marginale fordelinger. Men stokastisk uafhængighed kan også optræde som et fænomen, der kan opdages og bevises. Det giver vi et eksempel på nu.

EKSEMPEL 2.4.2. Idet vi stadig betragter situationen med de to terningkast, defineres

$$S = ((X_1 + X_2 - 1) \bmod 6) + 1 = \begin{cases} X_1 + X_2 \text{ for } X_1 + X_2 \leq 6 \\ X_1 + X_2 - 6 \text{ ellers} \end{cases}$$

(idet $n \bmod m$ betegner resten ved division af n med m for $n, m \in \mathbf{N}$). Denne variabel antager sine værdier i $\{1, \ldots, 6\}$. For fastholdt X_1 er det let at se, at når X_2 gennemløber sine seks mulige værdier vil S gøre det samme. Heraf følger let, at afbildningen $(X_1, X_2) \to (X_1, S)$ fra $\{1, \ldots, 6\} \times \{1, \ldots, 6\}$ ind i sig selv er bijektiv, hvoraf igen følger at (X_1, S) har samme fordeling som (X_1, X_2), nemlig ligefordelingen. Specielt er X_1 og S stokastisk uafhængige.

For de tre variable X_1, X_2 og S gælder åbenbart, at hvilke som helst to af dem er uafhængige. Da enhver af dem på den anden side kan skrives som funktion af de to andre, kan denne parvise uafhængighed ikke på nogen måde fortolkes som uafhængighed mellem dem alle tre (sammenlign et kast med tre terninger). Så vi får åbenbart brug for en definition mere:

Definition. Lad $(X_1, \ldots, X_n)$ være en stokastisk variabel på $E_1 \times \cdots \times E_n$. Vi siger, at $X_1, \ldots, X_n$ er stokastisk uafhængige, hvis det for vilkårlige $A_1 \subseteq E_1, \ldots, A_n \subseteq E_n$ gælder, at

$$P(\{X_1 \in A_1\} \cap \cdots \cap \{X_n \in A_n\}) = P(X_1 \in A_1) \ldots P(X_n \in A_n).$$

Definitionen er helt analog til definitionen i begyndelsen af denne paragraf (for $n = 2$).

Sætning 2.4.4. *Lad p betegne sandsynlighedsfunktionen for fordelingen af $(X_1, \ldots, X_n)$, og lad $p_1, \ldots, p_n$ betegne sandsynlighedsfunktionerne for fordelingerne af $X_1, \ldots, X_n$. Da er $X_1, \ldots, X_n$ uafhængige hvis og kun hvis det for alle $(x_1, \ldots, x_n) \in E_1 \times \cdots \times E_n$ gælder, at*

$$p(x_1, \ldots, x_n) = p_1(x_1) \ldots p_n(x_n).$$

Beviset er analogt til beviset for sætning 2.4.1. Men opskrivningen af det er ret besværlig, og overlades derfor til læseren.

Opgave 2.4.1. Betragt de seks stokastiske variable X_1, X_2, Y, Z_1, Z_2 og Z fra eksempel 1.4.1. Gør rede for, at de eneste to blandt disse, som er stokastisk uafhængige, er X_1 og X_2.

Opgave 2.4.2*. Lad $X_1, \ldots, X_4$ være uafhængige. Vis at

(a) X_1 og X_2 er uafhængige.

(b) (X_1, X_2) og (X_3, X_4) er uafhængige.

Formuler det generelle resultat, som (a) og (b) er specialtilfælde af.

Opgave 2.4.3*. Lad $X_1 \in E_1$ og $X_2 \in E_2$ være uafhængige. For givne afbildninger $t_1 \colon E_1 \to F_1$ og $t_2 \colon E_2 \to F_2$ betragter vi de afledte variable $Y_1 = t_1(X_1)$ og $Y_2 = t_2(X_2)$. Vis at Y_1 og Y_2 er uafhængige.

OPGAVE 2.4.4. (karakterisering af uafhængighed ved hjælp af parvis uafhængighed). For stokastiske variable $X_1 \in E_1, \ldots, X_n \in E_n$, vis at følgende betingelser er ækvivalente:

(1) $X_1, \ldots, X_n$ er uafhængige.

(2) De to variable $(X_1, \ldots, X_k)$ og X_{k+1} er uafhængige for ethvert $k = 1, \ldots, n-1$.

Benyt dette til (v.h.a. sætning 2.4.3) at give et kriterium for uafhængighed af n variable, som har at gøre med de betingede fordelinger af X_{k+1}, givet $\{X_1 = x_1, \ldots, X_k = x_k\}$.

OPGAVE 2.4.5. Under hvilke betingelser gælder det at X og $Y = t(X)$ er uafhængige?