Föreläsningsanteckningar till Sannolikhetsteori i forskarutbildningen i matematisk statistik

Johan Jonasson
Matematiska institutionen
CTH & GU
412 96 Göteborg

November 26, 1998

Motiverande exempel

Huvudproblemet för den första delen av denna kurs är att göra en fungerande definition av begreppet "mått". Mått är ett naturligt begrepp som vi alla hört talas om i varierande sammanhang, t.ex. inom sannolikhetsteorin där vi talar om sannolikhetsmått eller till vardags då vi talar om längd, area, volym, massa, etc. När man matematiskt vill definiera så till synes enkla ting visar det sig att man får tillgripa en djup och vid första anblicken ganska besvärlig teori. För att motivera oss för detta börjar vi med att titta på följande exempel.

Antag att vi vill definiera begreppet "längd". En längd är ju något som vi tillskriver delmängder till den reella linjen. Därför vore det naturligt om vi definierade längd som en funktion l med klassen $\mathcal{P}(\mathbf{R})$ av alla delmängder till $\mathbf{R}$ som definitionsrum och $[0,\infty]$ som värdemängd. Om l ska vara vettigt definierad kräver vi åtminstone att

(1) $l(\emptyset)=0$,

(2) $l(\cup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}l(A_i)$ då A_i:na är disjunkta,

(3) $l(A)=l(A+\{x\})$, för alla A och x,

(4) $l([0,1))=1$,

där $A+\{x\}=\{a+x : a\in A\}$ är x-translationen av A. Problemet är att detta är *omöjligt*. För att inse detta ska vi byta ut reella linjen mot intervallet [0,1) och vi ska ha som konvention att alla mängdtranslationer görs med toruskonvention, dvs att ändarna av intervallet är ihopknutna så att om en translation resulterar i att en del av mängden sticker ut utanför intervallet så kommer den biten tillbaka i andra änden. Det är helt klart att en vettig definition av längd måste uppfylla (3) även med denna konvention.

Dela nu in elementen i [0,1] i grupper (ekvivalensklasser) genom att säga att a och b hör till samma grupp om $a-b$ är ett rationellt tal. Bilda nu en mängd A genom att från varje grupp välja exakt ett element och låta A bestå av dessa valda element. Att man kan göra så är precis vad som antas i det s.k. *urvalsaxiomet*. För varje rationellt tal q i [0,1] sätter vi nu $A_q = A + \{q\}$.

Observera att (3) kräver att $l(A_{q_1}) = l(A_{q_2})$ för alla q_1 och q_2. Vi ska strax se dels att A_q:na är disjunkta så att (2) kräver att $l(\cup_q A_q) = \sum_q l(A_q)$ vilket medför att $l(\cup_q A_q)$ har värdet 0 eller ∞, dels att $\cup_q A_q = [0,1)$ så att att (4) kräver att $l(\cup_q A_q) = 1$, en motsägelse.

Att A_q:na är disjunkta inser vi genom att ta $q_1 \neq q_2$ och anta att $A_{q_1} \cap A_{q_2} \neq \emptyset$. Då gäller för ett godtyckligt element $x \in A_{q_1} \cap A_{q_2}$ att x dels kan skrivas som $a + q_1$ där $a \in A$, dels som $b + q_2$ där $b \in A$. Således är $a - b = q_1 - q_2$ och vi får att a och b ligger i samma grupp trots att $a \neq b$, en motsägelse mot konstruktionen av A som ju skulle bestå av endast ett element från varje grupp.

Att unionen av A_q:na är hela [0,1) ser vi genom att observera att för ett godtyckligt tal $x \in [0,1)$ finns det ett tal a som tillhör samma grupp som x och som tillhör A. Alltså kan vi skriva $x = a + q$ för något rationellt q, dvs $x \in A_q$.

Slutsatsen av detta exempel blir att det helt enkelt inte är rimligt att alla mängder ska ha en längd. Vissa mängder är helt enkelt inte *mätbara*. Som kuriosa kan nämnas att Banach och Tarski visade att för varje par av mängder, A och B i $\mathbf{R}^3$, sådana att A och B har icketomt inre gäller att A kan delas upp i ett ändligt antal bitar som sedan enbart via translationer och vridningar kan byggas om till B. Exempelvis kan en ärta delas upp i bitar och sedan pusslas ihop till ett klot stort som hela jorden!

1 Mängdfunktioner och mått

1.1 Algebror och σ-algebror

Det inledande exemplet visar på behovet av inskränka en definition av begreppet "mått" till en klass av delmängder (till ett visst rum, t.ex. $\mathbf{R}$) som ibland måste vara strikt mindre än klassen av alla delmängder till rummet i fråga. Här kommer begreppet σ-algebra in.

DEFINITION. Låt S vara en mängd, vilken som helst. Låt Σ_0 vara en familj av delmängder till S. Vi kallar Σ_0 en *algebra* om

(i) $S \in \Sigma_0$,

(ii) $F \in \Sigma_0 \Rightarrow F^c \in \Sigma_0$,

(iii) $F, G \in \Sigma_0 \Rightarrow F \cup G \in \Sigma_0$.

En *σ-algebra*, Σ, är en algebra som uppfyller en starkare version av (iii), nämligen att om $F_1, F_2, \ldots \in \Sigma$ så gäller också att $\cup_n F_n \in \Sigma$.

Observera att man genast ser att om Σ_0 är en algebra så gäller:

- $\emptyset \in \Sigma_0$ enligt (ii),
- $F, G \in \Sigma_0 \Rightarrow F \cap G = (F^c \cup G^c)^c \in \Sigma_0$ enligt (ii) och (iii).

Om Σ är en σ-algebra generaliserar sig den andra av dessa observationer till

- $F_n \in \Sigma, n = 1, 2, \ldots \Rightarrow \cap_n F_n = (\cup_n F_n^c)^c \in \Sigma$.

I ord brukar man säga att en algebra är en familj som är sluten under ändligt många mängdoperationer och en σ-algebra är en familj som är sluten under uppräkneligt många mängdoperationer.

Som vi anar och snart ska se är det så att ett mått definieras på någon σ-algebra av delmängder till S. Därför säger vi att paret (S, Σ), där Σ är en σ-algebra av delmängder till S, är ett *mätbart rum*. Om $F \in \Sigma$ kallar vi F för en Σ-mätbar mängd.

Observera att den minsta tänkbara σ-algebran till S är $\{\emptyset, S\}$ och att den största är $\mathcal{P}(S)$. (På den förra kan man alltid definiera icketriviala mått, medan man på $\mathcal{P}(S)$ som regel ej kan göra det, precis som vi redan sett exempel på. Dock kan man ju alltid definiera det triviala måttet som är 0 för alla mängder, men detta är ur praktisk synvinkel helt ointressant.)

1.2 Genererade σ-algebror

Eftersom det som regel är omöjligt att explicit uttrycka alla element i en σ-algebra är det lämpligt att tala om "σ-algebran genererad av viss klass av delmängder", där denna klass går att ange på någon sluten form.

DEFINITION. Låt $\mathcal{C}$ vara en klass av delmängder till S. Klassen $\sigma(\mathcal{C})$, *σ-algebran genererad av $\mathcal{C}$*, definieras som *den minsta σ-algebran som innehåller $\mathcal{C}$*, dvs

$$\sigma(\mathcal{C}) = \bigcap \{\Sigma : \Sigma \ \sigma\text{-algebra}\ , \Sigma \supseteq \mathcal{C}\}.$$

(**Övning**: Visa att ett snitt av flera σ-algebror är en σ-algebra.)

Exempel. Låt $S = [0, 1]$ och $\mathcal{C} = \{[0, 1/2)\}$. Då blir $\sigma(\mathcal{C}) = \{\emptyset, [0, 1], [0, 1/2), [1/2, 1]\}$.

1.3 Borel-σ-algebror och π-system

Ett viktigt specialfall av genererade σ-algebror är de sk *Borel-σ-algebrorna*. Antag att S är ett topologiskt rum med topologin (dvs klassen av alla öppna mängder) $\mathcal{J}$. Då ges Borel-σ-algebran $\mathcal{B}(S)$ på S av $\sigma(\mathcal{J})$. De viktigaste exemplen är förstås då S är den reella linjen eller delar därav såsom $[0,1]$ eller $\mathbf{R}_+$.

PROPOSITION 1.1 $\quad \mathcal{B}(\mathbf{R}) = \sigma(\pi(\mathbf{R}))$

där $\pi(\mathbf{R}) = \{(-\infty, x], x \in \mathbf{R}\}$.

Bevis. Det är välkänt att varje öppen mängd, G, i $\mathbf{R}$ kan skrivas som en uppräknelig union av öppna intervall: $G = \cup_{n=1}^{\infty}(a_n, b_n)$. För att visa att $\mathcal{B}(\mathbf{R}) \subseteq \sigma(\pi(\mathbf{R}))$ behöver vi alltså bara visa att $(a,b) \in \sigma(\pi(\mathbf{R}))$ för varje öppet intervall (a,b). Men

$$(a,b) = (-\infty, b) \setminus (-\infty, a] = (\cup_{n=1}^{\infty}(-\infty, b - 1/n]) \setminus (-\infty, a] \in \sigma(\pi(\mathbf{R}))$$

vilket bevisar denna inklusion.

Å andra sidan gäller att $(-\infty, x] = \cap_{n=1}^{\infty}(-\infty, x + 1/n) \in \mathcal{B}(\mathbf{R})$ vilket visar att $\sigma(\pi(\mathbf{R})) \subseteq \mathcal{B}(\mathbf{R})$ och saken är klar. $\square$

Klassen $\pi(\mathbf{R})$ i propositionen ovan är ett exempel på en annan typ av mängdklasser som ofta kommer att dyka upp i denna kurs, nämligen *π-system*. Dessa definieras som mängdklasser slutna under ändliga snitt:

DEFINITION. Låt $\mathcal{I}$ vara en klass av delmängder till S som är sådan att $F, G \in \mathcal{I} \Rightarrow F \cap G \in \mathcal{I}$. Då kallas $\mathcal{I}$ för ett π-system.

1.4 Mängdfunktioner och mått

Låt $\overline{\mathbf{R}}$ beteckna den utvidgade reella linjen, dvs $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, \infty\}$. En mängdfunktion är rent allmänt en $\overline{\mathbf{R}}$-värd funktion definierad på någon klass av delmängder till något rum S. Låt Σ_0 vara en algebra av delmängder till S och låt μ_0 vara en *ickenegativ* mängdfunktion definierad på Σ_0.

DEFINITION.

- Vi säger att μ_0 är *additiv* om $\mu_0(\emptyset) = 0$ och om $F, G \in \Sigma_0, F \cap G = \emptyset \Rightarrow \mu_0(F \cup G) = \mu_0(F) + \mu_0(G)$.

- Vi säger att μ_0 är *uppräkneligt additiv* om $\mu_0(\emptyset) = 0$ och om $F_n \in \Sigma_0, n = 1, 2, \ldots, \cup_n F_n \in \Sigma_0, F_n$:na disjunkta $\Rightarrow \mu_0(\cup_n F_n) = \sum_{n=1}^{\infty} \mu_0(F_n)$. (Observera att eftersom Σ_0 bara är en algebra och inte en σ-algebra så är det inte självklart att $\cup_n F_n \in \Sigma_0$.)

- Om Σ är en σ-algebra på S och μ är en *uppräkneligt additiv mängdfunktion definierad på* Σ så kallas μ för ett *mått*. Vi kallar då trippeln (S, Σ, μ) för ett *måttrum*. Om dessutom $\mu(S) < \infty$ kallas μ för ett *ändligt mått*. Vi kallar μ för ett *σ-ändligt mått* om det finns mängder $S_1, S_2, \ldots$ i Σ sådana att $\cup_n S_n = S$ och $\mu(S_n) < \infty$ för alla n.

- Ett *sannolikhetsmått* är ett mått μ med $\mu(S) = 1$.

En vanlig term inom måtteorin är uttrycket "nästan överallt" (eller "nästan säkert" som man ofta säger när man refererar till ett sannolikhetsmått). Antag att $\mathcal{S}$ är ett påstående om punkter i S, dvs för varje $s \in S$ är $\mathcal{S}(s)$ ett påstående om s. Om det gäller att $F = \{s \in S : \mathcal{S}(s) \text{ är falsk}\} \in \Sigma$ och det gäller för ett mått μ på Σ att $\mu(F) = 0$ säger man att $\mathcal{S}$ gäller *nästan överallt* (med avseende på μ).

1.5 Dynkins lemma och Unicitetssatsen

En oerhört nyttig sats är följande, som talar om när två ändliga mått är lika.

SATS 1.2 (UNICITETSSATSEN) *Låt Σ vara en σ-algebra på S och antag att $\mathcal{I}$ är ett π-system som genererar Σ. Om μ_1 och μ_2 är två mått definierade på Σ sådana att*

- $\mu_1(S) = \mu_2(S) < \infty$,
- $\mu_1(I) = \mu_2(I)$ *för alla* $I \in \mathcal{I}$,

gäller att $\mu_1 \equiv \mu_2$.

Omedelbar följd: *Om två sannolikhetsmått stämmer överens på $\mathcal{I}$ så är de lika.*

För att nå fram till ett bevis av unicitetssatsen behöver vi kunna *Dynkins lemma* som är ett av de klassiska resultaten inom måtteorin och som vi kommer att ha glädje av även i andra sammanhang än som nyckel till beviset av unicitetssatsen. Först måste vi veta vad ett d-system är:

DEFINITION. En klass $\mathcal{D}$ av delmängder till S är ett *d-system* om

(a) $S \in \mathcal{D}$,

(b) $A, B \in \mathcal{D}, A \subseteq B \Rightarrow B \setminus A \in \mathcal{D}$,

(c) $A_n \in \mathcal{D}, n = 1, 2, \dots, A_n \uparrow A \Rightarrow A \in \mathcal{D}$.

PROPOSITION 1.3 *Σ är en σ-algebra om och endast om Σ är att π-system och ett d-system.*

Bevis. Att en σ-algebra är ett π-system och ett d-system är uppenbart så vi koncentrerar oss på den andra implikationen. Vi ska då visa att om Σ är ett π-system och ett d-system så gäller att

(i) $S \in \Sigma$,

(ii) $F \in \Sigma \Rightarrow F^c \in \Sigma$,

(iii) $F_n \in \Sigma, n = 1, 2, \dots \Rightarrow \cup_n F_n \in \Sigma$.

Villkor (i) följer direkt av (a) i definitionen av ett d-system och (ii) följer av (b) eftersom $F^c = S \setminus F$. För att visa (iii), observera att för varje n har vi att $\cup_{k=1}^n F_k = (\cap_{k=1}^n F_k^c)^c = S \setminus (\cap_{k=1}^n F_k^c)$. Eftersom vi nyss visade att Σ är sluten under komplementbildning har vi att $F_k^c \in \Sigma$ så enligt π-systemegenskapen och (b) har vi att $\cup_{k=1}^n F_k \in \Sigma$ och (iii) följer av (c). $\square$

På samma sätt som för σ-algebror definieras d-systemet genererat av en klass $\mathcal{C}$ som snittet av alla d-system som innehåller $\mathcal{C}$. Beteckningen är $d(\mathcal{C})$. Notera att eftersom en σ-algebra alltid är ett d-system så gäller att $\sigma(\mathcal{C}) \supseteq d(\mathcal{C})$. Man frågar sig när det gäller att $\sigma(\mathcal{C}) = d(\mathcal{C})$. Dynkins lemma är ett svar på denna fråga:

SATS 1.4 (DYNKINS LEMMA) *Om $\mathcal{I}$ är ett π-system gäller att $d(\mathcal{I}) = \sigma(\mathcal{I})$.*

Bevis. Vi behöver visa att $d(\mathcal{I}) \supseteq \sigma(\mathcal{I})$ och för att göra det krävs det enligt propositionen ovan bara att vi visar att $d(\mathcal{I})$ är ett π-system. Med andra ord ska vi visa att klassen

$$\mathcal{D}_2 = \{B \in d(\mathcal{I}) : B \cap C \in d(\mathcal{I}) \text{ för alla } C \in d(\mathcal{I})\}$$

är just $d(\mathcal{I})$. Vi klarar inte detta direkt utan får ta det i två steg. Vi börjar med att bevisa att klassen

$$\mathcal{D}_1 = \{B \in d(\mathcal{I}) : B \cap C \in d(\mathcal{I}) \text{ för alla } C \in \mathcal{I}\}$$

sammanfaller med $d(\mathcal{I})$. Nu är ju $\mathcal{I}$ ett π-system så det är helt klart att $\mathcal{D}_1 \supseteq \mathcal{I}$. Därför räcker det att visa att $\mathcal{D}_1$ är ett d-system. Då gäller det att visa att

(a) $S \in \mathcal{D}_1$,

(b) $A \subseteq B, A, B \in \mathcal{D}_1 \Rightarrow B \setminus A \in \mathcal{D}_1$,

(c) $A_n \uparrow A, A_n \in \mathcal{D}_1 \Rightarrow A \in \mathcal{D}_1$.

Eftersom $S \cap C = C$ för alla C är del (a) uppenbar. Del (b) följer av att om $C \in \mathcal{I}$ och $A \subseteq B$ och $A, B \in \mathcal{D}_1$ så har vi att $(B \setminus A) \cap C = B \cap C \setminus A \cap C \in d(\mathcal{I})$, ty $A \cap C$ och $B \cap C$ ligger i $d(\mathcal{I})$ enligt definitionen av $\mathcal{D}_1$ och $d(\mathcal{I})$ är ett d-system.

Del (c) följer likadant; $A_n \in \mathcal{D}_1, A_n \uparrow A$ medför att $A_n \cap C \in d(\mathcal{I})$ för alla n då $C \in \mathcal{I}$ enligt definitionen av $\mathcal{D}_1$. Eftersom $A_n \cap C \uparrow A \cap C$ och $d(\mathcal{I})$ är ett d-system följer då att $A \cap C \in d(\mathcal{I})$ som vi ville.

När vi nu vet att $\mathcal{D}_1 = d(\mathcal{I})$ följer det precis på samma sätt som ovan att $\mathcal{D}_2 \supseteq \mathcal{I}$ och därför är vi klara om vi kan visa att även $\mathcal{D}_2$ är ett d-system. Att visa det är fullständigt analogt med beviset för $\mathcal{D}_1$ och görs lämpligen som övning. $\square$

Nu blir det strax en smal sak att bevisa unicitetssatsen, men vi sparar ändå detta tills vi har fått lite mer kött på benen.

1.6 Carathéodorys utvidgningssats och Lebesguemåttet

Vi har än så länge inte sett att det överhuvud taget *existerar* några icketriviala mått på σ-algebror som är stora nog för att vara intressanta. (Däremot har vi ju sett exempel på att det inte finns icketriviala mått på $\mathcal{P}(S)$ om $S = \mathbf{R}$.) Det är lätt att konstruera mått på σ-algebror som består av endast ett ändligt antal mängder (gör gärna detta) men för att klara det på t.ex. $\mathcal{B}(\mathbf{R})$ måste man använda följande djupa resultat:

SATS 1.5 (CARATHÉODORYS UTVIDGNINGSSATS) *Låt Σ_0 vara en algebra av delmängder till något rum S och sätt $\Sigma = \sigma(\Sigma_0)$. Om μ_0 är en uppräkneligt additiv mängdfunktion definierad på Σ_0 så finns det ett mått μ på Σ sådant att $\mu = \mu_0$ på Σ_0.*

Dessutom gäller att om $\mu_0(S) < \infty$ så är denna utvidgning unik.

Lästips: Man bör åtminstone en gång i livet ha läst och förstått beviset av Carathéodorys utvidgningssats då denna är en av grundbultarna i måtteorin och därmed även sannolikhetsteorin. Hoppa dock gärna över beviset vid en första genomläsning. Inga delar av beviset kommer att användas i förtsättningen, endast resultatet som sådant.

Bevis. Unikhetsdelen följer direkt av unicitetssatsen, vilket tillåter oss att koncentrera oss på existensdelen. Vi börjar med att definiera mängdfunktionen $\mu^* : \mathcal{P}(S) \to [0, \infty]$ på följande sätt:

$$\mu^*(A) = \inf\{\sum_{j=1}^{\infty} \mu_0(E_j) : E_j \in \Sigma_0, A \subseteq \cup_{j=1}^{\infty} E_j\},$$

för alla $A \subseteq S$. Eftersom $S \in \Sigma_0$ är μ^* väldefinierad. Första steget i beviset är att visa att μ^* är ett s.k. *yttre mått*, dvs att $\mu^*(\emptyset) = 0$, $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$ och att $\mu^*(\cup_{k=1}^\infty A_j) \leq \sum_{k=1}^\infty \mu^*(A_k)$ för alla följder $\{A_k\}$ av delmängder till S.

Att $\mu^*(\emptyset) = 0$ följer direkt av att ta $E_j = \emptyset$ för alla j i definitionen av μ^* och om $A \subseteq B$ är $\mu^*(B)$ ett infimum över en mindre mängd än $\mu^*(A)$ och således gäller att $\mu^*(A) \leq \mu^*(B)$. För att bevisa att det tredje kriteriet också är uppfyllt observerar vi att enligt definitionen av μ^* gäller det för varje $\epsilon > 0$ och varje k att vi kan finna mängder $E_1^{(k)}, E_2^{(k)}, \ldots$ i Σ_0 sådana att $A_k \subseteq \cup_{j=1}^\infty E_j^{(k)}$ och $\sum_{j=1}^\infty \mu_0(E_j^{(k)}) \leq \mu^*(A_k) + \epsilon 2^{-k}$. Eftersom $\cup_k A_k \subseteq \cup_{k,j} E_j^{(k)}$ gäller att $\mu^*(\cup_k A_k) \leq \sum_{k,j} \mu_0(E_j^{(k)}) \leq \sum_k \mu^*(A_k) + \epsilon$. Eftersom ϵ var godtyckligt bevisar detta den önskade subadditiviteten.

Definiera nu mängdklassen $\mathcal{F}$ genom att sätta

$$\mathcal{F} = \{A \subseteq S : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ för alla } E \subseteq S\}.$$

Steg två i beviset är att bevisa att $\mathcal{F}$ är en σ-algebra och att restriktionen av μ^* till $\mathcal{F}$ är ett mått. Observera att eftersom μ^* är ett yttre mått gäller alltid att $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$ så för att visa att $A \in \mathcal{F}$ räcker det att visa den omvända olikheten för godtyckligt E.

Att $S \in \mathcal{F}$ är uppenbart liksom att $\mathcal{F}$ är sluten under komplementbildning. För att kunna konstatera att $\mathcal{F}$ är en σ-algebra ska vi alltså visa att $\mathcal{F}$ är sluten under uppräknelig unionsbildning. Detta ska vi göra genom att dels visa att $\mathcal{F}$ är sluten under ändlig unionsbildning, dels att uppräkneliga unioner av *disjunkta* mängder i $\mathcal{F}$ tillhör $\mathcal{F}$. Om $A, B \in \mathcal{F}$ gäller för alla E att

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

$$= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)$$

$$\geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$

ty $(A \cap B) \cup (A^c \cap B) \cup (A \cap B^c) = A \cup B$. Alltså gäller att $A \cup B \in \mathcal{F}$ (så att $\mathcal{F}$ är en algebra). Dessutom gäller att om A och B är disjunkta får vi genom att sätta $E = A$ att $\mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) = \mu^*(A) + \mu^*(B)$, dvs μ^* är additiv på $\mathcal{F}$.

Antag nu att $\{A_k\}$ är en följd av disjunkta mängder i $\mathcal{F}$ och sätt $B_n = \cup_{k=1}^n A_k$, $n = 1, 2, \ldots$ och $B = \cup_{k=1}^\infty A_k$. Vi har att

$$\mu^*(E \cap B_n) = \mu^*(E \cap B_n \cap A_n) + \mu^*(E \cap B_n \cap A_n^c) = \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1}).$$

Genom ett enkelt induktionsargument följer att $\mu^*(E \cap B_n) = \sum_{k=1}^{n} \mu^*(E \cap A_k)$. Eftersom $B_n \in \mathcal{F}$ enligt vad vi visade ovan gäller nu att

$$\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c)$$

$$= \sum_{k=1}^{n} \mu^*(E \cap A_k) + \mu^*(E \cap B_n^c) \geq \sum_{k=1}^{n} \mu^*(E \cap A_k) + \mu^*(E \cap B^c).$$

Genom att låta $n \to \infty$ får vi att

$$\mu^*(E) \geq \sum_{k=1}^{\infty} \mu^*(E \cap A_n) + \mu^*(E \cap B^c) \geq \mu^*(E \cap B) + \mu^*(E \cap B^c) \geq \mu^*(E).$$

Detta bevisar att $\mathcal{F}$ är en σ-algebra och eftersom vi kan sätta $E = B$ i uttrycket ovan visar detta också att μ^* är uppräkneligt additiv på $\mathcal{F}$, dvs restriktionen av μ^* till $\mathcal{F}$ är ett mått.

Om vi nu kan visa dels att $\mu^*(A) = \mu_0(A)$ för alla $A \in \Sigma_0$, dels att $\Sigma_0 \subseteq \mathcal{F}$ är saken klar. Men för $A \in \Sigma_0$ får vi direkt att $\mu^*(A) \leq \mu_0(A)$ genom att ta $E_1 = A$ och $E_2 = E_3 = \ldots = \emptyset$ i definitionen av μ^* och om $A \subseteq \cup_{j=1}^{\infty} E_j$, $E_j \in \Sigma_0$, sätter vi $B_n = (E_n \setminus (\cup_{j=1}^{n-1} E_j)) \cap A$, $n = 1, 2, \ldots$. Då gäller att $\{B_n\}$ är en klass av disjunkta mängder sådana att $\cup_{n=1}^{\infty} B_n = A$ och $B_n \in \Sigma_0$ för alla n. Enligt den uppräkneliga additiviteten hos μ_0 gäller då att $\mu_0(A) = \sum_n \mu_0(B_n) \leq \sum_n \mu_0(E_n)$, dvs $\mu_0(A) \leq \mu^*(A)$.

Till sist, om $A \in \Sigma_0$ och $E \in \mathcal{P}(S)$ och ϵ är en godtyckligt positiv konstant, finns det en följd $\{B_j\}$ av mängder i Σ_0 sådan att $E \subseteq \cup_j B_j$ och $\sum_j \mu_0(B_j) \leq \mu^*(E) + \epsilon$. Eftersom μ_0 är additiv på Σ_0 får vi att för alla j att $\mu_0(B_j) = \mu_0(B_j \cap A) + \mu_0(B_j \cap A^c)$. Därför gäller enligt vad vi nyss visade och subadditiviten hos μ^* att

$$\mu^*(E) + \epsilon \geq \sum_j \mu_0(B_j \cap A) + \sum_j \mu_0(B_j \cap A^c) = \sum_j \mu^*(B_j \cap A) + \sum_j \mu^*(B_j \cap A^c)$$

$$\geq \mu^*((\cup_j B_j) \cap A) + \mu^*((\cup_j B_j) \cap A^c) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Eftersom ϵ var godtyckligt vald följer det att $A \in \mathcal{F}$ och saken är klar. $\square$

Det är värt att notera att σ-algebran $\mathcal{F}$ i beviset av Carathéodorys utvidgningssats kan vara strikt större än Σ.

Vi genomför nu konstruktionen av det vanliga längdmåttet, det sk *Lebesguemåttet*, på det halvöppna intervallet (0,1]:

Låt Σ_0 vara algebran som består av alla mängder av typen

$$F = (a_1, b_1] \cup (a_2, b_2] \cup \ldots \cup (a_n, b_n]$$

där $n \in \{1, 2, \ldots\}$ och $a_1 \leq b_1 \leq a_2 \leq \ldots \leq b_n$. Verifiera gärna som övning att Σ_0 är en algebra. Det är helt klart att $\Sigma = \sigma(\Sigma_0) = \mathcal{B}(0,1]$. Vi vill nu konstruera en uppräkneligt additiv mängdfunktion på Σ_0 som är lämplig som prototyp för ett längdmått. När det är gjort kan vi direkt använda Carathéodorys utvidgningssats till att utvidga denna prototyp till ett riktigt längdmått.

Naturligtvis låter vi

$$\mu_0(F) = \sum_{k=1}^{n} (b_k - a_k)$$

för alla $F \in \Sigma_0$. Det vi nu behöver göra är att kontrollera att μ_0 är uppräkneligt additiv. Att μ_0 är ändligt additiv följer direkt av definitionen, men den uppräkneliga additiviteten får vi arbeta lite med. (Här är ett **lästips** på sin plats igen: Hoppa gärna över beviset av den uppräkneliga additiviteten vid en första genomläsning.) Vi ska alltså visa att om $F \in \Sigma_0$ kan skrivas som $\cup_{n=1}^{\infty} F_n$ där $F_n \in \Sigma_0$ för alla n gäller att $\mu_0(F) = \sum_n \mu_0(F_n)$. Tack vare den ändliga additiviteten förlorar vi ingen generalitet om vi antar att $F = (a, b]$ och att $F_n = (a_n, b_n]$. För ett godtyckligt n gäller att $\cup_{j=1}^{n} F_j \in \Sigma_0$ och $F \setminus (\cup_{j=1}^{n} F_n) \in \Sigma_0$. Därför har vi att

$$\mu_0(F) = \mu_0(\cup_{j=1}^{n} F_j) + \mu_0(F \setminus (\cup_{j=1}^{n} F_j))$$

$$\geq \mu_0(\cup_{j=1}^{n} F_j) = \sum_{j=1}^{n} \mu_0(F_j).$$

Genom att låta $n \to \infty$ följer att $\mu_0(F) \geq \sum_n \mu_0(F_n)$. För att visa den omvända olikheten, fixera ett godtyckligt $\epsilon > 0$. Det gäller att klassen $\{(a_n, b_n + \epsilon 2^{-n})\}_{n=1}^{\infty}$ är en öppen övertäckning av det kompakta intervallet $[a + \epsilon, b]$. Denna övertäckning kan reduceras till en ändlig övertäckning, dvs för något index $N < \infty$ gäller att $\cup_{n=1}^{N} (a_n, b_n + \epsilon 2^{-n}) \supseteq [a+\epsilon, b]$. För ett intervall I låt $v(I)$ vara intervallets vänstra ändpunkt och $h(I)$ dess högra ändpunkt och omordna intervallen i den ändliga övertäckningen i en ordning $I_1, \ldots, I_N$ sådan att $v(I_1) \leq \ldots \leq v(I_N)$. Vi har då för alla m att $\cup_{j=1}^{m} I_m$ är ett öppet intervall. Det gäller att $h(I_1 \cup I_2) - v(I_1 \cup I_2) = h(I_1) - v(I_1) + h(I_1 \cup I_2) - h(I_1) \leq h(I_1) - v(I_1) + h(I_2) - v(I_2)$. Genom ett induktionsargument följer att $h(\cup_{m=1}^{N} I_m) - v(\cup_{m=1}^{N} I_m) \leq \sum_{m=1}^{N} (h(I_m) - v(I_m))$ vilket i sin tur implicerar att $b - a - \epsilon \leq \sum_{n=1}^{N} (b_n - a_n) + \epsilon \leq \sum_{n=1}^{\infty} (b_n - a_n) + \epsilon$. Eftersom ϵ var godtyckligt vald gäller alltså att $\mu_0(F) \leq \sum_n \mu_0(F_n)$ och den uppräkneliga additiviteten är påvisad.

M.hj.a. Carathéodorys utvidgningssats utvidgar vi nu alltså μ_0 till ett unikt (unikt p.g.a. μ_0 är ändlig) mått μ på $\mathcal{B}(0,1]$. Vi kallar μ för *Lebesguemåttet* på (0,1] och använder i fortsättningen beteckningen *Leb* för detta mått.

Om man vill kan man utvidga *Leb* från (0,1] till [0,1] genom att då $0 \in B$ sätta $Leb(B) = Leb(B \setminus \{0\})$.

För att konstruera Lebesguemåttet på hela $\overline{\mathbf{R}}$ ersätter vi (0,1] med $(-\infty, \infty]$ i konstruktionen ovan och utvidgar till $\overline{\mathbf{R}}$ på samma sätt.

1.7 Några grundläggande resultat

Här ska vi först gå igenom några basala fakta om egenskaper hos mått för att sedan bevisa unicitetssatsen. Hela tiden antar vi att (S, Σ, μ) är ett måttrum.

LEMMA 1.6 *(a)* $\mu(A \cup B) \leq \mu(A) + \mu(B)$,

(b) $\mu(\cup_{i=1}^n F_i) \leq \sum_{i=1}^n \mu(F_i)$,

(c) Om $\mu(S) < \infty$ *gäller att* $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$ *och om dessutom* $A \subseteq B$ *har vi att* $\mu(B \setminus A) = \mu(B) - \mu(A)$.

Bevisen har vi sett förr, t.ex. i Sannolikhetsteori I eller Grundkurs i matematisk statistik vid teknisk högskola. Del (c) generaliserar sig också till inklusions-exklusionsformeln som vi också sett i någon av de nämnda kurserna.

Nästa resultat är viktigt. Det stadfäster vad som brukar kallas *kontinuitet hos mått.*

LEMMA 1.7 (KONTINUITET HOS MÅTT) *(a)* $F_n \uparrow F \Rightarrow \mu(F_n) \uparrow \mu(F)$,

(b) $\mu(F_1) < \infty, F_n \downarrow F \Rightarrow \mu(F_n) \downarrow \mu(F)$,

(c) $\mu(N_n) = 0, n = 1, 2, \ldots \Rightarrow \mu(\cup_n N_n) = 0$.

Bevis.

(a) Att $F_n \uparrow F$ betyder att $F_1 \subseteq F_2 \subseteq \ldots$ och att $\cup_{n=1}^\infty F_n = F$. Vi ska dela upp F i disjunkta bitar och använda den uppräkneliga additiviteten hos μ.

Sätt $G_1 = F_1$, $G_2 = F_2 \setminus F_1$, $G_3 = F_3 \setminus (F_1 \cup F_2)$, $G_4 = F_4 \setminus (F_1 \cup F_2 \cup F_3)$, etc. Då är G_n:na disjunkta och för varje N har vi att $\cup_{n=1}^N G_n = F_N$ och således också att $\cup_n G_n = F$. Vi får

$$\mu(F) = \sum_{n=1}^\infty \mu(G_n) = \lim_{N\to\infty} \sum_{n=1}^N \mu(G_n) = \lim_{N\to\infty} \mu(F_N)$$

där den sista likheten följer av additiviteten hos μ.

(b) Låt $G_n = F_1 \setminus F_n$. Då gäller att $G_n \uparrow F_1 \setminus F$ så enligt (a) följer att $\mu(F_1 \setminus F_n) \uparrow \mu(F_1 \setminus F)$. Enligt del (c) av föregående lemma och antagandet att $\mu(F_1) < \infty$ betyder detta att $\mu(F_1) - \mu(F_n) \uparrow \mu(F_1) - \mu(F)$, dvs $\mu(F_n) \downarrow \mu(F)$.

(c) Enligt (b) i föregående lemma gäller för varje M att $\mu(\cup_{n=1}^M N_n) \leq \sum_{n=1}^M \mu(N_n) = 0$ så resultatet vi vill visa följer av (a).

□

Obs! Det är i del (b) viktigt att $\mu(F_1) < \infty$. I annat fall är slutsatsen falsk. Ett exempel är när $(S, \Sigma, \mu) = (\mathbf{R}, \mathcal{B}(\mathbf{R}), Leb)$ och $F_n = (n, \infty)$. Då gäller att $F_n \downarrow \emptyset$ men $\mu(F_n) = \infty$ för alla n.

Vi har nu kommit till beviset av unicitetssatsen som säger att två ändliga mått som överensstämmer på S och på ett π-system som genererar Σ måste vara lika.

Bevis av Unicitetssatsen. Enligt Dynkins lemma är $d(\mathcal{I}) = \sigma(\mathcal{I}) = \Sigma$ så om vi kan visa att klassen

$$\mathcal{D} = \{F \in \Sigma : \mu_1(F) = \mu_2(F)\}$$

är ett d-system så är vi klara. Att $S \in \mathcal{D}$ är en del av antagandet. Om $A \subseteq B$ och de två måtten överensstämmer på A och B får vi tack vare ändligheten hos måtten att $\mu_1(B \setminus A) = \mu_1(B) - \mu_1(A) = \mu_2(B) - \mu_2(A) = \mu_2(B \setminus A)$ och således att $B \setminus A \in \mathcal{D}$. Till sist, om $F_n \in \mathcal{D}$ och $F_n \uparrow F$ har vi att $\mu_1(F_n) = \mu_2(F_n)$ för alla n och det följer av kontinuiteten hos mått att $\mu_1(F) = \mu_2(F)$, dvs $F \in \mathcal{D}$.

2 Händelser

Inom sannolikhetsteorin har vi som bekant vårt eget språk och våra egna beteckningar. Vårt sannolikhetsrum brukar vi beteckna $(\Omega, \mathcal{F}, P)$ där Ω är ett *utfallsrum*, $\mathcal{F}$ är en σ-algebra av delmängder till Ω och P är ett sannolikhetsmått på $\mathcal{F}$.

Elementen $\omega \in \Omega$ kallar vi för *utfall*. En mängd $F \in \mathcal{F}$ kallar vi för en *händelse*. En händelse är alltså inget annat än en mängd som är mätbar m.a.p. en σ-algebra på vilken ett sannolikhetsmått finns definierat. Värdet $P(F)$ blir nu sannolikheten att det experiment det aktuella sannolikhetsrummet modellerar resulterar i ett utfall ω sådant att $\omega \in F$.

Exempel. Singla slant två gånger:

Utfallsrummet blir $\Omega = \{HH, HT, TH, TT\}$, där H står för "heads" och T för "tails", och vi kan låta $\mathcal{F} = \mathcal{P}(\Omega)$. Naturligt är nu att låta P definieras av att $P(\{\omega\}) = 1/4$ för alla $\omega \in \Omega$.

Slå tärning:

Ta lämpligen $\Omega = \{1, 2, 3, 4, 5, 6\}$ och $\mathcal{F} = \mathcal{P}(\Omega)$ och definiera P via $P(\{\omega\}) = 1/6$ för alla ω.

Välj ett tal på måfå i (0,1):

Sätt $\Omega = (0, 1)$, $\mathcal{F} = \mathcal{B}(0, 1)$ och $P = Leb$.

Exempel. Detta exempel sammanfattar avsnitt 2.3-2.4 i Williams och är till för att dels ge ytterligare ett exempel på ett sannolikhetsrum, dels visa att "överuppräkneligheter är farliga":

Antag att vi vill skapa ett sannolikhetsrum som är en lämplig modell för en oändlig följd av slantsinglingar. Utfallsrummet blir då $\Omega = \{\omega = (\omega_1, \omega_2, \ldots) : \omega_i \in \{H, T\}, i = 1, 2, \ldots\}$ eller om man så vill $\Omega = \{H, T\}^\infty$. Frågan är vilka mängder i Ω som ska vara mätbara. Genom att tänka på H som 0 och T som 1 kan vi associera varje utfall med en följd av nollor och ettor och varje sådan följd kan man ju se som en binärutveckling av ett tal i intervallet [0,1]. Eftersom det i stort sett är sant att det finns en 1-1-korrespondens mellan tal i [0,1] och följder av nollor och ettor kan vi associera Ω med [0,1]. (Inte riktigt sant ty t.ex. 1000... och 0111... svarar båda mot 1/2.) Nu vet vi ju redan att det inte går att definiera något sannolikhetsmått på $\mathcal{P}(0,1)$ och således är det också omöjligt på $\mathcal{P}(\Omega)$.

Om vi däremot låter

$$\mathcal{F} = \sigma(\{\omega \in \Omega : \omega_n = x\} : x \in \{H, T\}, n = 1, 2, \ldots),$$

dvs σ-algebran genererad av utsagor av typen "Klave i kast nr. n" och "Krona i kast nr. n", $n = 1, 2, \ldots$, så går det bra. Vi ska senare se precis hur man kan göra detta. Dock kan vi redan nu inse att en mängd av typen "Krona i kast nr. n" svarar mot en union av 2^{n-1} intervall om vi gör associationen mellan Ω och [0,1]. Därav inses att $\mathcal{F}$ svarar mot $\mathcal{B}[0,1]$ och vi kan låta P vara det mått som svarar mot *Leb*. Övertyga gärna dig själv om att P är ett lämpligt mått för att modellera slantsinglingar.

Vi ska senare visa att mängden

$$F = \{\omega \in \Omega : \text{ Proportionen klavar } \to 1/2\}$$

$$= \{\omega \in \Omega : \frac{\#\{k \leq n : \omega_k = H\}}{n} \to 1/2\}$$

är en händelse, dvs en $\mathcal{F}$-mätbar mängd. Vi känner till stora talens lag och enligt den gäller att $P(F) = 1$ (förutsatt förstås att P verkligen är lämpligt definierat). På samma sätt gäller att om $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$ är en godtycklig delföljd av $(1, 2, 3, \ldots)$ och vi sätter

$$F_\alpha = \{\omega \in \Omega : \frac{\#\{k \leq n : \omega_{\alpha_k} = H\}}{n} \to 1/2\}$$

så är $P(F_\alpha) = 1$.

Men för varje fixt $\omega = (\omega_1, \omega_2, \ldots)$ kan man ju låta $\alpha = (\alpha_1, \alpha_2, \ldots)$ bestå av av de k som är sådana att $\omega_k = H$ (eller T om det inte skulle finnas oändligt många H). Då är det klart att $\omega \notin F_\alpha$ för just detta α. Eftersom man kan finna ett sådant α till varje ω gäller att $\cap_\alpha F_\alpha = \emptyset$ så att $P(\cap_\alpha F_\alpha) = 0$ trots att $P(F_\alpha) = 1$ för varje enskilt α.

Detta till synes underliga faktum kanske skulle förbrylla den som inte är bevandrad i måtteori. Vi vet dock bättre; det är inte alls säkert att ett överuppräkneligt snitt av nästan säkra händelser inträffar nästan säkert. I vissa fall är det ju inte ens en mätbar mängd! Vad som dock är sant är att ett uppräkneligt snitt av nästan säkra händelser inträffar nästan säkert:

PROPOSITION 2.1 *Antag att $F_n \in \mathcal{F}$, $n = 1, 2, \ldots$, och att $P(F_n) = 1$ för alla n. Då gäller att $P(\cap_n F_n) = 1$.*

Bevis. Eftersom $P(F_n^c) = 0$ har vi att $P(\cap_n F_n) = 1 - P(\cup_n F_n^c) = 1 - 0$. □

Nu över till något annat. Kom ihåg hur man definierar lim sup och lim inf för följder av reella tal:

$$\limsup_n x_n = \inf_n(\sup_{m \geq n} x_m) = \lim_{n \to \infty}(\sup_{m \geq n} x_m),$$

$$\liminf_n x_n = \sup_n(\inf_{m \geq n} x_m) = \lim_{n \to \infty}(\inf_{m \geq n} x_m).$$

Eftersom $\{\sup_{m \geq n} x_m\}_{n=1}^\infty$ är en avtagande följd existerar alltid $\limsup_n x_n$, men kan vara oändligt. På samma sätt existerar alltid $\liminf_n x_n$, men kan vara oändligt.

Om $\limsup_n x_n$ och $\liminf_n x_n$ sammanfaller säger man att $\lim_{n \to \infty} x_n$ existerar och får detta gemensamma värde. (Vi tillåter alltså oändliga gränsvärden.) I annat fall existerar inte gränsvärdet.

Vi ska nu definiera lim sup och lim inf för följder av mängder: Låt $\{E_n\}$ vara en följd av delmängder till Ω. Vi sätter

$$\limsup_n E_n = \cap_n(\cup_{m \geq n} E_m),$$

$$\liminf_n E_n = \cup_n(\cap_{m \geq n} E_m).$$

Ett bra sätt att tänka på dessa begrepp är att om E_n;na är händelser så är $\limsup_n E_n$ händelsen att *oändligt många av E_n:na inträffar* och $\liminf_n E_n$ är händelsen att *alla utom ändligt många av E_n;na inträffar*. Det är en god **övning** att förstå och bevisa detta. Det är också viktigt att man inser skillnaden mellan "oändligt många" och "alla utom ändligt många".

(Som fotnot kan man observera att det finns en korrespondens mellan lim sup och lim inf för talföljder och mängdföljder, nämligen att $\limsup_n x_n$ svarar mot $\limsup_n[-\infty, x_n)$ och $\liminf_n x_n$ svarar mot $\liminf_n[-\infty, x_n)$.)

LEMMA 2.2 (FATOUS LEMMA OCH OMVÄNDA FATOUS LEMMA FÖR MÄNGDER) Dessa resultat är två enkla specialfall av de mer generella Fatous lemma vi ska se längre fram i kursen, men att visa dem är en god övning för oss:

Om E_n, $n = 1, 2, \ldots$ är händelser har vi att

(a) $P(\limsup_n E_n) \geq \limsup_n P(E_n)$,

(b) $P(\liminf_n E_n) \leq \liminf_n P(E_n)$.

Bevis. Observera först att eftersom $\mathcal{F}$ är sluten under uppräkneligt många mängdoperationer så är $\limsup_n E_n$ och $\liminf_n E_n$ mätbara och utsagorna är alltså meningsfulla.

Eftersom $\cup_{m\geq n} E_m \downarrow \limsup_n E_n$ följer det av kontinuiteten hos mått att $P(\limsup_n E_n) = \lim_{n\to\infty} P(\cup_{m\geq n} E_m)$. Då $E_{m_0} \subseteq \cup_{m\geq n} E_m$ för alla $m_0 \geq n$ har vi att $P(E_{m_0}) \leq P(\cup_{m\geq n} E_m)$ för alla $m_0 \geq n$ och således att $P(\cup_{m\geq n} E_m) \geq \sup_{m\geq n} P(E_m)$. Genom att kombinera dessa fakta får vi att

$$P(\limsup_n E_n) \geq \lim_{n\to\infty} (\sup_{m\geq n} P(E_m)) = \limsup_n P(E_n).$$

Del (b) är analog. □

Notera i beviset ovan att för del (a) måste vi använda kontinuitet uppifrån hos mått, vilket kräver ändlighet hos $P(\cup_n E_n)$ vilket följer automatiskt då P är ett sannolikhetsmått. Vi kan uppenbarligen ersätta P med ett godtyckligt ändligt mått, men inte med vilket annat mått som helst. För del (b) däremot är det kontinuitet nerifrån som gäller och det går bra att byta ut P mot vilket mått μ som helst och fortfarande få en sann utsaga.

Nästa resultat är ett mycket användbart resultat som dessutom har ett enkelt bevis:

SATS 2.3 (BOREL-CANTELLI I) *Om $\{E_n\}$ är en följd av händelser sådana att $\sum_{n=1}^{\infty} P(E_n) < \infty$ gäller att $P(\limsup_n E_n) = 0$, dvs sannolikheten att oändligt många E_n inträffar är 0.*

Bevis.

$$P(\limsup_n E_n) = P(\cap_n(\cup_{m\geq n} E_m)) \leq P(\cup_{m\geq n} E_m)$$

$$= \lim_{r\to\infty} P(\cup_{m=n}^{r} E_m) \leq \lim_{r\to\infty} \sum_{m=n}^{r} P(E_m) = \sum_{m=n}^{\infty} P(E_m)$$

för *varje* n. Genom att välja n tillräckligt stort kan vi få $\sum_{m=n}^{\infty} P(E_m)$ hur liten som helst och det följer att $P(\limsup_n E_n) = 0$. □

Exempel. Att spela ett rättvist spel kan vara lönsamt! (Om man har oändligt mycket pengar till att börja med.)

Antag att vi spelar på följande sätt: I spelomgång nr n ($n = 1, 2, \ldots$) vinner vi 1 krona med sannolikheten $1 - 1/2^n$ och förlorar $2^n - 1$ kronor med sannnolikheten $1/2^n$. Den förväntade vinsten i varje spelomgång blir då 0 så detta är vad vi brukar kalla för ett "rättvist spel".

Låt nu $E_n = \{\omega \in \Omega : \omega$ ger förlust i omgång nr. $n\}$. (Vi bekymrar oss inte om konstruktion av $(\Omega, \mathcal{F}, P)$ och mätbarhetsfrågor här.) Då har vi att $P(E_n) = 1/2^n$ så att $\sum_{n=1}^{\infty} P(E_n) = 1 < \infty$. Därför talar Borel-Cantelli I om för oss att med sannolikheten 1 kommer vi bara att förlora ändligt många gånger och således att i det långa loppet göra en oändlig vinst.

Gör nu gärna övning 2.9 i Williams för att bekanta dig närmare med definitionerna av $\limsup$ och $\liminf$ för mängder.

3 Mätbara funktioner och stokastiska variabler

Låt (S_1, Σ_1) och (S_2, Σ_2) vara två mätbara rum och antag att $f : S_1 \to S_2$ är en funktion. Man brukar säga att f är mätbar om "intressanta" utsagor om f är mätbara. Mer precist betyder detta att mängder av typen $\{s \in S_1 : f(s) \in A\}$ är mätbara delmängder av S_1 för vissa delmängder A av S_2. Av skäl som liknar dem som motiverar varför inte alltid alla delmängder av ett rum kan vara mätbara är det oftast för mycket begärt att detta ska gälla för alla delmängder A av S_2. Via σ-algebran Σ_2 talar man om vilka mängder som är "intressanta".

Definition. Låt $f : S_1 \to S_2$ där (S_1, Σ_1) och (S_2, Σ_2) är två mätbara rum. Vi säger att f är Σ_1/Σ_2-*mätbar* om $f^{-1}(A) \in \Sigma_1$ för alla $A \in \Sigma_2$.

En bild Williams ger av mätbarhetsbegreppet är den här:

$$S_1 \xrightarrow{f} S_2$$

$$\Sigma_1 \xleftarrow{f^{-1}} \Sigma_2$$

I den här kursen kommer vi nästan uteslutande att intressera oss för fallet då $(S_2, \Sigma_2) = (\overline{\mathbf{R}}, \mathcal{B}(\overline{\mathbf{R}}))$. Att vi arbetar med $\overline{\mathbf{R}}$-värda funktioner snarare än $\mathbf{R}$-värda beror på $\liminf$ och $\limsup$ av funktionsföljder alltid är definierade då, vilket gör vissa mätbarhetsfrågor lättare att besvara vilket vi alldeles strax ska se.

För enkelhets skull skriver i fortsättningen ibland bara $\mathcal{B}$ för $\mathcal{B}(\overline{\mathbf{R}})$ och när vi säger att en funktion är Σ-mätbar betyder det att funktionen i fråga är $\Sigma/\mathcal{B}$-mätbar.

Om S är ett topologiskt rum och $\Sigma = \mathcal{B}(S)$ och $h : S \to \overline{\mathbf{R}}$ är Σ-mätbar kallas h för en *Borelfunktion*. Vanligaste fallet är förstås då även $S = \overline{\mathbf{R}}$ och $\mathcal{B}(S) = \mathcal{B}$.

Williams använder följande beteckningar så låt även oss använda dem. Här är (S, Σ) ett mätbart rum:

- $m\Sigma$: Klassen av Σ-mätbara funktioner.
- $(m\Sigma)^+$: Klassen av ickenegativa Σ-mätbara funktioner.
- $b\Sigma$: Klassen av begränsade Σ-mätbara funktioner.

Innan vi tar den första satsen i det här kapitlet, låt oss minnas att alla mängdoperationer bevaras under inversa avbildningar, dvs $h^{-1}(\cup_\alpha A_\alpha) = \cup_\alpha h^{-1}(A_\alpha)$, $h^{-1}(A^c) = (h^{-1}(A))^c$, $h^{-1}(A \cap B) = h^{-1}(A) \cap h^{-1}(B)$, etc.

SATS 3.1 *Låt (S, Σ) vara ett mätbart rum och låt $h : S \to \overline{\mathbf{R}}$ vara en funktion.*

(a) Låt $\mathcal{C}$ vara en klass av delmängder till $\overline{\mathbf{R}}$ sådan att $\sigma(\mathcal{C}) = \mathcal{B}$. Då gäller att h är Σ-mätbar om $h^{-1}(C) \in \Sigma$ för alla $C \in \mathcal{C}$.

(b) Om S är ett topologiskt rum och h är kontinuerlig så är h en Borelfunktion.

(c) Om $\{s \in S : h(s) \leq c\} \in \Sigma$ för alla $c \in \overline{\mathbf{R}}$ så är h en Σ-mätbar funktion. På samma sätt räcker det att $\{h \geq c\} \in \Sigma$ för alla c eller att $\{h < c\} \in \Sigma$ för alla c, etc.

Bevis. Vi börjar med att observara att (c) följer från (a) genom att sätta $\mathcal{C} = \{[-\infty, c] : c \in \overline{\mathbf{R}}\}$ och att (b) följer från (a) om man låter $\mathcal{C}$ vara klassen av öppna mängder.

För att bevisa (b), sätt $\mathcal{E} = \{B \in \mathcal{B} : h^{-1}(B) \in \Sigma\}$. Vad vi ska visa är att $\mathcal{E} = \mathcal{B}$. Eftersom $\mathcal{E} \supseteq \mathcal{C}$ och $\mathcal{B} = \sigma(\mathcal{C})$ räcker det att visa att $\mathcal{E}$ är en σ-algebra. Vi kollar:

- $h^{-1}(\overline{\mathbf{R}}) = S \in \Sigma \Rightarrow \overline{\mathbf{R}} \in \mathcal{E}$.
- $A \in \mathcal{E} \Rightarrow h^{-1}(A) \in \Sigma \Rightarrow (h^{-1}(A))^c = h^{-1}(A^c) \in \Sigma \Rightarrow A^c \in \mathcal{E}$.
- $A_n \in \mathcal{E}, n = 1, 2, \ldots \Rightarrow h^{-1}(A_n) \in \Sigma, n = 1, 2, \ldots \Rightarrow \cup_n h^{-1}(A_n) = h^{-1}(\cup_n A_n) \in \Sigma \Rightarrow \cup_n A_n \in \mathcal{E}$.

$\square$

Det är värt att notera att del (a) generaliserar sig direkt till fallet där (S_1, Σ_1) och (S_2, Σ_2) är två generella mätbara rum.

Genom att utnyttja del (c) av ovanstående sats går det ganska smärtfritt att bevisa:

SATS 3.2 *Låt h, h_1 och h_2 vara Σ-mätbara funktioner och λ en reellvärd konstant. Då är även λh, $h_1 + h_2$ och $h_1 h_2$ Σ-mätbara.*

Bevis. Vi gör beviset för $h_1 + h_2$ och lämnar övriga fall som övning.

Enligt sats 3.1(c) räcker det om vi kan visa att $\{s \in S : h_1(s) + h_2(s) < c\} \in \Sigma$ för alla $c \in \overline{\mathbf{R}}$. Skriv

$$\{s \in S : h_1(s) + h_2(s) < c\} = \cup_{q \in \mathbf{Q}}(\{s \in S : h_1(s) < q\} \cap \{s \in S : h_2(s) < c - q\}).$$

Eftersom $\mathbf{Q}$ är uppräknelig och h_1 och h_2 är mätbara följer resultatet av de allmänna egenskaperna hos Σ. $\square$

LEMMA 3.3 *Antag att (S_i, Σ_i), $i = 1, 2, 3$, är mätbara rum och att $f : S_1 \to S_2$ och $h : S_2 \to S_3$ är Σ_1/Σ_2- respektive Σ_2/Σ_3-mätbara. Då gäller att sammansättningen $h \circ f$ är Σ_1/Σ_3-mätbar. Speciellt gäller att om $f : S \to \overline{\mathbf{R}}$ är Σ-mätbar och $h : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ är $\mathcal{B}$-mätbar så är $h \circ f$ Σ-mätbar.*

Bevis. Tag en godtycklig mängd $A \in \Sigma_3$. Då gäller att

$$(h \circ f)^{-1}(A) = f^{-1}(h^{-1}(A)) \in \Sigma_1$$

ty $h^{-1}(A) \in \Sigma_2$. $\square$

Nästa lemma säger att om man betraktar lim sup och lim inf av följder av mätbara funktioner så blir dessa i sin tur mätbara funktioner. Som vi ska se är detta mycket betydelsefullt eftersom det i sin tur innebär att allt som "rimligen är intressant" också är mätbart.

LEMMA 3.4 *Låt $\{h_n\}$ vara en följd av Σ-mätbara funktioner. Då är följande funktioner också mätbara:*

(i) $\sup_n h_n$,

(ii) $\inf_n h_n$,

(iii) $\limsup_n h_n$,

(iv) $\liminf_n h_n$.

Dessutom gäller att mängden $\{s \in S : \lim_{n\to\infty} h_n(s) \text{ existerar och är rellvärt}\}$ *är mätbar. Om $\lim_{n\to\infty} h_n(s)$ existerar för varje s är $\lim_{n\to\infty} h_n$ en mätbar funktion.*

Bevis. Del (i) följer av sats 3.1(c) då $\{\sup_n h_n \leq c\} = \cap_n \{h_n \leq c\}$. Tag nu del (ii) som övning.

Del (iii) och (iv) följer direkt av (i) och (ii), ty $\limsup_n h_n = \inf_n(\sup_{m \geq n} h_m)$ och $\liminf_n h_n = \sup_n(\inf_{m \geq n} h_m)$.

Att gränsvärdet är en mätbar funktion då den existerar ges direkt av (iii) eller (iv) då ju gränsvärdet sammanfaller med $\limsup_n h_n$ och $\liminf_n h_n$.

För att visa det påstående som återstår, sätt $g(s) = \limsup_n h_n(s) - \liminf_n h_n(s)$. Enligt (iii), (iv) och Lemma 3.2 är g mätbar och vi får att

$$\{lim_{n \to \infty} h_n \text{ existerar och är reellvärt}\}$$

$$= \{g = 0\} \cap \{\limsup_n h_n < \infty\} \cap \{\liminf_n h_n > -\infty\} \in \Sigma.$$

□

Om $(\Omega, \mathcal{F})$ är ett mätbart rum och Ω är ett utfallsrum till ett slumpförsök kallar vi en $\mathcal{F}$-mätbar funktion för en *stokastisk variabel.*

Exempel. Kom ihåg exemplet från kapitel 2 där vi började skapa en sannolikhetsteoretisk modell för en oändlig följd av slantsinglingar. Vi hade utfallrummet $\Omega = \{\omega = (\omega_1, \omega_2, \ldots) : \omega_n \in \{H, T\}, n = 1, 2, \ldots\}$. Som σ-algebra $\mathcal{F}$ på Ω tog vi

$$\mathcal{F} = \sigma(\{\omega : \omega_n = x\} : x \in \{H, T\}, n = 1, 2, \ldots\}.$$

Om vi sätter $X_n(\omega) = I_H(\omega_n)$, $n = 1, 2, \ldots$, dvs indikatorfunktionen för klave i kast nr. n så blir X_n trivialt $\mathcal{F}$-mätbar, dvs en stokastisk variabel. Låt nu $S_n = \sum_{i=1}^n X_n$, antalet klavar bland de första n singlingarna. Enligt Lemma 3.2 är även S_n/n en stokastisk variabel och således är mängden

$$\{\limsup_n S_n/n = 1/2\} \cap \{\liminf_n S_n/n = 1/2\}$$

$\mathcal{F}$-mätbar. Men denna mängd är ju precis den mängd av utfall för vilka andelen klavar konvergerar mot 1/2. Denna mängd är alltså en händelse och har alltså en sannolikhet. (Enligt stora talens lag är denna sannolikhet 1 under "rätt" förutsättningar, men det är en annan historia till vilken vi får återkomma senare.)

3.1 Fördelningsfunktioner

DEFINITION. Låt $(\Omega, \mathcal{F}, P)$ vara ett sannolikhetsrum och låt $X : \Omega \to \overline{\mathbf{R}}$ vara en stokastisk variabel. *Fördelningen* för X är sannolikhetsmåttet P_X på $\mathcal{B}(\overline{\mathbf{R}})$ givet av

$$P_X(B) = P(X^{-1}(B)) = P\{X \in B\}, B \in \mathcal{B}.$$

Fördelningfunktionen, F, till X definieras av

$$F_X(c) = P\{X \leq c\} = P_X[-\infty, c], c \in \overline{\mathbf{R}}.$$

Observera att eftersom klassen $\{[-\infty, c] : c \in \overline{\mathbf{R}}\}$ är ett π-system som genererar $\mathcal{B}$, bestäms fördelningen entydigt av F.

Det kommer inte som en överraskning att F är växande och högerkontinuerlig. Dessutom gäller att om $P\{X = \pm\infty\} = 0$ så gäller att $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$.

Övning. Använd kontinuitet hos mått för att bevisa det sistnämnda påståendet.

Existens av stokastiska variabler med önskad fördelningsfunktion: Antag att vi har en fördelningsfunktion, F, dvs en växande högerkontinuerlig funktion med $F(\infty) = 1$ och $F(-\infty) \geq 0$. Kan man alltid hitta ett sannolikhetsrum $(\Omega, \mathcal{F}, P)$ och en stokastisk variabel X definierad på Ω så att X har just F som fördelningsfunktion?

Jadå, här är två sätt på vilka man kan gå tillväga:

1. Härma konstruktionen av Lebesguemåttet: Bilda först ett mått μ på $\mathcal{B}(-\infty, \infty]$ genom att på algebran som består av mängder av typen

   $$A = (a_1, b_1] \cup (a_2, b_2] \cup \ldots \cup (a_n, b_n]$$

   $-\infty \leq a_1 \leq b_1 \leq a_2 \leq \ldots \leq b_n$, $n = 1, 2, \ldots$ låta $\mu_0(A) = \sum_{i=1}^n (F(b_i) - F(a_i))$ och utvidga μ_0 till ett unikt mått μ på $\mathcal{B}(-\infty, \infty]$ med hjälp av Carathéodorys utvidgningssats. Utvidga sedan μ till $\mathcal{B}[-\infty, \infty]$ genom att låta $\mu(B) = F(-\infty) + \mu(B \setminus \{-\infty\})$ för B sådana att $-\infty \in B$.

   Låt till sist $(\Omega, \mathcal{F}, P) = (\overline{\mathbf{R}}, \mathcal{B}, \mu)$ och $X(\omega) = \omega$ och saken är klar.

2. **Skorokhods representationssats:** Vi gör inte detta ordentligt utan beskriver bara idén. Vi ska anta att F är väldigt snäll, nämligen att F är kontinuerlig och *strängt* växande. Detta betyder speciellt att inversen F^{-1} existerar.

   Låt nu $(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}[0, 1], Leb)$. Nu sätter vi helt enkelt $X(\omega) = F^{-1}(\omega)$, $\omega \in [0, 1]$. Då får vi att $\{X \leq c\} = \{\omega : F^{-1}(\omega) \leq c\} = \{\omega : \omega \leq F(c)\} = [0, F(c)]$ och det följer dels att X är mätbar, dels att $P(X \leq c) = Leb[0, F(c)] = F(c)$

precis som önskat. (Att X är mätbar kan man också inse genom att observera att $X = F^{-1}$ är en kontinuerlig funktion och använda sats 3.1(b).)

I det allmänna fallet kan man låta X vara en s.k. generaliserad invers till F. Den brukar betecknas F^- och definieras genom

$$F^-(\omega) = \inf\{x \in \overline{\mathbf{R}} : F(x) \geq \omega\}.$$

Det krävs lite arbete med detaljerna för att visa att X verkligen får rätt fördelning. Detta finns gjort i Williams, avsnitt 3.12.

3.2 σ-algebror genererade av en klass av funktioner

I exemplet ovan med den oändliga följden av slantsinglingar definierade vi σ-algebran $\mathcal{F}$ som den minsta σ-algebra som gjorde alla funktionerna $X_n(\omega) = I_H(\omega_n)$ mätbara. Man skriver

$$\mathcal{F} = \sigma(X_n, n = 1, 2, \ldots).$$

Generellt, om $\{Y_c : c \in C\}$ är en klass av $\overline{\mathbf{R}}$-värda funktioner med gemensamt definitionsrum Ω och C är någon godtycklig indexmängd, kan vi låta

$$\mathcal{Y} = \sigma(Y_c^{-1}(B) : B \in \mathcal{B}, c \in C)$$

vara den minsta σ-algebran som gör alla Y_c mätbara. Vi säger att $\mathcal{Y}$ är *σ-algebran genererad av Y_c:na* och skriver $\mathcal{Y} = \sigma(Y_c : c \in C)$.

Observera att om alla Y_c:na är stokastiska variabler på $(\Omega, \mathcal{F})$ har vi automatiskt att $\mathcal{Y} \subseteq \mathcal{F}$. Observera också att om $\{X_n\}$ är en följd av funktioner på Ω och vi sätter $\chi_n = \sigma(X_1, \ldots, X_n)$ så är $\cup_{n=1}^{\infty} \chi_n$ en algebra som genererar $\sigma(X_1, X_2, \ldots)$. Det är dock inte alls säkert att $\cup_{n=1}^{\infty} \chi_n$ är en σ-algebra. (Varför?)

En intuitiv tolkning av σ-algebran $\mathcal{Y}$ är följande: Ett slumpförsök består i att slumpen på något sätt väljer ett utfall $\omega \in \Omega$. Antag att vi som information om ω får veta alla värdena $Y_c(\omega)$, $c \in C$. Då vet vi alltså om mängderna $\{Y_c \in B\}$ har inträffat eller ej för alla $B \in \mathcal{B}$ (och även för $B \notin \mathcal{B}$ men det anser vi som "ointressant" information). Men $\mathcal{Y} = \sigma(\{Y_c \in B\} : B \in B, c \in C)$ så man kan säga att $\mathcal{Y}$ består av alla mängder vi vet om de inträffat eller ej om vi vet värdena på alla Y_c och använder högst "uppräkneligt mycket information". Kort sagt kan man alltså säga att σ-algebran genererad av en klass av funktioner består av "den information som dessa funktioner ger".

PROPOSITION 3.5 *(a) Låt (S, Σ) vara ett mätbart rum och låt $Y : \Omega \to S$. Det gäller att funktionen $X : \Omega \to \mathbf{R}$ är $\sigma(Y)$-mätbar om och endast om det finns en $\Sigma/\mathcal{B}$-mätbar funktion $f : S \to \mathbf{R}$ sådan att $X = f \circ Y$.*

(b) *Specialfall av (a): Om Y är reellvärd så gäller att X är $\sigma(Y)$-mätbar om och endast om $X = f \circ Y$ för någon Borelfunktion f.*

Bevis. Eftersom (b) är ett specialfall av (a) räcker det att visa (a). Vi ser direkt att om $X = f \circ Y$ och f är Σ-mätbar så följer det att X är $\sigma(Y)$-mätbar av sats 3.2.

Antag nu att X är $\sigma(Y)$-mätbar. Vi ska då hitta funktionen f. Enligt $\sigma(Y)$-mätbarheten hos X gäller det för varje heltal k och varje positivt heltal n att

$$\{X \in [k2^{-n}, (k+1)2^{-n})\} = \{Y \in F_{nk}\}$$

för någon mängd $F_{nk} \in \Sigma$. Definiera nu funktionen f_n genom att sätta $f_n(s) = k2^{-n}$ då $s \in F_{nk}$. Då blir f_n uppenbarligen Σ-mätbar och $f_n \leq f_{n+1}$, dvs följden $\{f_n\}$ är växande. Således existerar funktionen $f = \lim_{n\to\infty} f_n$. Eftersom gränsvärden av följder av mätbara funktioner är mätbara är f mätbar och eftersom $|f_n(Y(\omega)) - X(\omega)| < 2^{-n}$ gäller att $f \circ Y = X$. □

3.3 Monoton-klass-satsen

Kom ihåg Dynkins lemma som hjälper oss att dra slutsatser om hela σ-algebror utifrån resultat för genererande π-system. Monoton-klass-satsen är en utvidgning av Dynkins lemma. Denna utvidgning hjälper oss att dra slutsatser om alla begränsade Σ-mätbara funktioner utifrån resultat för indikatorfunktioner för mängder i ett π-system som genererar Σ.

DEFINITION. Låt S vara något rum och låt $\mathcal{H}$ vara en klass av begränsade reellvärda funktioner definierade på S. Om

(i) $\mathcal{H}$ är ett reellt vektorrum,

(ii) $f \equiv 1 \Rightarrow f \in \mathcal{H}$,

(iii) $f_n \in \mathcal{H}$, $f_n \uparrow f$, f begränsad $\Rightarrow f \in \mathcal{H}$,

så kallas $\mathcal{H}$ för en *monoton klass.*

SATS 3.6 (MONOTON-KLASS-SATSEN) *Låt $\mathcal{H}$ vara en monoton klass och antag att $I_A \in \mathcal{H}$ för alla mängder A i ett π-system $\mathcal{I}$. Då gäller att $\mathcal{H} \supseteq b\sigma(\mathcal{I})$.*

Bevis. Huvudverktyget i detta bevis är Dynkins lemma.

Låt $\mathcal{D}$ vara klassen av mängder $F \subseteq S$ sådana att $I_F \in \mathcal{H}$. Det följer av (i)-(iii) att $\mathcal{D}$ är ett d-system, så enligt Dynkins lemma har vi att $\mathcal{D} \supseteq \sigma(\mathcal{I})$, dvs alla indikatorfunktioner för mängder i $\sigma(\mathcal{I})$ ingår i $\mathcal{H}$.

Tag nu en $\sigma(\mathcal{I})$-mätbar begränsad funktion f och antag att f är ickenegativ så att $0 \leq f < K$ för något heltal $K < \infty$. Sätt för $i = 0, \ldots, K2^n - 1$

$$A_{ni} = \{s \in S : i2^{-n} \leq f(s) < (i+1)2^{-n}\}$$

och låt

$$f_n(s) = \sum_{i=0}^{K2^n-1} i2^{-n} I_{A_{ni}}(s).$$

Enligt vad vi nyss visade gäller att $I_{A_{ni}} \in \mathcal{H}$ för alla n och i så enligt (i) har vi att $f_n \in \mathcal{H}$ och eftersom $f_n \uparrow f$ gäller att $f \in \mathcal{H}$.

(Notera likheten mellan konstruktionen av f_n här och motsvarande konstruktion i beviset av Proposition 3.5.)

Till sist om f är en generell begränsad $\sigma(\mathcal{I})$-mätbar funktion, skriv $f = f^+ - f^-$ där $f^+(s) = \max(f(s), 0)$ och $f^-(s) = \max(-f(s), 0)$. Eftersom vi nu vet att f^+ och f^- tillhör $\mathcal{H}$ så gäller enligt (i) detta också f och vi är klara. $\square$

4 Oberoende

4.1 Definitionen och Borel-Cantellis andra lemma

I hela detta kapitel antas att $(\Omega, \mathcal{F}, P)$ är det sannolikhetsrum vi arbetar med. Du har i tidigare kurser sett definitioner av oberoende händelser och oberoende stokastiska variabler. Vi definierar dessa begrepp i termer av σ-algebror:

DEFINITION.

- Låt $\mathcal{G}_1, \mathcal{G}_2, \ldots$ vara en ändlig eller uppräknelig familj av del-klasser till $\mathcal{F}$. Vi säger att $\mathcal{G}_1, \mathcal{G}_2, \ldots$ är *oberoende* om det gäller för alla $G_1 \in \mathcal{G}_1$, $G_2 \in \mathcal{G}_2, \ldots$ och alla ändliga uppsättningar av distinkta index $i_1, i_2, \ldots, i_n$ att

$$P(\cap_{k=1}^n G_{i_k}) = \prod_{k=1}^n P(G_{i_k}).$$

- De stokastiska variablerna $X_1, X_2, \ldots$ sägs vara oberoende om σ-algebrorna $\sigma(X_1)$, $\sigma(X_2), \ldots$ är oberoende.
- Händelserna $E_1, E_2, \ldots$ sägs vara oberoende om $\sigma(E_1)$, $\sigma(E_2), \ldots$ är oberoende, där $\sigma(E_j) = \sigma(I_{E_j}) = \{\emptyset, \Omega, E_j, E_j^c\}$.

Verifiera gärna att dessa definitioner stämmer överens med vad du tidigare lärt dig.

Även i dessa sammanhang är vi hjälpta av att använda π-system.

LEMMA 4.1 *Låt $\mathcal{G}$ och $\mathcal{H}$ vara två del-σ-algebror till $\mathcal{F}$ genererade av π-systemen $\mathcal{I}$ repektive $\mathcal{J}$. Då gäller att $\mathcal{G}$ och $\mathcal{H}$ är oberoende om och endast om $\mathcal{I}$ och $\mathcal{J}$ är oberoende.*

Bevis. Att oberoende mellan $\mathcal{G}$ och $\mathcal{H}$ implicerar oberoende mellan $\mathcal{I}$ och $\mathcal{J}$ är självklart, så låt oss koncentrera oss på den andra implikationen.

Vi ska använda ett tvåstegsargument som till viss del påminner om det vi gjorde i beviset av Dynkins lemma. Fixera först en händelse $I \in \mathcal{I}$. Definiera två mått μ_1 och μ_2 på $\mathcal{H}$ genom att sätta

$$\mu_1(H) = P(H \cap I)$$

$$\mu_2(H) = P(H)P(I)$$

för alla $H \in \mathcal{H}$. Eftersom $\mathcal{I}$ och $\mathcal{J}$ är oberoende har vi att $\mu_1(H) = \mu_2(H)$ för alla $H \in \mathcal{J}$. Dessutom har vi att $\mu_1(\Omega) = P(\Omega \cap I) = P(\Omega)P(I) = \mu_2(\Omega) \leq 1 < \infty$. Vi kan alltså tillämpa unicitetssatsen som talar om för oss att $\mu_1 \equiv \mu_2$. Eftersom I var godtyckligt vald medför detta att $\mathcal{H}$ och $\mathcal{I}$ är oberoende.

Fixera nu $H \in \mathcal{H}$ och definiera

$$\mu_3(G) = P(G \cap H)$$

$$\mu_4(G) = P(G)P(H)$$

för alla $G \in \mathcal{G}$. Genom ett fullständigt analogt resonemang följer det av unicitetssatsen och det nyss bevisade oberoendet mellan $\mathcal{H}$ och $\mathcal{I}$ att $\mu_3 \equiv \mu_4$ och eftersom H var godtyckligt vald betyder detta att $\mathcal{G}$ och $\mathcal{H}$ är oberoende. $\square$

Övning (Williams 4.1): Antag att $\mathcal{G}_1, \ldots, G_n$ är del-σ-algebror som genereras av π-systemen $\mathcal{I}_1, \mathcal{I}_2, \ldots$ respektive $\mathcal{I}_n$ och antag att $\Omega \in \mathcal{I}_k$ för alla k. Visa att $\mathcal{G}_1, \ldots, \mathcal{G}_n$ är oberoende om och endast om $\mathcal{I}_1, \ldots, \mathcal{I}_n$ är oberoende.

Exempel. Låt X och Y vara två stokastiska variabler. Enligt sats 3.1(c) genereras $\sigma(X)$ och $\sigma(Y)$ av de två π-systemen $\{\{X \leq x\} : x \in \overline{\mathbf{R}}\}$ respektive $\{\{Y \leq y\} : y \in \overline{\mathbf{R}}\}$. Enligt Lemma 4.1 gäller alltså att X och Y är oberoende om och endast om

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

för alla $x, y \in \overline{\mathbf{R}}$.

I övningen ovan utvidgas Lemma 4.1 till n stycken σ-algebror. Detta kan vi använda för att på samma sätt visa att $X_1, X_2, \ldots$ är oberoende om och endast om

$$P(\cap_{j=1}^{k}\{X_{i_j} \leq x_{i_j}\}) = \prod_{j=1}^{k} P(X_{i_j} \leq x_{i_j})$$

för alla ändliga indexdelmängder och alla x_{i_j}:n.

Sats 4.2 (Borel-Cantelli II) *Om $\{E_n\}$ är en följd av* oberoende *händelser sådana att $\sum_{n=1}^{\infty} P(E_n) = \infty$ gäller att $P(\limsup_n E_n) = 1$.*

Bevis. Sätt $p_n = P(E_n)$. Observera att

$$(\limsup_n E_n)^c = (\cap_n \cup_{m \geq n} E_m)^c = \cup_n \cap_{m \geq n} E_m^c.$$

Vi ska alltså visa att $P(\cup_n \cap_{m \geq n} E_m^c) = 0$ vilket är det samma som att $P(\cap_{m \geq n} E_m^c) = 0$ för alla n.

Enligt kontinuitet hos mått och oberoendet har vi att

$$P(\cap_{m \geq n} E_m^c) = \lim_{r \to \infty} P(\cap_{m=n}^{r} E_m^c) = \lim_{r \to \infty} \prod_{m=n}^{r} P(E_m^c) = \prod_{m=n}^{\infty} (1 - p_m).$$

Vi utnyttjar nu olikheten $1 - x \leq e^{-x}$ och får att

$$P(\cap_{m \geq n} E_m^c) \leq \prod_{m=n}^{\infty} e^{-p_m} = e^{-\sum_{m=n}^{\infty} p_m} = e^{-\infty} = 0.$$

□

Exempel. Låt $X_1, X_2, \ldots$ vara en iid-följd av stokastiska variabler som har en exponentialfördelning med parameter $\lambda = 1$. Då har vi att $P(X_n > x) = e^{-x}$. Vi ska visa att $\limsup_n (X_n/\log n) = 1$ n.s. I ord betyder detta att rekordvärdet vid tid n för stort n med mycket stor sannolikhet kommer att ligga mycket nära $\log n$.

För ett positivt heltal α har vi att

$$P(X_n/\log n > \alpha) = P(X_n > \alpha \log n) = n^{-\alpha}.$$

Eftersom $\sum_{n=1}^{\infty} n^{-\alpha}$ är ändlig då $\alpha > 1$ och oändlig då $\alpha \leq 1$ säger Borel-Cantelli I och II att $P(X_n/\log n > \alpha\, o.o.)$ blir 0 då $\alpha > 1$ och 1 då $\alpha \leq 1$. Här (och i fortsättningen) använder vi *o.o.* som en förkortning av "oändligt ofta". Händelsen $(E_n\, o.o.)$ är alltså $\limsup_n E_n$.

Som konsekvens får vi att

$$P(\limsup_n (X_n/\log n) \geq 1) \geq P(X_n > \log n\, o.o.) = 1$$

och att

$$P(\limsup_n (X_n/\log n) > 1) = P(\cup_k \{\limsup_n (X_n/\log n) > 1 + 1/k\})$$

$$\leq \sum_k P(\limsup_n (X_n/\log n) > 1 + 1/k)$$

$$\leq \sum_k P(X_n/\log n > 1 + 1/(2k)\, o.o.) = 0.$$

4.2 Konstruktion av sannolikhetsteoretiska modeller

Vi har nu talat om följder av oberoende stokastiska variabler med givna fördelningsfunktioner definierade på något gemensamt sannolikhetsrum $(\Omega, \mathcal{F}, P)$. Återigen infinner sig frågan om och i så fall hur man konstruerar ett sådant sannolikhetsrum.

Även här kan vi utgå från Lebesguemåttet på $\mathcal{B}[0,1]$ som vi vet existerar. Sätt $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}[0,1], Leb)$.

För varje $\omega \in [0,1]$ skriver vi $0.\omega_1\omega_2\omega_3\ldots$ för dess binärutveckling. (Här väljer man alltid utvecklingar av typen $0.10000\ldots$ framför $0.01111\ldots$.) Låt nu $\{\alpha_{11}, \alpha_{12}, \alpha_{13}, \ldots\}$, $\{\alpha_{21}, \alpha_{22}, \alpha_{23}, \ldots\}$, ... vara *disjunkta* delföljder av de positiva heltalen och sätt $X_i(\omega) = 0.\omega_{\alpha_{i1}}\omega_{\alpha_{i2}}\ldots$, $i = 1, 2, \ldots$. Eftersom olika X_i:n använder olika koordinater och dessa är sinsemellan oberoende är X_i:na oberoende och det är intuitivt klart att X_i har en likformig fördelning på [0,1] för varje i. (Verifiera gärna dessa påståenden rigoröst.)

Om vi nu vill att Y_i ska ha fördelninsfunktion F_i använder vi Skorokhods representationssats och låter $Y_i = F_i^{-}(X_i)$.

4.3 Konstruktion av Markovkedjor

En *stokastisk process* indexerad av en mängd C är en familj $\{Y_c : c \in C\}$ av stokastiska variabler definierade på ett sannolikhetsrum $(\Omega, \mathcal{F}, P)$.

Existensen av en generell stokastisk process gör vi inte i denna kurs, men vi ska göra ett specialfall, nämligen konstruktionen av en Markovkedja. I det fallet är $C = \{1, 2, \ldots\}$ och vi antar att tillståndsrummet, E, för Markovkedjan vi vill konstruera är ändligt eller uppräkneligt. Vi vill alltså konstruera en process som

- startar i tillstånd i, $i \in E$, med sannolikhet $\mu(i)$, där μ är något givet sannolikhetsmått på E,
- hoppar från i till j med sannolikhet p_{ij}, där $[p_{ij}]_{i,j \in E}$ är någon given övergångsmatris, oberoende av vad som hänt tidigare.

Om vi låter Z_n beteckna kedjans tillstånd vid tidpunkt n betyder detta att vi vill att $P(Z_0 = i_0, Z_1 = i_1, \ldots, Z_n = i_n) = \mu_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}$ för alla $i_0, \ldots, i_n$ och alla n.

Men vi klarar ju av att konstruera ett sannolikhetsrum med en uppräknelig uppsättning av obeorende stokastiska variabler med önskad fördelning. Låt nu Z_0 ha fördelning μ_0 och låt för varje $i \in E$ och varje n, Y_{in} vara en stokastisk variabel sådan att $P(Y_{in} = j) = p_{ij}$. Sätt nu bara $Z_1 = Y_{Z_0 1}, Z_2 = Y_{Z_1 2}, Z_3 = Y_{Z_2 3}, \ldots$ och saken är klar.

4.4 Kolmogorovs 0-1-lag

Antag att $X_1, X_2, \ldots$ är en följd av oberoende stokastiska variabler. Vi ska i detta avsnitt bevisa ett resultat som säger att en händelse som inte påverkas av värdet på ändligt många X_i:n måste ha sannolikhet 0 eller 1. Exempel på sådana händelser är

- $\{\lim_{n\to\infty} X_n$ existerar$\}$,
- $\{\sum_n X_n$ konvergent$\}$,
- $\{\lim_{n\to\infty} \sum_{k=1}^n X_k/n$ existerar$\}$.

Denna typ av händelser kallas *svanshändelser*, dvs händelser som tillhör *svans-σ-algebran* till X_i:na. Den formella definitionen är följande:

DEFINITION. Låt $X_1, X_2, \ldots$ vara en följd av stokastiska variabler och låt $\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$. Sätt $\mathcal{T} = \cap_n \mathcal{T}_n$. Då kallas $\mathcal{T}$ för svans-σ-algebran till följden $\{X_n\}$.

En stokastisk variabel som är $\mathcal{T}$-mätbar kallas för en svansfunktion och är alltså sådan att den inte påverkas av värdena av ändligt många X_i:n. Exempel är $\limsup_n X_n$, $\liminf_n X_n$ och $\limsup_n (\sum_{k=1}^n X_k/n)$.

SATS 4.3 (KOLMOGOROVS 0-1-LAG) *Låt $X_1, X_2, \ldots$ vara en följd av* oberoende *stokastiska variabler och låt $\mathcal{T}$ vara svans-σ-algebran till $\{X_n\}$. Det gäller att*

(i) $F \in \mathcal{T} \Rightarrow P(F) \in \{0,1\}$,

(ii) om ξ är en svansfunktion är ξ n.s. konstant.

Bevis. Vi börjar med del (i). Låt $\chi_n = \sigma(X_1, \ldots, X_n)$ och $\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$. Vi har tidigare sett att om vi låter $\mathcal{K} = \{\{X_1 \leq x_1, \ldots, X_n \leq x_n\} : x_i \in \overline{\mathbf{R}}, i = 1, \ldots, n\}$ så är $\mathcal{K}$ ett π-system som genererar χ_n. På motsvarande sätt gäller att klassen $\mathcal{J} = \{\{X_{n+1} \leq x_{n+1}, \ldots, X_{n+r} \leq x_{n+r}\} : x_j \in \overline{\mathbf{R}}, j = n+1, \ldots, n+r, r = 1, 2, \ldots\}$ är ett π-system som genererar $\mathcal{T}_n$. Eftersom X_i:na är oberoende följer direkt att $\mathcal{K}$ och $\mathcal{J}$ är oberoende och enligt Lemma 4.1 är då även χ_n och $\mathcal{T}_n$ oberoende.

Eftersom $\mathcal{T} \subseteq \mathcal{T}_n$ följer det nu automatiskt att χ_n och $\mathcal{T}$ är oberoende.

Sätt $\chi_\infty = \sigma(X_1, X_2, \ldots) = \sigma(\cup_n \chi_n)$. Eftersom χ_n är oberoende av $\mathcal{T}$ för varje n har vi att $\cup_n \chi_n$ är oberoende av $\mathcal{T}$. Men $\cup_n \chi_n$ är ju en algebra och därmed ett π-system och det följer därför från Lemma 4.1 att χ_∞ och $\mathcal{T}$ är oberoende.

Genom att slutligen observera att $\mathcal{T} \subseteq \chi_\infty$ ser man att $\mathcal{T}$ är oberoende av sig själv, vilket per definition betyder att

$$F \in \mathcal{T} \Rightarrow P(F) = P(F \cap F) = P(F)^2$$

vilket i sin tur betyder att $P(F)$ är 0 eller 1 för alla $F \in \mathcal{T}$.

För att bevisa att ξ är n.s. konstant, låt $c = \sup\{x : P(\xi \leq x) = 0\}$. Enligt (i) är $P(\xi \leq x)$ 0 eller 1. Om nu $c = \infty$ följer direkt att $P(\xi = \infty) = 1$ och likadant om $c = -\infty$ att $P(\xi = -\infty) = 1$. Om $c \in \mathbf{R}$ har vi enligt definitionen av c att $P(\xi \leq c - 1/n) = 0$ för alla c och således att $P(\xi < c) = P(\cup_n \{\xi \leq c - 1/n\}) = 0$, och $P(\xi \leq c + 1/n) = 1$ för alla n så att $P(\xi \leq c) = P(\cap_n \{\xi \leq c + 1/n\}) = 1$. Alltså gäller $P(\xi = c) = 1$. □

Exempel. Apa som skriver Shakespeare: Antag att en apa sitter vid en skrivmaskin och skriver ett på måfå valt tecken varje sekund. Vi ska se att förr eller senare kommer apan (om den har ett oändligt liv och aldrig tröttnar på att skriva maskin) att skriva Shakespeares samlade verk.

Låt $X_1, X_2, \ldots$ vara tecknen som apan skriver. Vi antar att dessa är oberoende stokastiska variabler som är likformigt fördelade på mängden av, säg, N möjliga tecken. Låt H vara händelsen att apan med sin oändliga följd av tecken skriver Shakespeares samlade verk inte bara en utan oändligt många gånger. Då är H en svanshändelse så enligt Kolmogorovs 0-1-lag gäller att $P(H)$ är 0 eller 1.

Om nu Shakespeares samlade verk omfattar M tecken så är ju chansen minst $p = 1/N^M$ att apan lyckas med en gång. Likadant är ju chansen p att apan lyckas i tidsintervallet $\{M+1, \ldots, 2M\}$ och i intervallet $\{2M+1, \ldots, 3M\}$ osv. Eftersom dessa intervall är disjunkta har vi oberoende mellan vad som händer i de olika intervallen och eftersom $\sum_{i=1}^{\infty} p = \infty$ talar Borel-Cantelli II om för oss att $P(H) = 1$.

(Observera att Kolmogorov 0-1-lag bara talade om för oss att P(H) är 0 eller 1. För att avgöra vilket var vi tvungna att ta till andra medel för att visa att $P(H) > 0$ så att $P(H) = 1$. I just detta fall talade dock Borel-Cantelli II om direkt för oss att $P(H) = 1$ så vi behövde egentligen inte Kolmogorovs 0-1-lag alls.)

Observera att Kolmogorovs 0-1-lag kräver oberoende. Den är falsk annars. Ett exempel på det får man från en superkritisk Galton-Watson-process. Låt Z_n beteckna generationsstorleken för generation n i en sådan. Sätt $\mu = \mathbf{E}[Z_1]$. Från tidigare kurser känner vi till att $M_\infty = \lim_{n\to\infty}(Z_n/\mu^n)$ existerar n.s. och har en icketrivial fördelning. Det är ändå dock så att M_∞ är en svansfunktion till $\{Z_n\}$. Alltså ser vi att (ii) och således även (i) i Kolmogorovs 0-1-lag spricker i avsaknad av oberoende.

5 Integraler

Vi ska här definiera den s.k. *Lebesgueintegralen*, som är den integral som "gäller" i den moderna matematiken.

Låt oss först påminna oss om hur man definierar Riemannintegralen, dvs den integral som man lär sig om på gymnasiet och i grundutbildningen.

5.1 Riemannintegralen

Låt $f : [a, b) \to \mathbf{R}$ vara en funktion definierad på ett intervall av den reella linjen. Vi kallar en funktion h på $[a, b)$ för *styckvis konstant* om om det finns en partition av $[a, b)$ i delintervall $[t_0, t_1), [t_1, t_2), \ldots, [t_{n-1}, t_n)$ (där $t_0 = a$ och $t_n = b$) sådan att h är konstant på alla delintervall. Vi kallar G för en *överfunktion* till f om G är styckvis konstant och $G(x) \geq f(x)$ för alla $x \in [a, b)$. En styckvis konstant funktion g sådan att $g(x) \leq f(x)$ för alla $x \in [a, b)$ kallas för en *underfunktion* till f.

Riemannintegralen av en styckvis konstant funktion h definieras naturligt genom att man sätter $\int_a^b h(x)dx = \sum_{i=1}^n h(t_{i-1})(t_i - t_{i-1})$. För f sätter vi nu

$$\overline{I}(f) = \inf\{\int_a^b G(x)dx : G \text{ överfunktion till } f\}$$

och

$$\underline{I}(f) = \inf\{\int_a^b g(x)dx : g \text{ underfunktion till } f\}.$$

Om $\overline{I}(f) = \underline{I}(f) = I(f)$ säger vi att Riemannintegralen för f existerar och vi sätter $\int_a^b f(x)dx = I(f)$.

5.2 Lebesgueintegralen

Lebesgueintegralen är en generalisering på två sätt; dels kan definitionsrummet $[a, b)$ bytas ut mot vilket måttrum (S, Σ, μ) som helst, dels frångår vi inskränkningen att dela upp definitionsrummet i intervall (vad nu det skulle betyda) utan delar istället upp i godtyckliga partitioner av mätbara mängder.

(Vi uppfattar $[a, b)$ som måttrummet $([a, b), \mathcal{L}[a, b), Leb)$. Här är $\mathcal{L}[a, b)$ den sk *Lebesgue-σ-algebran*. Denna innehåller förutom alla mängder i $\mathcal{B}$ också alla delmängder till mängder i $\mathcal{B}$ med Lebesguemått 0 och unioner av par av mängder av de två typerna, dvs $\mathcal{L} = \{A \cup B : B \in \mathcal{B}, \exists C \in \mathcal{B} : Leb(C) = 0, A \subseteq C\}$. Att verifiera att $\mathcal{L}$ är en σ-algebra är en övning som kräver en viss list. Gör gärna detta själv. När det är gjort kan man utvidga Lebesguemåttet från $\mathcal{B}$ till $\mathcal{L}$ genom att sätta $Leb(A \cup B) = Leb(B)$ för alla $A \cup B$ av ovan angivna typ. Rummet $([a, b), \mathcal{L}[a, b), Leb)$ är ett exempel på ett s.k. *fullständigt* måttrum, se övningarna.)

Definitionen av Lebesgueintegralen görs i tre steg:

1. För *enkla funktioner* (vilka är motsvarigheten till de styckvis konstanta funktionerna).
2. För *ickenegativa* mätbara funktioner.
3. För generella mätbara funktioner.

Låt oss börja. Vi har en funktion $f : (S, \Sigma, \mu) \to (\overline{\mathbf{R}}, \mathcal{B})$, dvs en Σ-mätbar funktion, och vi vill skapa en definition av integralen $\int_S f(s)\mu(ds)$ vilket också ibland skrivs $\int_S f(s)d\mu(s)$ eller, enklare, $\int_S f d\mu$.

1. Enkla funktioner:

Antag att $\phi : (S, \Sigma, \mu) \to (\overline{\mathbf{R}}, \mathcal{B})$ kan skrivas på formen

$$\phi(s) = \sum_{i=1}^{n} x_i I_{A_i}(s)$$

där $x_i \in [0, \infty)$ och $\{A_1, \ldots, A_n\}$ är en partition av S sådan att $A_i \in \Sigma$ för alla i. Då säger vi att ϕ är en *enkel* funktion. Observera att en enkel funktion automatiskt är mätbar, ickenegativ och ändlig. Integralen av en enkel funktion ϕ definierar vi som

$$\int_S \phi d\mu = \sum_{i=1}^{n} x_i \mu(A_i).$$

Exempel. Redan här ser man att Lebesgueintegralen klarar vissa fall som Riemannintegralen inte klarar. Låt t.ex. $(S, \Sigma, \mu) = ([0,1], \mathcal{L}, Leb)$ och låt $\phi(s) = 0 \cdot I_{\mathbf{Q}}(s) + 1 \cdot I_{\mathbf{Q}^c}(s) = I_{\mathbf{Q}^c}(s)$, dvs indikatorfunktionen för att s är irrationellt. Då följer det direkt av definitionen att $\int_S \phi d\mu = Leb(\mathbf{Q}^c) = 1$, medan Riemannintegralen är odefinierad eftersom alla överfunktioner måste vara minst 1 och alla underfunktioner kan vara högst 0.

SATS 5.1 *Låt ϕ och ψ vara enkla funktioner på (S, Σ, μ) och låt a och b vara två ickenegativa konstanter.*

(a) $\int_S (a\phi + b\psi) d\mu = a \int_S \phi d\mu + b \int_S \psi d\mu$,

(b) $\phi \le \psi \Rightarrow \int_S \phi d\mu \le \int_S \psi d\mu$,

(c) $\phi = \psi$ *n.ö.* $\Rightarrow \int_S \phi d\mu = \int_S \psi d\mu$.

Bevis. Vi gör bara del (b) och lämnar (a) och (c) som övning. Vi har att $\phi = \sum_{i=1}^n x_i I_{A_i}$ och $\psi = \sum_{j=1}^m y_j I_{B_j}$ där x_i:na och y_j:na är ickenegativa reella tal och $\{A_i\}$ och $\{B_j\}$ är två partitioner av S sådana att $A_i \in \Sigma$ och $B_j \in \Sigma$. Tricket är att skapa en gemensam partition för ϕ och ψ genom att helt enkelt sätta $C_{ij} = A_i \cap B_j$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. (Vissa C_{ij} kan bli tomma, men det spelar ingen roll.) Vi kan nu skriva $\phi = \sum_{i,j} x_i I_{C_{ij}}$ och $\psi = \sum_{i,j} y_j I_{C_{ij}}$. Eftersom $\phi \le \psi$ har vi att $x_i \le y_j$ för alla i, j sådana att C_{ij} är icketom och vi får

$$\int_S \phi d\mu = \sum_{i,j} x_i \mu(C_{ij}) \le \sum_{i,j} y_j \mu(C_{ij}) = \int_S \psi d\mu.$$

□

DEFINITION. Låt ϕ vara en enkel funktion och låt $A \in \Sigma$. Vi definierar integralen av ϕ över A som

$$\int_A \phi d\mu = \int_S \phi I_A d\mu.$$

Observera att eftersom ϕ är enkel är givetvis också ϕI_A enkel och således är $\int_A \phi d\mu$ väldefinierad.

PROPOSITION 5.2 *Låt ϕ vara en enkel funktion och definiera mängdfunktionen γ genom att sätta $\gamma(A) = \int_A \phi d\mu$, $A \in \Sigma$. Då är γ ett mått.*

Bevis. Att γ är ickenegativ och att $\gamma(\emptyset) = 0$ är uppenbart så det gäller att visa att om $A_1, A_2, \ldots$ är disjunkta mängder i Σ så är $\gamma(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \gamma(A_i)$. Men eftersom vi kan skriva $\phi = \sum_{k=1}^{n} x_i I_{E_k}$ för någon partition $\{E_k\}_{k=1}^{n}$ har vi

$$\gamma(\cup_{i=1}^{\infty} A_i) = \int_S \phi I_{\cup_{i=1}^{\infty} A_i} d\mu = \int_S (\sum_{k=1}^{n} x_k I_{\cup_{i=1}^{\infty}(E_k \cap A_i)}) d\mu = \sum_{k=1}^{n} x_k \sum_{i=1}^{\infty} \mu(E_k \cap A_i)$$

$$= \sum_{i=1}^{\infty} \sum_{k=1}^{n} x_k \mu(E_k \cap A_i) = \sum_{i=1}^{\infty} \int_{A_i} \phi d\mu = \sum_{i=1}^{\infty} \gamma(A_i).$$

□

2. Ickenegativa funktioner

När man definierar integralen för en ickenegativ funktion gör man det i ljuset av följande resultat som säger att en ickenegativ mätbar funktion kan approximeras godtyckligt väl underifrån av enkla funktioner.

Sats 5.3 *Låt f vara en ickenegativ Σ-mätbar funktion. Då existerar en följd $\{\phi_n\}$ av enkla Σ-mätbara funktioner sådana att $\phi_n \uparrow f$ (dvs $\phi_n(s) \uparrow f(s)$ för alla $s \in S$).*

Bevis. Vi har vid det här laget sett varianter av följande konstruktion vid två tillfällen tidigare: För varje $n = 1, 2, \ldots$, sätt för $k = 0, 1, 2, \ldots, n2^n - 1$

$$A_{n,k} = \{s \in S : k2^{-n} \leq f(s) < (k+1)2^{-n}\}$$

och sätt $A_{n,n2^n} = \{s \in S : f(s) \geq n\}$. Eftersom f är mätbar har vi att $A_{n,k}$:na är mätbara mängder. Låt nu

$$\phi_n(s) = \sum_{k=0}^{n2^n} k2^{-n} I_{A_{n,k}}(s).$$

Då är ϕ_n uppenbarligen enkel och mätbar och det följer av konstruktionen av ϕ_n att $\phi_n(s) \uparrow f(s)$ för alla s. □

I ljuset av ovanstående resultat är det naturligt att tänka sig att man ska ha $\int_S f d\mu = \lim_{n\to\infty} \int_S \phi_n d\mu$. Nu är ju inte följden $\{\phi_n\}$ den enda följd som approximerar f och eftersom den ena följden inte är "bättre" än den andra använder man följande definition:

Definition. Låt $f : (S, \Sigma, \mu) \to (\overline{\mathbf{R}}, \mathcal{B})$ vara ickenegativ. Integralen av f ges av

$$\int_S f d\mu = \sup\{\int_S \phi d\mu : \phi \text{ enkel }, \phi \leq f\}.$$

SATS 5.4 (MONOTON-KONVERGENS-SATSEN) *Låt $\{f_n\}$ vara en följd av funktioner i $(m\Sigma)^+$ sådana att $f_n \uparrow f$. Då gäller att*

$$\int_S f_n d\mu \uparrow \int_S f d\mu.$$

Bevis. Som en direkt följd av definitionen av integralen för en ickenegativ funktion och Sats 5.1(b) gäller att $\int_S f_n d\mu \leq \int_S f d\mu$ för alla n och att $\int_S f_n d\mu$ är växande i n. Således räcker det att visa att $\sup_n \int_S f_n d\mu \geq \int_S f d\mu$.

Fixera nu ett godtyckligt tal $\alpha \in (0,1)$ och en godtycklig enkel funktion ϕ sådan att $\phi \leq f$. Eftersom $f_n(s) \uparrow f(s)$ för alla s gäller det för alla s att $f_n(s) \geq \alpha\phi(s)$ för alla n som är tillräckligt stora. Med andra ord gäller att om vi låter

$$A_n = \{s \in S : f_n(s) \geq \alpha\phi(s)\}$$

har vi att $A_n \uparrow S$. Detta medför att

$$\int_S f_n d\mu \geq \int_{A_n} f_n d\mu \geq \alpha \int_{A_n} \phi d\mu.$$

Kom nu ihåg Proposition 5.2 som sade att $\int_A \phi d\mu$ som funktion av A är ett mått. Enligt kontinuitet hos mått gäller alltså att högerledet konvergerar mot $\alpha \int_S \phi d\mu$ då $n \to \infty$. Vi har alltså visat att $\sup_n \int_S f_n d\mu \geq \alpha \int_S \phi d\mu$. Men α var ju godtycklig så olikheten kvarstår när vi låter $\alpha \uparrow 1$. Till sist konstaterar vi att eftersom även ϕ var godtycklig gäller enligt definitionen av $\int_S f d\mu$ att $\sup_n \int_S f_n d\mu \geq \int_S f d\mu$. $\square$

KOROLLARIUM 5.5 *(a) Om $f, g \in (m\Sigma)^+$ och $f = g$ n.ö. gäller att $\int_S f d\mu = \int_S g d\mu$.*

(b) Om $f_n \uparrow f$ n.ö. (dvs $\exists N \in \Sigma : \mu(N) = 0, f_n(s) \uparrow f(s)$ för alla $s \in N^c$) gäller att $\int_S f_n d\mu \uparrow \int_S f d\mu$.

Bevis.

(a) Det räcker att visa att om ϕ är en godtycklig enkel funktion sådan att $\phi \leq f$ gäller att $\int_S \phi d\mu \leq \int_S g d\mu$. Sätt nu $N = \{s \in S : \phi(s) > g(s)\}$. Enligt förutsättning är $\mu(N) = 0$ så om vi sätter $\psi = \phi I_{N^c}$ så är ψ en enkel funktion sådan att $\psi = \phi$ n.ö. och $\psi \leq g$ och vi har enligt Sats 5.1(c) att $\int_S \phi d\mu = \int_S \psi d\mu \leq \int_S g d\mu$.

(b) Eftersom $f_n I_{N^c} \uparrow f I_{N^c}$ kan vi använda (a) tillsammans med monoton-konvergens-satsen.

□

Korollariet ovan tar vi fortsättningen som "självklart", dvs vi kommer att i vissa bevis att låtsas att "nästan överallt" betyder "överallt" utan att be om ursäkt för det.

SATS 5.6 *Tag* $f, g \in (m\Sigma)^+$ *och* $a, b \in [0, \infty]$. *Då gäller att* $\int_S (af + bg) d\mu = a \int_S f d\mu + b \int_S g d\mu$.

Bevis. Enligt Sats 5.3 kan vi hitta följder $\{\phi_n\}$ och $\{\psi_n\}$ av enkla funktioner sådana att $\phi_n \uparrow f$ och $\psi_n \uparrow g$. Eftersom vi redan vet att $\int_S (a\phi + b\psi) d\mu = a \int_S \phi_n d\mu + b \int_S \psi_n d\mu$ för alla n så följer resultatet av monoton-konvergens-satsen. □

I ljuset av monoton-konvergens-satsen är det naturligt att undra om det inte räcker att $f_n \to f$ för att $\int_S f_n d\mu$ ska konvergera mot $\int_S f d\mu$. Detta är *inte* generellt sant. Sätt t.ex. $(S, \Sigma, \mu) = (\mathbf{R}, \mathcal{B}, Leb)$ och sätt $f_n = I_{[n,\infty)}$. Då gäller att $f_n \to 0$ men $\int_S f_n d\mu = \infty$ för alla n. Vad som däremot gäller är följande viktiga resultat.

SATS 5.7 (FATOUS LEMMA) *Låt* $\{f_n\}$ *vara en följd av ickenegativa* Σ*-mätbara funktioner. Då har vi att*

$$\int_S (\liminf_n f_n) d\mu \leq \liminf_n \int_S f_n d\mu.$$

Bevis.

$$\int_S (\liminf_n f_n) d\mu = \int_S \sup_n (\inf_{m \geq n} f_m) d\mu = \sup_n \int_S (\inf_{m \geq n} f_m) d\mu$$

$$\leq \sup_n \inf_{m \geq n} \int_S f_n d\mu = \liminf_n \int_S f_n d\mu$$

där den andra likheten följer av monoton-konvergens-satsen. □

Det finns också en omvänd version av Fatous lemma, men den *kräver att följden* $\{f_n\}$ *är dominerad*:

Om det finns en funktion $g \in (m\Sigma)^+$ sådan att $g \geq f_n$ för alla n och $\int_S g d\mu < \infty$ gäller att

$$\int_S (\limsup_n f_n) d\mu \geq \limsup_n \int_S f_n d\mu.$$

Beviset klarar man genom att tillämpa Fatous lemma på följden $\{g - f_n\}$. Man måste då använda att $\int_S (g - f_n) d\mu = \int_S g d\mu - \int_S f_n d\mu$ vilket följer av Sats 5.6 genom att skriva g som $(g - f_n) + f_n$.

3. Generella mätbara funktioner

För en generell funktion $f \in m\Sigma$ kan vi skriva

$$f = f^+ - f^-$$

där $f^+(s) = \max(f(s), 0)$ och $f^-(s) = \max(-f(s), 0)$. Eftersom f^+ och f^- är ickenegativa är deras integraler definierade enligt vad vi gjort tidigare. Om $\int_S f^+ d\mu < \infty$ och $\int_S f^- d\mu < \infty$ säger vi att f är *μ-integrerbar* och vi *definierar*

$$\int_S f d\mu = \int_S f^+ d\mu - \int_S f^- d\mu.$$

Familjen av alla μ-integrerbara funktioner $f \in m\Sigma$ betecknas med $L^1(S, \Sigma, \mu)$. Oftast skrivs detta på någon kortform där man bara specificerar det man behöver utan att riskera att missuppfattas, t.ex. $L^1(S)$, $L^1(\mu)$, $L^1(\Sigma)$ eller helt enkelt L^1.

Vi observerar omedelbart att eftersom $|f| = f^+ + f^-$ så har vi att $|\int_S f d\mu| \leq \int_S |f| d\mu$. Vi ser också att integralen är linjär. Detta följer omedelbart från linjariteten hos integraler av ickenegativa funktioner, Sats 5.6.

Sats 5.8 (Dominerad-konvergens-satsen) *Låt $\{f_n\}$ vara en följd i $L^1(S, \Sigma, \mu)$ sådan att $f_n \to f$ (punktvis) och antag att det finns en funktion $g \in (m\Sigma)^+$ sådan att $\int_S g d\mu < \infty$ och $|f_n| \leq g$ för alla n. Då gäller att*

$$\int_S |f_n - f| d\mu \to 0$$

då $n \to \infty$.

Omedelbar följd: $\int_S f_n d\mu \to \int_S f d\mu$, ty $|\int_S f_n d\mu - \int_S f d\mu| = |\int_S (f_n - f) d\mu| \leq \int_S |f_n - f| d\mu$.

Bevis. Enligt triangelolikheten har vi att $|f_n - f| \leq |f_n| + |f| \leq 2g$ så vi kan tillämpa omvända Fatous lemma och får att

$$\limsup_n \int_S |f_n - f| d\mu \leq \int_S \limsup_n |f_n - f| d\mu = 0.$$

□

Vi har redan sett exempel på att domineringsvillkoret i dominerad-konvergens-satsen är nödvändigt. Vi hade $(S, \Sigma, \mu) = (\mathbf{R}, \mathcal{B}, Leb)$ och $f_n = I_{[n,\infty)}$ och fick $f_n \to 0$ men $\int_S f_n d\mu = \infty$ för alla n. En variant får man om man istället sätter $f_n = I_{[n,n+1]}$. Då har vi $\int_S f_n d\mu = 1$ för alla n trots att $f_n \to 0$.

Exempel. Vi ska i detta exempel visa att Lebesgueintegralen är "bättre" än Riemannintegralen i den mening att varje Riemannintegrerbar funktion också har en Lebesgueintegral som sammanfaller med dess Riemannintegral.

Vi ska alltså ha $(S, \Sigma, \mu) = ([a, b], \mathcal{L}[a, b], Leb)$. för reella tal a och b där $a < b$. Vi förlorar ingen generalitet på att anta att $a = 0$ och $b = 1$ så låt oss göra det. Antag att f är en Riemannintegrerbar funktion definierad på [0,1] och sätt

$$I(f) = \int_0^1 f(x)dx,$$

dvs Riemannintregralen av f. Eftersom f är Riemannintegrerbar kan vi per definition hitta en följd $\{G_n\}$ av överfunktioner till f och en följd $\{g_n\}$ av underfunktioner till f sådana att $I(G_n) \downarrow I(f)$ och $I(g_n) \uparrow I(f)$.

Genom att ta partiella minima respektive maxima kan vi också anta att G_n är avtagande i n och att g_n är växande i n. Således existerar gränsfunktionerna G och g givna av $G(s) = \lim_{n\to\infty} G_n(s)$ och $g(s) = \lim_{n\to\infty} g_n(s)$. Vi noterar att $g \leq f \leq G$.

Eftersom G_n och g_n är styckvis konstanta är de uppenbarligen $\mathcal{B}$-mätbara och eftersom mätbarhet bevaras under gränsbildning är G och g $\mathcal{B}$-mätbara.

Kom ihåg att en Riemannintegrerbar funktion alltid är begränsad så vi kan anta att G_1 är begränsad och vi kan alltså tillämpa dominerad-konvergens-satsen och får

$$I(G_n) = \int_S G_n d\mu \to \int_S G d\mu$$

och eftersom $I(G_n) \to I(f)$ har vi att $\int_S G d\mu = I(f)$. Genom att tillämpa monotonkonvergens-satsen får vi på samma sätt att

$$I(g_n) = \int_S g_n d\mu \to \int_S g d\mu$$

och att $\int_S g d\mu = I(f)$. Vi har alltså att $I(f) = \int_S G d\mu = \int_S g d\mu$ och $g \leq f \leq G$ så det enda som återstår är att visa att f är $\mathcal{L}$-mätbar så att $\int_S f d\mu$ existerar.

Men $\{s : f(s) \leq x\} = \{s : G(s) \leq x\} \cup A$, där A är någon mängd sådan att $A \subseteq \{s : G(s) \neq g(s)\}$. Eftersom $G - g \geq 0$ och $\int_S (G - g)d\mu = 0$ följer av Sats 5.10 nedan att $G - g = 0$ n.ö. dvs $G = g$ n.ö. Således är $\mu\{s : G(s) \neq g(s)\} = 0$ så enligt definition av $\mathcal{L}$ gäller att $A \in \mathcal{L}$ och det följer av Sats 3.1(b) att f är $\mathcal{L}$-mätbar.

5.3 Standardmaskineriet och Radon-Nikodyms sats

Tag en funktion $f \in (m\Sigma)^+$ och definiera mängdfunktionen γ på Σ genom att sätta $\gamma(A) = \int_A f d\mu$, $A \in \Sigma$. Vi känner till att om f är enkel så är γ ett mått. Genom att approximera f med enkla funktioner enligt Sats 5.3 och använda monoton-konvergens-satsen bevisar man lätt att γ är ett mått även då f är en godtycklig ickenegativ mätbar funktion.

Intuitivt resonerat gäller på "infinitesimal" nivå att $\gamma(ds) = f(s)\mu(ds)$. Om nu h är en funktion i $L^1(S, \Sigma, \gamma)$ så gäller ju då att $h(s)\gamma(ds) = h(s)f(s)\mu(ds)$ så det verkar rimligt att tro att $\int_S h d\gamma = \int_S h f d\mu$. Detta är ett exempel på ett resultat som man kan visa enligt följande arbetsschema som Williams kallar för *standardmaskineriet*:

- Bevisa i fallet då h är en indikatorfunktion.
- Använd linjaritet för att bevisa för godtyckliga enkla funktioner.
- Använd monoton-konvergens-satsen för att bevisa för godtyckliga funktioner i $(m\Sigma)^+$.
- Använd linjaritet för att bevisa för godtyckliga funktioner i L^1.

Ta det som en övning att applicera standardmaskineriet på detta exempel. (Alternativt kan man använda monoton-klass-satsen. Gör gärna även detta som övning.)

Måttet γ ovan är definierat så att $\gamma(A) = \int_A f d\mu$ för alla $A \in \Sigma$. Man kan om man vill se det hela omvänt och se måttet γ som "givet" och se f som definierad för att uppfylla denna ekvation för varje $A \in \Sigma$. Man kallar f för *Radon-Nikodym-derivatan* av γ m.a.p. μ och skriver

$$f = \frac{d\gamma}{d\mu}.$$

Det är naturligt att fråga sig om det för varje givet par av mått, γ och μ finns en sådan Radon-Nikodym-derivata av γ m.a.p. μ. Svaret är ja, under förutsättning att γ och μ är σ-ändliga och att γ är absolutkontinuerligt m.a.p. μ:

DEFINITION. Låt γ och μ vara två mått på samma mätbara rum (S, Σ). Man säger att γ är *absolutkontinuerligt* m.a.p. μ och skriver $\gamma << \mu$ om $A \in \Sigma, \mu(A) = 0 \Rightarrow \gamma(A) = 0$.

SATS 5.9 (RADON-NIKODYMS SATS) *Antag att γ och μ är två σ-ändliga mått på (S, Σ) sådana att $\gamma << \mu$. Då finns det en funktion $f \in (m\Sigma)^+$ sådan att*

$$\gamma(A) = \int_A f d\mu$$

för alla $A \in \Sigma$. Om g är en annan sådan funktion gäller att $f = g$ n.ö. m.a.p. μ.

Vi delar in beviset i två olika bitar. Unikhetsdelen av Radon-Nikodyms sats följer genom att tillämpa följande resultat (en direkt tillämpning om γ är ändligt och genom att dela upp S i γ-ändliga bitar om γ är σ-ändligt):

Sats 5.10 *Antag att $f, g \in L^1(S, \Sigma, \mu)$ är sådana att*

$$\int_A f d\mu \leq \int_A g d\mu$$

för alla $A \in \Sigma$. Då gäller att $f \leq g$ n.ö.

Bevis. Antag att $\mu\{f > g\} > 0$. Eftersom $\{f > g\} = \cup_n \{f > g + 1/n\}$ har vi för något heltal n att $\mu\{f > g + 1/n\} > 0$. Kalla den sistnämnda mängden för A. Då har vi att

$$\int_A f d\mu \geq \int_A (g + 1/n) d\mu = \int_A g d\mu + \frac{1}{n}\mu(A) > \int_A g d\mu,$$

en motsägelse. □

Observera att en direkt följd av Sats 5.10 är att om f är ickenegativ och $\int_S f d\mu = 0$ gäller att $f = 0$ n.ö.

Bevis av existensdelen. Observera först att det räcker att bevisa satsen då γ och μ är ändliga mått ty i det σ-ändliga fallet kan vi sedan dela upp S i ett uppräkneligt antal bitar på vilka γ och μ är ändliga och sedan summera respektive Radon-Nikodym-derivator f.

Antag nu att γ och μ är ändliga. Definiera funktionsklassen $\mathcal{F}$ som

$$\mathcal{F} = \{f \in (m\Sigma)^+ : \int_A f d\mu \leq \gamma(A) \text{ för alla } A \in \Sigma\}.$$

Klassen $\mathcal{F}$ är sluten under maxima ty om $f, g \in \mathcal{F}$ kan vi skriva $B = \{s \in S : f(s) \leq g(s)\}$ och det gäller för alla A att $\int_A \max(f, g) d\mu = \int_{A \cap B} g d\mu + \int_{A \cap B^c} f d\mu \leq \mu(A \cap B) + \mu(A \cap B^c) = \mu(A)$. Låt nu $m = \sup\{\int_S f d\mu : f \in \mathcal{F}\}$. Välj en följd $\{f_n\}$ av funktioner i $\mathcal{F}$ sådan att $\int_S f_n d\mu \to m$. Genom att sätta $g_n = \max(f_1, \ldots, f_n)$ får vi en växande följd av funktioner i $\mathcal{F}$ och vi kan sätta $f = \lim_{n\to\infty} g_n$. Enligt monoton-konvergens-satsen gäller att $\int_S f d\mu = m$.

Vi ska nu visa att f är en Radon-Nikodym-derivata. Sätt $\nu(A) = \int_A f d\mu$, $A \in \Sigma$. Då är ν ett mått och det följer att $\gamma - \nu$ är ett mått. Om f inte är en Radon-Nikodym-derivata gäller att det finns en mängd $A \in \Sigma$ med $\mu(A) > 0$ sådan att $\gamma(A) - \nu(A) > 0$. Med andra ord finns $\epsilon > 0$ sådant att $\gamma(A) - \nu(A) > 2\epsilon\nu(A)$. Definiera nu mängdfunktionen $\rho(A) = \gamma(A) - (1 + \epsilon)\nu(A)$, $A \in \Sigma$. Observera att ρ inte behöver vara ett mått eftersom $\rho(A)$ kan anta negativa värden, men eftersom ρ är differensen mellan två ändliga mått gäller att ρ är uppräkneligt additiv. Vi påstår nu att för varje A med $\rho(A) > \epsilon$ (vi observerade just att minst en sådan mängd existerar) gäller det att för varje $\delta > 0$ finns en mängd $B \subseteq A$ sådan att $\rho(B) \geq \rho(A)$ och $E \subseteq B \Rightarrow \rho(E) > -\delta$. Om så inte är fallet finns det nämligen en mängd $E_1 \subseteq A$ sådan att $\rho(E_1) \leq -\delta$. Eftersom ρ är additiv gäller då att $\rho(A \backslash E_1) > \epsilon$ och således finns en mängd $E_2 \subset A \backslash E_1$ sådan att $\rho(E_2) \leq -\delta$. Genom att förtsätta induktivt hittar vi disjunkta mängder $E_1, E_2, \dots$ sådana att $\rho(E_n) < -\delta$ för alla n. Men då gäller ju att $\rho(\cup_{n=1}^{\infty} E_n) = -\infty$, en omöjlighet eftersom ν är ett ändligt mått.

Fixera nu en mängd A med $\rho(A) > \epsilon$ och välj för $n = 1, 2, \dots$ mängden B_n så att $E \subseteq B_n \Rightarrow \rho(E) > -1/n$ och sätt $B = \cap_n B_n$. Då är B sådan att $E \subseteq B \Rightarrow \rho(B) \geq 0$ och dessutom gäller att $\rho(B) \geq \epsilon$, ty $\rho(B_n) \to \rho(B)$ enligt kontinuiteten hos γ och ν.

Vi har alltså visat att om f inte är en Radon-Nikodym-derivata finns det en mängd B med $\gamma(B) \geq (1 + 2\epsilon)\nu(B)$ sådan att för alla $E \subseteq B$ gäller att $\gamma(E) \geq (1 + \epsilon)\nu(E) = (1 + \epsilon) \int_E f d\mu$. Men då gäller ju att funktionen $h = f + \epsilon I_B \in \mathcal{F}$. Eftersom $\gamma << \mu$ måste det gälla att $\mu(B) > 0$ och vi får att $\int_S h d\mu > m$, en motsägelse mot definitionen av m. $\square$

6 Väntevärden

Vi byter nu ut vårt måttrum (S, Σ, μ) mot sannolikhetsrummet $(\Omega, \mathcal{F}, P)$. Kom ihåg att en stokastisk variabel per definition är en $\mathcal{F}$-mätbar $\overline{\mathbf{R}}$-värd funktion.

Definition. *Väntevärdet* av en stokastisk variabel X ges av

$$\mathbf{E}[X] = \int_{\Omega} X(\omega) P(d\omega).$$

Definitionen fungerar för varje ickenegativ stokastisk variabel, men generellt kräver vi naturligtvis att $X \in L^1(\Omega, \mathcal{F}, P)$.

Vi skriver om några av de viktiga resultaten från integrationsteorin i termer av stokastiska variabler:

- Monoton-konvergens-satsen: Om $\{X_n\}$ är en följd av ickenegativa stokastiska variabler sådan att $X_n \uparrow X$ n.s. gäller att $\mathbf{E}[X_n] \uparrow \mathbf{E}[X]$.
- Fatous lemma: Om $\{X_n\}$ är en följd av ickenegativa stokastiska variabler gäller att $\mathbf{E}[\liminf_n X_n] \leq \liminf_n \mathbf{E}[X_n]$.
- Dominerad-konvergens-satsen: Om $\{X_n\}$ är en följd av stokastiska variabler sådan att $X_n \to X$ n.s. och det existerar en ickenegativ stokastisk variabel Y med $\mathbf{E}[Y] < \infty$ sådan att $|X_n| \leq Y$ n.s. för alla n gäller att $\mathbf{E}[X_n] \to \mathbf{E}[X]$.

Notera att för att domineringsvillkoret i dominerad-konvergens-satsen ska vara uppfyllt räcker det att det existerar en ändlig konstant K sådan att $|X_n| \leq K$ n.s. för alla n. (Dominering av en konstant på detta sätt duger bra eftersom P är ett ändligt mått, men duger inte generellt. Exempelvis är ju $\int_{\mathbf{R}} K dLeb = K \cdot Leb(\mathbf{R}) = \infty$.)

Lite notation: För en händelse $A \in \mathcal{F}$ skriver vi $\mathbf{E}[X; A]$ för $\mathbf{E}[XI_A] = \int_A XdP$.

Sats 6.1 (Markovs olikhet) *Låt Z vara en stokastisk variabel och låt $g : \overline{\mathbf{R}} \to [0, \infty]$ vara en ickeavtagande mätbar funktion. Då gäller att*

$$\mathbf{E}[g(Z)] \geq g(c)P(Z \geq c)$$

för alla $c \in \overline{\mathbf{R}}$.

Bevis.

$$\mathbf{E}[g(Z)] = \int_\Omega (g \circ Z)dP \geq \int_{\{Z \geq c\}} (g \circ Z)dP$$

$$\geq \int_{\{Z \geq c\}} g(c)dP = g(c)P\{Z \geq c\}.$$

□

Det vanligaste specialfallet av Markovs olikhet får man om man sätter $Z = |X|$ och $g(x) = x^+$ (där $x^+ = \max(x, 0)$) och får att $\mathbf{E}|X| \geq cP(|X| \geq c)$. Om man sätter $g(x) = e^{\theta x}$ får man

$$P(Y \geq c) \leq e^{-\theta c}\mathbf{E}[e^{\theta Y}] = e^{-\theta c}M_Y(\theta)$$

där M_Y är den momentgenererande funktionen för Y. Detta kan ge en ganska skarp olikhet om θ väljs optimalt.

En direkt följd av Markovs olikhet är också den välkända *Chebyshevs* olikhet: $\epsilon^2 P(|X - \mathbf{E}[X]| \geq \epsilon) \leq E[(X - \mathbf{E}[X])^2]$ som fås genom att sätta $Z = |X - \mathbf{E}[X]|$ och $g(x) = (x^+)^2$. Om högersidan är ändlig definierar vi den i kapitel 6 som $Var(X)$.

Följande sats ger några små användbara resultat om väntevärden.

Sats 6.2 *(a) $E|X| < \infty \Rightarrow |X| < \infty$ n.s.*

(b) Om $Z_1, Z_2, \ldots$ är ickenegativa gäller att $\mathbf{E}[\sum_{n=1}^\infty Z_n] = \sum_{n=1}^\infty \mathbf{E}[Z_n]$.

(c) *Om* $Z_1, Z_2, \dots$ *är ickenegativa och* $\sum_{n=1}^{\infty} \mathbf{E}[Z_n] < \infty$ *gäller att* $\sum_{n=1}^{\infty} Z_n < \infty$ *n.s.*

Bevis. Del (a) är självklar, del (b) är en tillämpning av monoton-konvergens-satsen och del (c) är en direkt kombination av (a) och (b). $\square$

Exempel. Antag att $F_1, F_2, \dots$ är händelser sådana att $\sum_{n=1}^{\infty} P(F_n) < \infty$. Sätt $Z_n = I_{F_n}$. Då följer det av del (c) i satsen ovan att $\sum_{n=1}^{\infty} I_{F_n} < \infty$ n.s., dvs Borel-Cantellis första lemma.

6.1 Jensens olikhet

DEFINITION. Låt G vara ett öppet intervall i $\mathbf{R}$ och låt $c : G \to \mathbf{R}$ vara en funktion. Om c är sådan att

$$p, q \in [0,1], p + q = 1 \Rightarrow c(px + qy) \leq pc(x) + qc(y)$$

för alla $x, y \in G$ sägs c vara en *konvex* funktion.

I ord kan man säga att c är konvex ifall ingen del av en rät linje dragen mellan två punkter på c's graf hamnar under grafen. Ett mer naivt sätt att beskriva det hela på är att c är konvex om ingen del av grafen böjer av nedåt.

Observera att intervallet G i definitionen kan vara $\mathbf{R}$ självt.

SATS 6.3 (JENSENS OLIKHET) *Antag att* $c : G \to \mathbf{R}$ *är konvex och att* X *är en integrerbar stokastisk variabel sådan att* $X \in G$ *n.s. och* $\mathbf{E}|c(X)| < \infty$. *Då gäller att*

$$\mathbf{E}[c(X)] \geq c(\mathbf{E}[X]).$$

Bevis. Sätt $\Delta(u,v) = \frac{c(v)-c(u)}{v-u}$, $u < v, u, v \in G$. Vi ska börja med att visa att $u_1 \leq u_2, v_1 \leq v_2 \Rightarrow \Delta(u_1, v_1) \leq \Delta(u_2, v_2)$. För att göra detta räcker det att visa att $u < v < w \Rightarrow \Delta(u,v) \leq \Delta(v,w)$. Vi har:

$$\Delta(v,w) - \Delta(u,v) = \frac{1}{(v-u)(w-v)}[(v-u)(c(w)-c(v)) - (w-v)(c(v)-c(u))]$$

$$= \frac{1}{(v-u)(w-v)}[-(w-u)c(v) + (w-v)c(u) + (v-u)c(w)]$$

$$= \frac{w-u}{(v-u)(w-v)}[\frac{w-v}{w-u}c(u) + \frac{v-u}{w-u}c(w) - c(v)].$$

Men

$$\frac{w-v}{w-u}u + \frac{v-u}{w-u}w = v$$

så enligt konvexiteten hos c gäller att $\Delta(v,w) - \Delta(u,v) \geq 0$ som önskat.

Om vi antar att $\Delta(u,v) = \infty$ för något par (u,v) följer det av det vi nyss visade att $c(v) = \infty$, en motsägelse. På motsvarande sätt följer att $\Delta(u,v) > -\infty$. (Av detta följer att c är kontinuerlig.)

Fixera nu v och sätt $(D_-c)(v) = \lim_{u\uparrow v}\Delta(u,v)$ och $(D_+c)(v) = \lim_{w\downarrow v}\Delta(v,w)$. Enligt vad vi visat ovan har vi att dessa gränser existerar och är reellvärda och att $(D_-c)(v) \leq (D_+c)(v)$. Fixera nu att tal m sådant att $(D_-c)(v) \leq m \leq (D_+c)(v)$. För ett godtyckligt $x \in G$ gäller nu att om $x \geq v$ så är $c(x) - c(v) = \Delta(v,x)(x-v) \geq m(x-v)$ och om $x \leq v$ så är $c(x) - c(v) = -\Delta(x,v)(x-v) \geq m(x-v)$, dvs $c(x) \geq c(v) + m(x-v)$ för alla $x, v \in G$.

Sätt nu $x = X$ och $v = \mathbf{E}[X]$ och vi får:

$$c(X) \geq c(\mathbf{E}[X]) + m(X - \mathbf{E}[X]).$$

Tag nu väntevärde av bägge sidor och saken är klar. $\square$

Låt oss göra följande observation: Om man i beviset ovan sätter $m_q = (D_-c)(q)$ för varje $q \in G$ följer det av kontinuiteten hos c och det faktum att $c(x) \geq c(q) + m_q(x-q)$ att

$$c(x) = \sup_{q\in G}(m_q(x-q) + c(q)).$$

Tack vare c's kontinuitet spelar det heller ingen roll om vi låter q löpa genom de rationella talen i G, dvs

$$c(x) = \sup_{q\in\mathbf{Q}\cap G}(m_q(x-q) + c(q)).$$

Genom att skriva $a_q = m_q$ och $b_q = c(q) - m_q q$ har vi att

$$c(x) = \sup_{q\in\mathbf{Q}\cap G}(a_q x + b_q)$$

dvs vi har bevisat att det existerar två följder $\{a_n\}$ och $\{b_n\}$ sådana att för alla $x \in G$ gäller att

$$c(x) = \sup_n(a_n x + b_n).$$

Detta resultat är känt under namnet *stödlinjessatsen.* Vi kommer att ha användning för stödlinjessatsen längre fram.

6.2 L^p-normer

Vi har redan tidigare definierat $L^1(\Omega, \mathcal{F}, P)$ som klassen av alla $\mathcal{F}$-mätbara funktioner X sådana att $\int_\Omega |X| dP = \mathbf{E}|X| < \infty$. Generellt definierar vi $L^p(\Omega, \mathcal{F}, P)$ som klassen av $\mathcal{F}$-mätbara funktioner X sådana att $\int_\Omega |X|^p dP = \mathbf{E}[|X|^p] < \infty$, $1 \leq p < \infty$. *L^p-normen* $||X||_p$ definieras som $\mathbf{E}[|X|^p]^{1/p}$.

SATS 6.4 (MONOTONICITET HOS L^p-NORMER) *Om $1 \leq p \leq r < \infty$ gäller att $||X||_p \leq ||X||_r$.*

Bevis. Vi ska använda Jensens olikhet med $c(x) = x^{r/p}$ och skulle gärna vilja tillämpa den på $|X|^p$. Nu finns det ju inga garantier för att $|X|^p$ uppfyller förutsättningarna i Jensens olikhet, men om vi sätter $X_n = (|X| \wedge n)^p$, $n = 1, 2, \ldots$ har vi att $X_n \uparrow |X|^p$ och $E[X_n] < \infty$ och $\mathbf{E}[X_n^{r/p}] < \infty$. Därför ger Jensens olikhet att

$$\mathbf{E}[X_n]^{r/p} \leq \mathbf{E}[X_n^{r/p}]$$

och genom att tillämpa monoton-konvergens-satsen får vi att $\mathbf{E}[|X|^p]^{r/p} \leq \mathbf{E}[|X|^r]$. Tag nu r:te roten ur bägge sidor för att avsluta beviset. □

Den metod vi använde i beviset nyss, att "kapa av toppen" på X, brukar kallas för *trunkering* och är ofta användbar.

SATS 6.5 *$L^p(\Omega, \mathcal{F}, P)$ är ett reellt vektorrum.*

Bevis. Det gäller att visa att L^p är slutet under multiplikation med konstanter och under addition, så antag att $X, Y \in L^p$ och att $a \in \mathbf{R}$. Att $aX \in L^p$ följer av att $\mathbf{E}[|aX|^p] = |a|^p \mathbf{E}[|X|^p] < \infty$ och eftersom $\mathbf{E}[|X+Y|^p] \leq E[(2(|X| \vee |Y|))^p] \leq 2^p \mathbf{E}[|X|^p + |Y|^p] < \infty$ gäller även att $X + Y \in L^p$. □

Vi ska nu ett tag koncentrera oss på specialfallet L^2 där man kan göra en rad intressanta saker.

SATS 6.6 (SCHWARZ OLIKHET) *Antag att $X, Y \in L^2$. Då gäller att $|\mathbf{E}[XY]| \leq \mathbf{E}|XY| \leq \mathbf{E}[X^2]^{1/2}\mathbf{E}[Y^2]^{1/2} = ||X||_2 ||Y||_2$.*

Eftersom Schwarz olikhet är ett specialfall av Hölders olikhet som vi visar i slutet av detta kapitel hoppar vi över beviset.

Från Schwarz olikhet följer *triangelolikheten*: $X, Y \in L^2 \Rightarrow ||X+Y||_2 \leq ||X||_2 + ||Y||_2$. (Detta är egentligen en del av *definitionen* av en norm, men i denna kurs har vi ju ännu inte visat ett L^p-normen verkligen är en norm.)

Bevis av triangelolikheten.

$$||X+Y||_2^2 = \mathbf{E}[|X+Y||X+Y|] \leq \mathbf{E}[|X||X+Y|] + \mathbf{E}[|Y||X+Y|]$$

$$\leq ||X||_2||X+Y||_2 + ||Y||_2||X+Y||_2 = ||X+Y||_2(||X||_2 + ||Y_2||)$$

där den första olikheten följer av den vanliga triangelolikheten för reella tal och den andra följer av Schwarz olikhet. Dividera nu bägge sidor med $||X+Y||_2$. $\square$

Antag att $X, Y \in L^2$ så att $\mathbf{E}[X^2]$ och $\mathbf{E}[Y^2]$ är ändliga. Vi kan då definiera varianser och kovarians:

Från monotoniciteten hos L^p-normer följer att $X, Y \in L^1$ så att $\mu_X = \mathbf{E}[X]$ och $\mu_Y = \mathbf{E}[Y]$ bägge existerar och är reellvärda. Eftersom L^2 är ett vektorrum har vi att de stokastiska variablerna $\bar{X} = X - \mu_X$ och $\bar{Y} = Y - \mu_Y$ bägge ligger i L^2. Alltså kan vi använda Schwarz olikhet till att inse att $\mathbf{E}[|\bar{X}\bar{Y}|] \leq ||\bar{X}||_2||\bar{Y}||_2 < \infty$, dvs att $\bar{X}\bar{Y} \in L^1$ så att $\mathbf{E}[\bar{X}\bar{Y}]$ existerar.

DEFINITION. $Cov(X,Y) = \mathbf{E}[\bar{X}\bar{Y}] = \mathbf{E}[(X-\mu_X)(Y-\mu_Y)] = \mathbf{E}[XY] - \mu_X\mu_Y$.

Notera att vi måste använda Schwarz olikhet för att få den sista likheten; den behövs ju för att tala om för oss att $XY \in L^1$ så att $\mathbf{E}[XY]$ existerar.

DEFINITION. $Var(X) = Cov(X,X) = \mathbf{E}[(X-\mu_X)^2] = \mathbf{E}[X^2] - \mu_X^2$.

6.3 Om geometrin hos $L^2(\Omega, \mathcal{F}, P)$

Man kan definiera inre produkt på L^2: Sätt $\langle U, V\rangle = \mathbf{E}[UV] = \int_\Omega U(\omega)V(\omega)P(d\omega)$, $U, V \in L^2$. Observera att $\langle\cdot,\cdot\rangle$ är linjär i sina argument och att $\langle U, U\rangle = ||U||_2$. Därav följer *parallellogramlagen*:

$$||U+V||_2^2 + ||U-V||_2^2 = \langle U+V, U+V\rangle + \langle U-V, U-V\rangle$$

$$= 2\langle U, U\rangle + 2\langle V, V\rangle = 2||U||_2^2 + 2||V||_2^2$$

och Pytagoras sats:

$$\langle U, V\rangle = 0 \Rightarrow ||U+V||_2^2 = ||U||_2^2 + ||V||_2^2.$$

Genom att tillämpa Pytagoras sats på $U = X - \mu_x$ och $V = Y - \mu_Y$ får man det välkända faktum att $Cov(X,Y) = 0 \Rightarrow Var(X+Y) = Var(X) + Var(Y)$.

Not. Ett *inre produktrum* är är ett linjärt rum där det finns en inre produkt definierad sådan att om man sätter $||x|| = \langle x, x\rangle$ så är $||\cdot||$ en norm. Ett av kraven för att $||\cdot||$ ska få kallas för en norm är att $||x|| = 0 \Leftrightarrow x = 0$. Härav ser vi att $L^2(\Omega, \mathcal{F}, P)$ inte är ett alldeles äkta inre produktrum efterom $||U||_2 = \mathbf{E}[U^2]^{1/2} = 0$ endast är ekvivalent med att $U = 0$ *nästan säkert*. Man kan komma undan detta problem genom att dela in L^2 i ekvivalensklasser enligt relationen $U \sim V$ om $U = V$ n.s. och sedan betrakta varje ekvivalensklass som ett enda element i ett nytt rum. Detta förorsakar inga problem i denna kurs men kan göra det i mer avancerade sammanhang. I fortsättningen betraktar vi alltså L^2 som ett äkta inre produktrum utan att bekymra oss om den saken.

SATS 6.7 *För alla $p \geq 1$ gäller att $L^p(\Omega, \mathcal{F}, P)$ är ett fullständigt rum.*

Bevis. Kom ihåg att fullständighet betyder att alla Cauchyföljder konvergerar. Vi ska alltså visa att om $\{X_n\}$ är en följd av stokastiska variabler i L^p sådan att $||X_r - X_s||_p \to 0$ då $r, s \to \infty$ så finns en stokastisk variabel $X \in L^p$ sådan att $||X_n - X||_p \to 0$.

Antag att $\{X_n\}$ är en sådan följd och låt $k_1, k_2, \ldots$ vara en följd av index sådan att $r, s \geq k_n \Rightarrow ||X_r - X_s||_p \leq 2^{-n}$. Eftersom

$$\mathbf{E}[|X_{k_{n+1}} - X_{k_n}|] = ||X_{k_{n+1}} - X_{k_n}||_1 \leq ||X_{k_{n+1}} - X_{k_n}||_p \leq 2^{-n}$$

enligt monotoniciteten hos L^p-normer, gäller att

$$\mathbf{E}[\sum_{n=1}^{\infty} |X_{k_{n+1}} - X_{k_n}|] = \sum_{n=1}^{\infty} \mathbf{E}[|X_{k_{n+1}} - X_{k_n}|] \leq 1 < \infty.$$

Enligt Sats 6.2(c) gäller alltså att $\sum_{n=1}^{\infty} |X_{k_{n+1}} - X_{k_n}|$ är n.s. konvergent. Men det betyder ju att den teleskoperande summan $\sum_{n=1}^{\infty}(X_{k_{n+1}} - X_{k_n})$ också är n.s. konvergent, dvs X_{k_n} konvergerar n.s. då $n \to \infty$. Genom att sätta $X(\omega) = \limsup_n X_{k_n}(\omega)$, $\omega \in \Omega$, har vi alltså konstruerat en stokastisk variabel X sådan att $X_{k_n} \to X$ n.s.

Det som nu återstår att visa är dels att $X \in L^p$, dels att $||X_n - X||_p \to 0$. Om vi väljer $r \geq k_n$ och $m \geq n$ gäller att

$$\mathbf{E}[|X_r - X_{k_m}|^p] = ||X_r - X_{k_m}||_p^p \leq 2^{-np}.$$

Låt nu $m \to \infty$. Då gäller att $|X_r - X_{k_m}|^p \to |X_r - X|^p$ n.s. så enligt Fatous lemma har vi att

$$\mathbf{E}[|X_r - X|^p] \leq \liminf_m \mathbf{E}[|X_r - X_{k_m}|^p] \leq 2^{-np}.$$

Detta betyder för det första att $X_r - X \in L^p$ så att $X \in L^p$ enligt vektorrumsegenskapen hos L^p och för det andra att $r \geq k_n \to \mathbf{E}[|X_r - X|^p] \leq 2^{-np}$, dvs $\mathbf{E}[|X_r - X|^p] \to 0$ då $r \to \infty$ och saken är klar. $\square$

Den som är intresserad av funktionalanalys noterar att Sats 6.7 säger att $L^p(\Omega, \mathcal{F}, P)$ är ett Banachrum om man inför ekvivalensklasser så som indikerat i noten ovan.

Vårt främsta intresse för fullständigheten hos L^p-rum i denna kurs ligger i det att den kan utnyttjas för att påvisa existensen av *ortogonala projektioner* i L^2:

SATS 6.8 *Låt $\mathcal{K}$ vara ett fullständigt vektorunderrum till $L^2(\Omega, \mathcal{F}, P)$ och tag $X \in L^2(\Omega, \mathcal{F}, P)$. Då finns det en stokastisk variabel $Y \in \mathcal{K}$ sådan att*

(i) $||X - Y|| = \inf\{||X - W|| : W \in \mathcal{K}\}$,

(ii) $\langle X - Y, Z \rangle = 0$ *för alla* $Z \in \mathcal{K}$.

Egenskaperna (i) och (ii) är ekvivalenta och om Y' är en annan stokastisk variabel i $\mathcal{K}$ som uppfyller (i) och (ii) gäller att $Y = Y'$ n.s.

Observera att för att förenkla notationen en aning skriver vi här $||\cdot||$ för $||\cdot||_2$ och låter tvåan vara underförstådd.

Bevis. Vi börjar bakifrån med att bevisa att om Y och Y' bägge satisfierar (ii) så är $Y = Y'$ n.s. Om Y och Y' satisfierar (ii) har vi för alla $Z \in \mathcal{K}$ att $\langle X - Y', Z \rangle - \langle X - Y, Z \rangle = \langle Y - Y', Z \rangle = 0$. Eftersom $\mathcal{K}$ är ett vektorrum gäller att $Y - Y' \in \mathcal{K}$ så genom att sätta $Z = Y - Y'$ får vi att $\langle Y - Y', Y - Y' \rangle = ||Y - Y'||^2 = 0$ och således att $Y = Y'$ n.s.

För att visa att (i) implicerar (ii), tag en godtycklig stokastisk variabel $Z \in \mathcal{K}$ och $t \in \mathbf{R}$ och antag att Y satisfierar (i). Eftersom $Y + tZ \in \mathcal{K}$ har vi då att

$$||X - Y - tZ||^2 = \langle X - Y - tZ, X - Y - tZ \rangle$$

$$= ||X - Y||^2 + t^2||Z||^2 - 2t\langle X - Y, Z \rangle \geq ||X - Y||^2$$

dvs att

$$t^2||Z||^2 - 2t\langle X - Y, Z \rangle \geq 0$$

för alla t. Men för att detta skall vara sant för alla t nära 0 måste det gälla att $\langle X-Y, Z\rangle = 0$, vilket bevisar (ii).

Å andra sidan gäller att om Y satisfierar (ii) har vi för godtyckligt $W \in \mathcal{K}$ att

$$||X - W||^2 = ||(X - Y) - (W - Y)||^2$$

$$= ||X - Y||^2 + ||W - Y||^2 + 2\langle X - Y, W - Y\rangle$$

$$= ||X - Y||^2 + ||W - Y||^2 \geq ||X - Y||^2$$

där den andra likheten följar av det faktum att $W - Y \in \mathcal{K}$.

Nu över till huvudnumret, dvs existensen av $Y \in \mathcal{K}$ som uppfyller (i). Sätt $\Delta = \inf\{||X - W|| : W \in \mathcal{K}\}$. Välj en följd $\{Y_n\}$ av stokastiska variabler i $\mathcal{K}$ sådan att $||X - Y_n|| \to \Delta$. Kom nu ihåg parallellogramlagen som sade att för $U, V \in L^2$ gäller

$$||U + V||^2 + ||U - V||^2 = 2||U||^2 + 2||V||^2.$$

Genom att sätta $U = X - Y_r/2 - Y_s/2$ och $V = Y_r/2 - Y_s/2$ får man

$$||X - Y_s||^2 + ||X - Y_r||^2 = 2||X - (Y_r + Y_s)/2||^2 + 2||(Y_r - Y_s)/2||^2.$$

Eftersom $(Y_r + Y_s)/2 \in \mathcal{K}$ har vi att $||X - (Y_r + Y_s)/2||^2 \geq \Delta^2$. Detta tillsammans med det faktum att $||X - Y_r||^2$ och $||X - Y_s||^2$ bägge konvergerar mot Δ^2 då $r, s \to \infty$ leder till slutsatsen att $||Y_r - Y_s||^2 \to 0$ då $r, s \to \infty$, dvs att följden $\{Y_n\}$ är en Cauchyföljd. Enligt antagande är $\mathcal{K}$ ett fullständigt underrum till L^2 varför det existerar en stokastisk variabel $Y \in \mathcal{K}$ sådan att $||Y_n - Y|| \to 0$. Enligt triangelolikheten gäller nu att

$$||X - Y|| \leq ||X - Y_n|| + ||Y_n - Y|| \to \Delta$$

och således att $||X - Y|| = \Delta$ som önskat. $\square$

6.4 Den omedvetne statistikerns lag

SATS 6.9 (DEN OMEDVETNE STATISTIKERNS LAG) *Låt X vara en stokastisk variabel och låt P_X vara fördelningen för X, dvs $P_X(B) = P(X \in B)$, $B \in \mathcal{B}$. Om $h : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ är en Borelfunktion sådan att $h \circ X \in L^1(\Omega, \mathcal{F}, P)$ gäller att*

$$\mathbf{E}[h(X)] = \int_{\overline{\mathbf{R}}} h(x) P_X(dx).$$

Bevis. Detta är en bra övning på att använda standardmaskineriet:

Antag först att $h = I_B$ för någon mängd $B \in \mathcal{B}$. Då är $\mathbf{E}[h(X)] = \mathbf{E}[I_B(X)] = \mathbf{E}[I_{\{X \in B\}}] = P(X \in B) = P_X(B) = \int_{\overline{\mathbf{R}}} I_B(x) P_X(dx)$.

Nu följer det direkt av linjaritet hos integraler att satsen gäller för alla enkla funktioner h.

Antag nu att $h \in (m\mathcal{B})^+$. Då kan vi finna en följd $\{h_n\}$ av enkla funktioner sådan att $h_n \uparrow h$. Då gäller att $h_n(X) \uparrow h(X)$ och det följer av monoton-konvergens-satsen att $\mathbf{E}[h(X)] = \lim_{n\to\infty} \mathbf{E}[h_n(X)] = \lim_{n\to\infty} \int_{\overline{\mathbf{R}}} h_n(x) P_X(dx) = \int_{\overline{\mathbf{R}}} h(x) P_X(dx)$.

Till sist, då $h(X) \in L^1$, delar vi upp h i h^+ och h^- och använder linjaritet hos integraler för att avsluta beviset. $\square$

Om vi skriver F för X's fördelningsfunktion skriver man ofta integralen $\int_{\overline{\mathbf{R}}} h(x) P_X(dx)$ som $\int_{\overline{\mathbf{R}}} h(x) dF(x)$.

Exempel. Antag att X är en stokastisk variabel sådan att $P_X << Leb$. Enligt Radon-Nikodyms sats finns det då en funktion $f \in (m\mathcal{B})^+$ sådan att $P_X(B) = \int_B f dLeb = \int_B f(x)dx$, dvs $P(X \in B) = \int_B f(x)dx$ för alla $B \in \mathcal{B}$. Detta betyder att f är en *täthetsfunktion* för X. Enligt Sats 5.10 är f n.ö. unik. Den omedvetne statistikerns lag talar om för oss att om $\int_{-\infty}^{\infty} |h(x)| f(x)dx < \infty$ så gäller att $\mathbf{E}[h(X)] = \int_{-\infty}^{\infty} h(x)f(x)dx$.

6.5 Hölder och Minkowski

Vid presentationen av Schwarz olikhet lovades att det längre fram skulle komma ett bevis på den mer generella Hölders olikhet och det är just detta vi ska gripa oss an i detta avsnitt. Vi håller oss här med ett generellt måttrum (S, Σ, μ). Vi säger att en funktion $f : S \to \overline{\mathbf{R}}$ ligger i $L^p(S, \Sigma, \mu)$ om $\int_S |f|^p d\mu < \infty$. Vi definierar precis som förut L^p-normen $||f||_p$ som $(\int_S |f|^p d\mu)^{1/p}$.

Sats 6.10 (Hölders olikhet) *Antag att $p, q \in (1, \infty)$ är sådana att $1/p + 1/q = 1$. Om $f \in L^p(S, \Sigma, \mu)$ och $h \in L^q(S, \Sigma, \mu)$ gäller att $fh \in L^1(S, \Sigma, \mu)$ och att*

$$|\int_S fh d\mu| \leq \int_S |fh| d\mu \leq ||f||_p ||h||_q.$$

Bevis. Den första olikheten är självklar så vi koncentrerar oss på den andra. Vi antar utan att förlora generalitet att $f, h \geq 0$ och att $\int_S f^p d\mu > 0$. Definiera ett sannolikhetsmått P genom att sätta

$$P(A) = \frac{\int_A f^p d\mu}{\int_S f^p d\mu}$$

för alla $A \in \Sigma$. Definiera också en stokastisk variabel U genom att låta

$$U(s) = \frac{h(s)}{f(s)^{p-1}} I_{\{f>0\}}(s), s \in S.$$

Nu är det ju så att antingen är $E[U^q] = \infty$ och då gäller trivialt att $\mathbf{E}[U^q] \geq \mathbf{E}[U]^q$ eller så har vi $\mathbf{E}[U^q] < \infty$ och i så fall är $\mathbf{E}[U] = \mathbf{E}[U; U < 1] + \mathbf{E}[U; U \geq 1] \leq 1 + \mathbf{E}[U^q] < \infty$ och vi kan använda Jensens olikhet till att få att $\mathbf{E}[U^q] \geq \mathbf{E}[U]^q$.

Enligt definitionen av U och P gäller att

$$\mathbf{E}[U^q] = \frac{\int_S (h(s)^q / f(s)^{q(p-1)}) f(s)^p \mu(ds)}{\int_S f(s)^p \mu(ds)} = \frac{\int_S h^q d\mu}{\int_S f^p d\mu}$$

ty $pq - p - q = 0$, medan

$$\mathbf{E}[U]^q = (\frac{\int_S (h(s)/f(s)^{p-1}) f(s)^p \mu(ds)}{\int_S f(s)^p \mu(ds)})^q = \frac{(\int_S fh d\mu)^q}{(\int_S f^p d\mu)^q}.$$

Genom att sätta in dessa uttryck i olikheten ovan får vi att

$$(\int_S fh d\mu)^q \leq \int_S h^q d\mu (\int_S f^p d\mu)^{q-1}$$

och genom att ta q:te roten ur bägge sidor får vi att

$$\int_S fh d\mu \leq ||h||_q (\int_S f^p d\mu)^{1-1/q} = ||h||_q ||f||_p$$

ty $1 - 1/q = 1/p$. $\square$

För att få Schwarz olikhet sätter vi $p = q = 2$ i Hölders olikhet. Nästa resultat vi ska visa är Minkowskis olikhet som man också kan referera till som triangelolikheten för L^p-norm:

Sats 6.11 (Minkowskis olikhet) *Tag $p > 1$ och antag att $f, g \in L^p(S, \Sigma, \mu)$. Det gäller att $||f + g||_p \leq ||f||_p + ||g||_p$.*

Bevis. Vi sätter $q = (1 - 1/p)^{-1}$ och härmar beviset av triangelolikheten för L^2-norm:

$$||f + g||_p^p = \int |f + g|^p d\mu \leq \int |f||f + g|^{p-1} d\mu + \int |g||f + g|^{p-1} d\mu$$

$$\leq (||f||_p + ||g||_p)|||f+g|^{p-1}||_q$$

där den sista olikheten följer av Hölders olikhet. Men

$$|||f+g|^{p-1}||_q = (\int |f+g|^{q(p-1)} d\mu)^{1/q} = ||f+g||_p^{p/q} = ||f+g||_p^{p-1}$$

ty $q(p-1) = p$, så genom att dividera bägge sidor av olikheten med $||f+g||_p^{p-1}$ får vi det önskade resultatet. □

7 Väntevärden av oberoende stokastiska variabler

Sats 7.1 *Om X och Y är oberoende stokastiska variabler i $L^1(\Omega, F, P)$ gäller att $XY \in L^1$ och att $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.*

Observera att som en direkt följd av detta resultat gäller att om X och Y är oberoende har vi att $Cov(X, Y) = 0$.

Bevis. Antag först att X och Y är enkla funktioner, dvs $X = \sum_{i=1}^n x_i I_{A_i}$ och $Y = \sum_{j=1}^m y_j I_{B_j}$. Vi antar utan att förlora generalitet att alla x_i:n är olika och att det samma gäller för y_j:na. Detta betyder att $A_i \in \sigma(X)$ för alla i och $B_j \in \sigma(Y)$ för alla j och således enligt oberoendeantagandet att $P(A_i \cap B_j) = P(A_i)P(B_j)$ för alla i och j. Eftersom $XY = \sum_{i=1}^n \sum_{j=1}^m x_i y_j I_{A_i \cap B_j}$ får vi då att

$$\mathbf{E}[XY] = \sum_{i=1}^n \sum_{j=1}^m x_i y_j P(A_i \cap B_j) = \sum_{i=1}^n x_i P(A_i) \sum_{j=1}^m y_j P(B_j) = \mathbf{E}[X]\mathbf{E}[Y].$$

Antag nu att X och Y är ickenegativa. Då finns följder $\{\phi_n\}$ och $\{\psi_n\}$ av enkla funktioner sådana att ϕ_n:na är $\sigma(X)$-mätbara, $\phi_n \uparrow X$, ψ_n:na är $\sigma(Y)$-mätbara och $\psi_n \uparrow Y$. (Att man kan få ϕ_n:na och ψ_n:na att vara $\sigma(X)$ respektive $\sigma(Y)$-mätbara följer implicit av konstruktionen man gör i beviset av Sats 5.3.) Eftersom varje ϕ_n är oberoende av varje ψ_n följer det av det ovan visade och monoton-konvergens-satsen att

$$\mathbf{E}[XY] = \lim_{n\to\infty} \mathbf{E}[\phi_n \psi_n] = \lim_{n\to\infty} \mathbf{E}[\phi_n]\mathbf{E}[\psi_n] = \mathbf{E}[X]\mathbf{E}[Y].$$

Slutligen gäller för generella $X, Y \in L^1$ att

$$E[XY] = \mathbf{E}[(X^+ - X^-)(Y^+ - Y^-)] = \mathbf{E}[X]\mathbf{E}[Y]$$

då ju $X^{+/-}$ är oberoende av $Y^{+/-}$. □

Den mest generella formen av stora talens för oberoende och likafördelade stokastiska variabler har ett långt och därmed ganska besvärligt bevis. Nedan följer en inte fullt lika generell form som dock har den fördelen att beviset är betydligt kortare utan att för den skull vara mindre elementärt. Det enda extra villkoret är att de stokastiska variablerna har begränsat fjärdemoment, vilket i någon mening är en ganska liten inskränkning eftersom de flesta av de "vanliga" fördelningarna har alla moment. Dessutom kan man frångå antagandet om likafördelning.

Sats 7.2 (Stora talens lag med begränsat fjärdemoment) *Antag att $X_1, X_2, \ldots$ är oberoende stokastiska variabler med $\mathbf{E}[X_k] = 0$ och sådana att det existerar en konstant $K < \infty$ sådan att $\mathbf{E}[X_k^4] \leq K$ för alla k. Med $S_n = \sum_{k=1}^n X_k$ gäller att $S_n/n \to 0$ n.s.*

Bevis. Vi utvecklar $S_n^4 = (\sum_{k=1}^n X_k)^4$:

$$S_n^4 = \sum_k X_k^4 + 4\sum\sum_{i\neq j} X_i^3 X_j + 12\sum\sum\sum_{i\neq j, i\neq k, j<k} X_i^2 X_j X_k$$

$$+24\sum\sum\sum\sum_{i<j<k<l} X_i X_j X_k X_l + 6\sum\sum_{i<j} X_i^2 X_j^2.$$

Enligt oberoendet och Sats 7.1 gäller att $\mathbf{E}[X_i^3 X_j] = 0$ och detsamma gäller för $\mathbf{E}[X_i^2 X_j X_k]$ och $\mathbf{E}[X_i X_j X_k X_l]$. (Att Sats 7.1 är tillämpbar följer av monotoniciteten hos L^p-normer och begränsningen i fjärdemomentet.) Därför får vi att

$$\mathbf{E}[S_n^4] = \mathbf{E}[\sum_k X_k^4 + 6\sum\sum_{i<j} X_i^2 X_j^2].$$

Eftersom $||X_i||_2 \leq ||X_i||_4$ gäller att $\mathbf{E}[X_i^2] \leq \mathbf{E}[X_i^4]^{1/2} \leq K^{1/2}$ och därför att $\mathbf{E}[X_i^2 X_j^2] = \mathbf{E}[X_i^2]\mathbf{E}[X_j^2] \leq K$. Alltså gäller att $\mathbf{E}[S_n^4] \leq nK + 3n(n-1)K \leq 3n^2 K$ och därmed att $\mathbf{E}[(S_n/n)^4] \leq 3K/n^2$. Genom att summera över n får vi att $\mathbf{E}[\sum_{n=1}^\infty (S_n/n)^4] \leq 3K\sum_{n=1}^\infty n^{-2} < \infty$. Detta leder till slutsatsen att $\sum_{n=1}^\infty (S_n/n)^4 < \infty$ n.s. och speciellt att $(S_n/n)^4 \to 0$ n.s. dvs $S_n/n \to 0$ n.s. $\square$

Resultatet ovan generaliserar sig till fallet då $\mathbf{E}[X_k] = \mu$ för alla k och säger då att $S_n/n \to \mu$ n.s. Detta följer genom att tillämpa satsen på $Y_k = X_k - \mu$. Det gäller bara att kontrollera så att Y_k:na har begränsat fjärdemoment, vilket följer av Minkowskis olikhet då $\mathbf{E}[Y_k^4] = ||Y_k||_4^4 \leq (||\mu||_4 + ||X_k||_4)^4 = (|\mu| + ||X_k||_4)^4$.

8 Produktmått

8.1 Produkter av två måttrum

Antag att vi vill definiera begreppet "area" i planet $\mathbf{R}^2 = \mathbf{R}\times\mathbf{R}$. Vi vill alltså skapa ett lämpligt mått μ definierat på en lämplig σ-algebra av delmängder av $\mathbf{R}^2$. Ett naturligt kriterium på ett sådant mått är att om B_1 och B_2 är Borelmätbara delmängder av $\mathbf{R}$ så ska det gälla att $\mu(B_1 \times B_2) = Leb(B_1) \cdot Leb(B_2)$. (Kom ihåg produktbeteckningen för mängder: Om A och B är godtyckliga mängder så definieras $A \times B$ som $\{(x,y) : x \in A, y \in B\}$.) Implicit betyder detta också att $B_1 \times B_2$ måste tillhöra den σ-algebra på vilken areamåttet är definierat för alla $B_1, B_2 \in \mathcal{B}$. Vi låter σ-algebran av delmängder som vi definierar area för vara just σ-algebran genererad av alla $B_1 \times B_2$, $B_1, B_2 \in \mathcal{B}$ och vårt problem blir att visa att det existerar att unikt mått som uppfyller kriteriet ovan. Vi ska förstås göra detta i en mer generell situation. Vi antar att vi har måttrummen (S_1, Σ_1, μ_1) och (S_2, Σ_2, μ_2) och betraktar $S = S_1 \times S_2$.

Definition. *Produkt-σ-algebran* $\Sigma = \Sigma_1 \times \Sigma_2$ ges av

$$\Sigma = \sigma\{S_1 \times B_1, B_1 \times S_2 : B_1 \in \Sigma_1, B_2 \in \Sigma_2\}.$$

Alternativt kan man definiera Σ med hjälp av koordinatavbildningar: Definiera $\rho_i : S_1 \times S_2 \to S_i$, $i = 1, 2$ genom att sätta $\rho_i(s_1, s_2) = s_i$. Då blir $\Sigma = \sigma(\rho_1, \rho_2)$.

Genom att observera att $(S_1 \times B_2) \cap (B_1 \times S_2) = B_1 \times B_2$ följer det direkt att Σ är genererad av klassen $\mathcal{I} = \{B_1 \times B_2, B_1 \in \Sigma_1, B_2 \in \Sigma_2\}$, som är ett π-system.

Definiera nu en algebra Σ_0 på S genom att låta Σ_0 bestå av alla *ändliga disjunkta* unioner av element i $\mathcal{I}$. Det är helt klart att $\sigma(\Sigma_0) = \Sigma$. På Σ_0 definierar vi nu en lämplig prototyp till det *produktmått* som vi eftersträvar genom att för en mängd $\cup_{k=1}^n (A_k \times B_k) \in \Sigma_0$ sätta

$$\pi(\cup_{k=1}^n (A_k \times B_k)) = \sum_{k=1}^n \mu_1(A_k)\mu_2(B_k).$$

Om vi kan visa att π är en uppräkneligt additiv mängdfunktion kan vi använda Carathéodorys utvidgningssats till att dra slutsatsen att π kan utvidgas till ett mått på Σ. Om μ_1 och μ_2 är σ-ändliga följer det av unicitetssatsen att utvidgningen är unik.

Proposition 8.1 *Mängdfunktionen π är uppräkneligt additiv.*

Bevis. Vi ska visa att om $A_n \in \Sigma_1$, $B_n \in \Sigma_2$, $A_n \times B_n$ är disjunkta för olika n och $\cup_{n=1}^\infty (A_n \times B_n) \in \Sigma_0$ gäller att $\pi(\cup_n (A_n \times B_n)) = \sum_n \mu_1(A_n)\mu_2(B_n)$. Vi kan utan att förlora generalitet anta att $\cup_n (A_n \times B_n)$ är av formen $A \times B$, $A \in \Sigma_1$, $B \in \Sigma_2$ och det gäller då att visa att $\mu_1(A)\mu_2(B) = \sum_n \mu_1(A_n)\mu_2(B_n)$. Vi gör det i två steg: Först observerar vi som steg 1 att för varje $s_2 \in S_2$ har vi att

$$\mu_1(A) I_B(s_2) = I_B(s_2) \int_{S_1} I_A(s_1)\mu_1(ds_1) = \int_{S_1} I_B(s_2) I_A(s_1)\mu_1(ds_1)$$

$$= \sum_n \int_{S_1} I_{B_n}(s_2) I_{A_n}(s_1) \mu_1(ds_1) = \sum_n I_{B_n}(s_2) \int_{S_1} I_{A_n}(s_1) \mu_1(ds_1)$$

$$= \sum_n I_{B_n}(s_2) \mu_1(A_n).$$

Med hjälp av detta faktum får vi som steg 2 att

$$\mu_1(A)\mu_2(B) = \int_{S_2} \mu_1(A) I_B(s_2) \mu_2(ds_2) = \sum_n \mu_1(A_n) \int_{S_2} I_{B_n}(s_2) \mu_2(ds_2)$$

$$= \sum_n \mu_1(A_n) \mu_2(B_n)$$

och saken är klar. $\square$

En utvidgning av π till Σ kallar vi för ett produktmått av μ_1 och μ_2. Om μ_1 och μ_2 är σ-ändliga finns det som sagt ett unikt produktmått och detta brukar betecknas $\mu_1 \times \mu_2$. Vi sätter $\mu = \mu_1 \times \mu_2$. Vi har alltså konstruerat $(S, \Sigma, \mu) = (S_1 \times S_2, \Sigma_1 \times \Sigma_2, \mu_1 \times \mu_2)$.

8.2 Produkter av uppräkneligt många sannolikhetsrum

Konstruktionen ovan generaliserar sig på ett uppenbart sätt till alla produkter av ett godtyckligt men ändligt antal måttrum $(S_1, \Sigma_1, \mu_1), \ldots, (S_n, \Sigma_n, \mu_n)$.

Om vi har måttrummen (S_i, Σ_i, P_i), $i = 1, 2, \ldots$ där P_i:na är sannolikhetsmått och vi vill skapa måttrummet $(\prod_{i=1}^{\infty} S_i, \prod_{i=1}^{\infty} \Sigma_i, \prod_{i=1}^{\infty} P_i)$ fungerar det på i stort sett samma sätt; $\Sigma = \prod_{i=1}^{\infty} \Sigma_i$ definieras som $\sigma(\rho_i, i = 1, 2, \ldots)$. Vi har att Σ genereras av π-systemet

$$\mathcal{I} = \{(\prod_{i=1}^{n} B_i) \times (\prod_{i>n} S_i) : B_i \in \Sigma_i, i = 1, \ldots, n, n = 1, 2, 3, \ldots\}.$$

och också av algebran Σ_0 som består av ändliga disjunkta unioner av mängder i $\mathcal{I}$. På en typisk mängd i $\mathcal{I}$ definierar vi mängdfunktionen π genom att sätta

$$\pi((\prod_{i=1}^{n} B_i) \times (\prod_{i>n} S_i)) = \prod_{i=1}^{n} P_i(B_i).$$

och utvidgar π till Σ_0 på uppenbart sätt. Man visar på samma sätt som ovan att π är uppräkneligt additiv och således kan utvidgas till ett unikt sannolikhetmått $P = \prod_{i=1}^{\infty} P_i$ på Σ. (Observera att det är viktigt att P_i:na är sannolikhetsmått, vilket framgår av definitionen av π.)

8.3 Funktioner och integraler på ett produktrum

Vi går nu tillbaka till fallet med produkten av två måttrum (S_i, Σ_i, μ_i), $i = 1, 2$ och vi antar för enkelhets skull att μ_1 och μ_2 är ändliga. Vi sätter som förut $(S, \Sigma, \mu) = (S_1 \times S_2, \Sigma_1 \times \Sigma_2, \mu_1 \times \mu_2)$.

I grundkurser i matematik får man på ett överslätande sätt lära sig hur man beräknar dubbelintegraler över ett givet område. Dessa är då definierade enligt det Riemannska tillvägagångssättet, dvs via över- och underfunktioner som är konstanta på rektanglar. Man får lära sig att man under vissa förutsättningar kan beräkna dubbelintegralen genom att integrera koordinatvis och att man också får kasta om integrationsordningen när man gör detta. Den sats som säger att detta är tillåtet går under namnet *Fubinis sats*, men den bevisas inte i grundkurser.

Nu, när vi har försett oss själva med en hel del kraftfull mått- och integrationsteori, skall vi bevisa Fubinis sats. Detta gör vi förstås i vår generella Lebesgueintegralssituation, dvs vi ska visa att under vissa förutsättningar gäller för en Σ-mätbar funktion f att $\int_S f(s_1, s_2)\mu(d(s_1, s_2)) = \int_{S_1}(\int_{S_2} f(s_1, s_2)\mu_2(ds_2))\mu_1(ds_1) = \int_{S_2}(\int_{S_1} f(s_1, s_2)\mu_1(ds_1))\mu_2(ds_2)$. Det första av dessa tre uttryck är väldefinierat av det enkla faktum att (S, Σ, μ) är ett måttrum. Det två andra uttrycken vet vi däremot i nuläget inte om de överhuvudtaget har någon mening. Vi måste för det första visa att $f(s_1, s_2)$ sedd som en funktion av endast en av koordinaterna är mätbar m.a.p. respektive σ-algebra och för det andra, när detta är gjort, visa att det man får när man integrerat ut den ena koordinaten är en mätbar funktion av den andra koordinaten.

Sätt för alla $s_1 \in S_1$ och $s_2 \in S_2$, $f_{s_1}(s_2) = f_{s_2}(s_1) = f(s_1, s_2)$. För varje fixt s_2 har vi då att $f_{s_2} : S_1 \to \overline{\mathbf{R}}$ och motsvarande för varje fixt s_1.

LEMMA 8.2 *Låt $f : S_1 \times S_2 \to \overline{\mathbf{R}}$ vara en Σ-mätbar funktion. För varje fixt $s_2 \in S_2$ gäller att f_{s_2} är en Σ_1-mätbar funktion och för varje fixt $s_1 \in S_1$ är f_{s_1} en Σ_2-mätbar funktion.*

Bevis. Låt först $\mathcal{H}$ vara klassen av begränsade Σ-mätbara funktioner f sådana att f_{s_1} är Σ_2-mätbar för alla s_1 och f_{s_2} är Σ_1-mätbar för alla s_2. Eftersom $\mathcal{H}$ trivialt innehåller klassen av alla funktioner av typen $I_{B_1 \times B_2}$, $B_1 \in \Sigma_1, B_2 \in \Sigma_2$ räcker det enligt monotonklass-satsen att visa att $\mathcal{H}$ är en monoton klass för att kunna dra slutsatsen att $\mathcal{H} = b\Sigma$. Att verifiera villkoren för detta är trivialt.

Om nu f är en godtycklig ickenegativ Σ-mätbar funktion låter vi på vanligt sätt en följd $\{f_n\}$ av begränsade funktioner vara sådan att $f_n \uparrow f$. Eftersom $f_{n,s_i} \uparrow f_{s_i}$ gäller då att f_{s_i} blir Σ_j-mätbar, $i, j = 1, 2$, $i \neq j$. Slutligen delar man upp en generell funktion f i f^+ och f^-. $\square$

Enligt Lemma 8.2 är följande funktioner väldefinierade:

$$g(s_1) = \int_{S_2} f(s_1, s_2)\mu_2(ds_2),$$

$$h(s_2) = \int_{S_1} f(s_1, s_2)\mu_1(ds_1).$$

förutsatt att f_{s_1} och f_{s_2} är antingen ickenegativa eller integrerbara.

LEMMA 8.3 *Låt f vara en ickenegativ eller integrerbar funktion i $m\Sigma$. Då gäller att $g \in m\Sigma_1$ och att $h \in m\Sigma_2$.*

Bevis. Låt först $\mathcal{H} = \{f \in b\Sigma : g \in m\Sigma_1, h \in m\Sigma_2\}$. För $f = I_{B_1 \times B_2}$, $B_1 \in \Sigma_1, B_2 \in \Sigma_2$, har vi att $g(s) = I_{B_1}(s_1)\mu_2(B_2)$ och $h(s_2) = I_{B_2}(s_2)\mu_1(B_1)$ vilka trivialt är mätbara m.a.p. respektive σ-algebra, dvs $f \in \mathcal{H}$. Använd nu monoton-klass-satsen för att visa att $\mathcal{H} = b\Sigma$.

Att lemmat håller för godtyckliga ickenegativa f följer nu av monoton-konvergens-satsen och genom att dela upp generella integrerbara f i f^+ och f^- följer lemmat av linjaritet hos integraler. □

SATS 8.4 (FUBINIS SATS) *(a) Om $f \in (m\Sigma)^+$ gäller att*

$$\int_S f d\mu = \int_{S_2} (\int_{S_1} f(s_1, s_2)\mu_1(ds_1))\mu_2(ds_2) = \int_{S_1} (\int_{S_2} f(s_1, s_2)\mu_2(ds_2))\mu_1(ds_1).$$

(b) Påståendet i (a) gäller även då $f \in L^1(S, \Sigma, \mu)$.

Obs! Man kan använda (a) för att kontrollera om $f \in L^1$, ty det är ju sant för varje $f \in m\Sigma$ att $|f| \in (m\Sigma)^+$ och att $f \in L^1$ betyder ju att $\int_S |f| d\mu < \infty$.

Bevis. Låt först $f = I_{B_1 \times B_2}$, $B_1 \in \Sigma_1, B_2 \in \Sigma_2$. Vi har då att

$$\int_{S_1} (\int_{S_2} I_{B_1 \times B_2}(s_1.s_2)\mu_2(ds_2))\mu_1(ds_1) = \int_{S_1} (I_{B_1}(s_1) \int_{S_2} I_{B_2}(s_2)\mu_2(ds_2))\mu_1(ds_1)$$

$$= \mu_1(B_1)\mu_2(B_2) = \int_S I_{B_1 \times B_2} d\mu.$$

Låt nu $\mathcal{H}$ vara klassen av alla begränsade funktioner $f \in b\Sigma$ som satisfierar påståendet i (a). Att konstantfunktionen 1 tillhör $\mathcal{H}$ är uppenbart och om $f_n \in \mathcal{H}$, $n = 1, 2, \ldots$, $f_n \uparrow f$ och f är begränsad följer det av monoton-konvergenssatsen att $f \in \mathcal{H}$. Att $\mathcal{H}$ är ett vektorrum är också uppenbart. (Observera dock att för att detta ska gälla är det viktigt att μ_1 och μ_2 är ändliga mått.) Nu följer satsen av monoton-klass-satsen för alla begränsade f. Detta utvidgas nu till (a) via monoton-konvergens-satsen och sedan till (b) via linjaritet hos integraler. □

Om μ_1 och μ_2 är σ-ändliga är även μ σ-ändligt och man kan dela upp rummet S i en uppräknelig union av delmängder på vilka μ är ändligt. Genom att tillämpa vår version av Fubinis sats på dessa delmängder och sedan summera ser man att satsen håller även i det σ-ändliga fallet. Om μ_1 eller μ_2 inte är σ-ändligt spricker däremot Fubinis sats. Detta är ingen överraskning eftersom produktmåttet då inte ens alltid är entydigt definierat. Följande exempel är standard för att visa hur det kan gå snett när ett av de två måtten inte är σ-ändligt.

Exempel. Låt $S_1 = S_2 = [0, 1]$ och $\Sigma_1 = \Sigma_2 = \mathcal{B}$. Låt μ_1 vara Lebesguemåttet och låt μ_2 vara det sk *räknemåttet*, dvs $\mu_2(F) = \#(F) =$ antal punkter i F. Då är μ_2 inte ett σ-ändligt mått. Låt nu F vara diagonalen, dvs $F = \{(x, y) \in S_1 \times S_2 : x = y\}$. (**Övning:** Kontrollera att $F \in \mathcal{B} \times \mathcal{B}$.) Då har vi att

$$\int_{S_1} (\int_{S_2} I_F d\mu_2) d\mu_1 = \int_{S_1} 1 d\mu_1 = 1$$

medan

$$\int_{S_2} (\int_{S_1} I_F d\mu_1) d\mu_2 = \int_{S_2} 0 d\mu_2 = 0.$$

Låt oss nu se på ett mer konstruktivt exempel i en situation där Fubinins sats fungerar.

Exempel. Låt $X : (\Omega, \mathcal{F}, P) \to ([0, \infty), \mathcal{B})$ vara en ickenegativ stokastisk variabel. Låt nu $S = \Omega \times [0, \infty)$ och $\Sigma = \mathcal{F} \times \mathcal{B}$ och låt $\mu = P \times Leb.$ Definiera mängden $A = \{(\omega, x) \in \Omega \times [0, \infty) : 0 \le x \le X(\omega)\}$. Vi har att $A \in \mathcal{F} \times \mathcal{B}$ (Verifiera!) och Fubinis sats säger att å ena sidan gäller att

$$\mu(A) = \int_S I_A d\mu = \int_0^\infty (\int_\Omega I_A(\omega, x) P(d\omega)) dx = \int_0^\infty P(X \ge x) dx$$

och å andra sidan gäller att

$$\mu(A) = \int_\Omega (\int_0^\infty I_A(\omega, x) dx) P(d\omega) = \int_\Omega (\int_0^\infty I_{[0, X(\omega)]}(x) dx) P(d\omega)$$

$$= \int_\Omega X(\omega) P(d\omega) = \mathbf{E}[X],$$

dvs vi har härlett den välkända formeln $\mathbf{E}[X] = \int_0^\infty P(X \ge x) dx$ för ickenegativa stokastiska variabler.

8.4 Tvådimensionella fördelningar och produktmått

På samma sätt som i det endimensionella fallet kan man om man har två stokastiska variabler, X och Y, definierade på samma sannolihetsrum $(\Omega, \mathcal{F}, P)$ definiera fördelningen för (X,Y) som $P_{X,Y}(A) = P((X,Y) \in A)$, $A \in \mathcal{B} \times \mathcal{B}$. Eftersom klassen $\{[-\infty, x] \times [-\infty, y] : x, y \in \overline{\mathbf{R}}\}$ är ett π-system som genererar $\mathcal{B} \times \mathcal{B}$ följer det att fördelningen entydigt bestäms av fördelningsfunktionen $F_{X,Y}$ given av

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y),$$

$x, y \in \overline{\mathbf{R}}$.

Om det gäller att $P_{X,Y} << Leb \times Leb$, dvs om (X,Y) har en kontinuerlig fördelning, finns det enligt Radon-Nikodyms sats en $\mathcal{B} \times \mathcal{B}$-mätbar funktion $f : \mathbf{R}^2 \to [0, \infty)$ sådan att

$$P_{X,Y}(A) = \int_A f d(Leb \times Leb) = \int_A f(x,y) dx dy$$

för alla $A \in \mathcal{B} \times \mathcal{B}$. Funktionen f kallas för täthetsfunktionen för (X,Y). Eftersom f är ickenegativ gäller enligt Fubinis sats att

$$P_{X,Y}(A) = \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} f(x,y) I_A(x,y) dy) dx.$$

Genom att sätta $A = B \times \mathbf{R}$, $B \in \mathcal{B}$ får man att $P(X \in B) = \int_B (\int_{-\infty}^{\infty} f(x,y) dy) dx$, dvs avbildningen $x \to \int_{-\infty}^{\infty} f(x,y) dy$ är en täthetsfunktion för X. Motsvarande gäller för Y.

Följande tre påståenden är ekvivalenta:

(i) X och Y är oberoende,

(ii) $P_{X,Y} = P_X \times P_Y$,

(iii) $F_{X,Y}(x,y) = F_X(x) F_Y(y)$ för alla $x, y \in \overline{\mathbf{R}}$.

Om dessutom $P_{X,Y} << Leb \times Leb$ är även följande påstående ekvivalent med de tre ovanstående påståendena:

(iv) $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ för nästan alla (m.a.p. $Leb \times Leb$) $x, y \in \mathbf{R}$.

Ta det som en övning att bevisa dessa ekvivalenser.

8.5 Produkter av Borel-σ-algebror

Antag att S_1 och S_2 är två topologiska rum och sätt $S = S_1 \times S_2$. Den s.k. produkttopologin på S definieras av att en delmängd av S sägs vara öppen om den kan skrivas som en union av mängder av typen $G_1 \times G_2$ där G_i är öppen i S_i. Borel-σ-algebran $\mathcal{B}(S)$ ges sedan som σ-algebran genererad av alla öppna delmängder av S som vanligt.

Låt oss nu jämföra $\mathcal{B}(S)$ med $\mathcal{B}(S_1) \times \mathcal{B}(S_2)$. Den senare av dessa är definierad som den minsta σ-algebra som gör koordinatavbildningarna $(s_1, s_2) \to s_i$, $i = 1, 2$, till mätbara funktioner. Koordinatavbildningarna är kontinuerliga funktioner så det följer av Sats 3.1(b) att dessa är $\mathcal{B}(S)$-mätbara. Således gäller att $\mathcal{B}(S_1) \times \mathcal{B}(S_2) \subseteq \mathcal{B}(S)$. För att visa den andra inklusionen skulle det räcka att visa att varje öppen delmängd, G, av S ligger i $\mathcal{B}(S_1) \times \mathcal{B}(S_2)$. Detta skulle följa om vi kunde visa att G kan skrivas som en *uppräknelig* union av mängder av typen $G_1 \times G_2$ där G_i är öppen i S_i. Men definitionen av produkttopologin säger ingenting om uppräknelighet och faktum är att i det generella fallet är det *falskt* att $\mathcal{B}(S) = \mathcal{B}(S_1) \times \mathcal{B}(S_2)$.

I de flesta tillämpningar är rummen S_1 och S_2 *sekundärt uppräkneliga*. Detta betyder att det finns en uppräknelig *bas* för topologin, dvs en uppräknelig klass $\{B_i\}$ av öppna mängder sådan att varje öppen mängd kan skrivas som en union av B_i:n. Om nu S_1 och S_2 är sekundärt uppräkneliga så finns det alltså uppräkneliga baser $\{B_i^{(1)}\}_{i=1}^{\infty}$ och $\{B_i^{(2)}\}_{i=1}^{\infty}$ för toplogin på S_1 respektive S_2. Enligt definition av produkttopologin på S är nu $\{B_i^{(1)} \times B_j^{(2)}\}_{i,j=1}^{\infty}$ en uppräknelig bas för denna (ty $(\cup_i A_i) \times (\cup_j C_j) = \cup_{i,j}(A_i \times C_j)$). Men detta betyder ju just att varje öppen delmängd av S kan skrivas som en uppräknelig union av produkter av öppna mängder precis som vi önskade oss ovan.

Slutsatsen är alltså att generellt gäller att $\mathcal{B}(S_1) \times B(S_2) \subseteq \mathcal{B}(S)$ och att likhet gäller då S_1 och S_2 är sekundärt uppräkneliga. Eftersom rummet $\mathbf{R}^n$ är sekundärt uppräkneligt för alla n (varje öppen mängd kan ju skrivas som en union av öppna rektanglar med hörn vars koordinater är rationella) har vi som följdsats:

PROPOSITION 8.5 $\quad \mathcal{B}(\mathbf{R}^n) = \mathcal{B}(\mathbf{R})^n.$

I fallet då S_1 och S_2 ej är sekundärt uppräkneliga kan en liten varning vara på sin plats. Antag att X_i, för $i = 1, 2$, är S_i-värda stokastiska variabler, dvs $X_i : \Omega \to S_i$ är $\mathcal{F}/\mathcal{B}(S_i)$-mätbar, $i = 1, 2$. Paret $X = (X_1, X_2)$ är då en avbildning från Ω till $S = S_1 \times S_2$. Vad har X för mätbarhetsegenskaper? För $B_1 \in \mathcal{B}(S_1)$ och $B_2 \in \mathcal{B}(S_2)$ gäller att $X^{-1}(B_1 \times B_2) = X_1^{-1}(B_1) \cap X_2^{-1}(B_2) \in \mathcal{F}$ och eftersom $\mathcal{B}(S_1) \times \mathcal{B}(S_2)$ är genererad av alla rektanglar av formen $B_1 \times B_2$, $B_i \in \mathcal{B}(S_i)$, följer det av Sats 3.1(c) (i dess generella form) att X är $\mathcal{F}/(\mathcal{B}(S_1) \times \mathcal{B}(S_2))$-mätbar. Man kan säga att X är en $(S, \mathcal{B}(S_1) \times \mathcal{B}(S_2))$-värd stokastisk variabel. Men (och det är här varningen kommer) eftersom $\mathcal{B}(S)$ kan vara en strikt större σ-algebra än $\mathcal{B}(S_1) \times \mathcal{B}(S_2)$ är det inte säkert att X är en $(S, \mathcal{B}(S))$-värd stokastisk variabel.

8.6 Alternativ konstruktion av en följd av oberoende stokastiska variabler

Tidigare konstruerade vi en följd av obereonde stokastiska variabler med hjälp av binärutvecklingar av talen i [0,1]. Här kommer ett mer elegant sätt. Vi vet sedan förut hur man med hjälp av Lebesguemåttet på [0,1] kan konstruera stokastiska variabler X_i på ett sannolikhetsrum $(\Omega_i, \mathcal{F}_i, P_i)$ på så sätt att X_i får någon önskad fördelningsfunktion F_i. Detta gör vi för $i = 1, 2, \ldots$. Problemet är nu att X_i:na är funktioner på olika utfallsrum och vi vill att de ska vara definierade på samma utfallsrum. Men sätt nu

$$(\Omega, \mathcal{F}, P) = (\prod_{i=1}^{\infty} \Omega_i, \prod_{i=1}^{\infty} \mathcal{F}_i, \prod_{i=1}^{\infty} P_i).$$

Definiera nu stokastiska variabler $Y_1, Y_2, \ldots$ genom att sätta

$$Y_i(\omega_1, \omega_2, \ldots) = X_i(\omega_i).$$

Enligt definitionen av produkt-σ-algebran är Y_i:na $\mathcal{F}$-mätbara och enligt definitionen av produktmåttet gäller för alla i att Y_i har samma fördelning som X_i och att Y_i:na är oberoende.

9 Betingade väntevärden

Antag att Z och X är diskreta stokastiska variabler sådana att Z tar sina värden i $\{z_1, \ldots, z_n\}$ och X tar sina värden i $\{x_1, \ldots, x_m\}$. Vi brukar definiera $\mathbf{E}[X|Z = z_j]$ som $\sum_i x_i P(X = x_i | Z = z_j) = \mathbf{E}[X; Z = z_j]/P(Z = z_j)$. Vi säger sedan att $\mathbf{E}[X|Z]$ är den stokastiska variabel Y sådan att $Y(\omega) = \mathbf{E}[X|Z = z_j]$ för de ω för vilka $Z(\omega) = z_j$.

Att definiera $\mathbf{E}[X|Z]$ på detta sätt går bra eftersom X och Z är diskreta, men om Z är kontinuerlig gäller alltid att $P(Z = z) = 0$ och då är det inte lika lätt att veta vad $\mathbf{E}[X|Z = z]$ ska betyda. I grundkurser brukar man lösa detta via betingade tätheter. Här ska vi använda en teori baserad på Radon-Nikodyms sats som ger en gemensam definition för alla typer av stokastiska variabler, såväl diskreta som kontinuerliga och alla blandformer.

Observera om den stokastiska variabeln $Y = \mathbf{E}[X|Z]$ som vi definierade i det diskreta fallet ovan att

- Y är $\sigma(Z)$-mätbar, ty $\{\omega \in \Omega : Y(\omega) = \mathbf{E}[X|Z = z_j]\} = \{\omega \in \Omega : Z(\omega) = z_j\}$.
- $\int_{\{Z=z_j\}} Y dP = \mathbf{E}[X|Z = z_j] P(Z = z_j) = \mathbf{E}[X; Z = z_j] = \int_{\{Z=z_j\}} X dP$.

Detta kan vi skriva om som:

- $\mathbf{E}[X|Z]$ är $\sigma(Z)$-mätbar.
- $\int_G \mathbf{E}[X|Z] dP = \int_G X dP$ för alla $G \in \sigma(Z)$.

För en vettig definition av betingad väntevärde borde dessa observationer gälla generellt. Därför *definierar* vi helt enkelt det betingade väntevärdet så:

DEFINITION. Låt $(\Omega, \mathcal{F}, P)$ vara ett sannolikhetsrum och låt $\mathcal{G}$ vara en del-σ-algebra till $\mathcal{F}$. För en stokastisk variabel $X \in L^1(\Omega, \mathcal{F}, P)$ säger vi att Y är en *version* av det betingade väntevärdet $\mathbf{E}[X|\mathcal{G}]$ av X givet $\mathcal{G}$ om

- Y är $\mathcal{G}$-mätbar,
- $\int_G Y dP = \int_G X dP$ för alla $G \in \mathcal{G}$.

Existensen av $\mathbf{E}[X|\mathcal{G}]$ är en direkt konsekvens av Radon-Nikodyms sats då X är ickenegativ, ty mängdfunktionen $\int_G X dP$, $G \in \mathcal{G}$ är ett mått $\mathcal{G}$ som är absolutkontinuerligt m.a.p. P. I det generella fallet delar vi upp X i X^+ och X^- och använder Radon-Nikodyms sats på dessa.

Att vi talar om en "version" av det betingade väntevärdet beror på att $\mathbf{E}[X|\mathcal{G}]$ inte är helt entydigt definierad; om Y och Y' är två versioner av $\mathbf{E}[X|\mathcal{G}]$ följer det av Sats 5.10 att $Y = Y'$ *nästan säkert*. Detta betyder automatiskt att alla generella utsagor om det betingade väntevärdet kommer att vara på nivån "nästan säkert".

Om Z är en stokastisk variabel definierar vi $\mathbf{E}[X|Z]$ som $\mathbf{E}[X|\sigma(Z)]$.

Obs! Vid en första anblick verkar det kanske som om man skulle kunna ta $\mathbf{E}[X|\mathcal{G}] = X$, men eftersom X oftast inte är $\mathcal{G}$-mätbar stämmer inte detta.

Existensen av det betingade väntevärdet följer som sagt av Radon-Nikodyms sats. Om det nu är så att man har hoppat över beviset av Radon-Nikodyms sats så hänger alltså existensen av betingade väntevärden i luften. Som tur är finns det i det läget ett alternativt bevis baserat på ortogonalprojektioner i L^2:

SATS 9.1 *Antag att $X \in L^1(\Omega, \mathcal{F}, P)$ och att $\mathcal{G}$ är en del-σ-algebra till $\mathcal{F}$. Då existerar en stokastisk variabel $Y \in L^1(\Omega, \mathcal{G}, P)$ sådan att*

$$\int_G Y dP = \int_G X dP$$

för alla $G \in \mathcal{G}$. Om Y' är en annan sådan stokastisk variabel gäller att $Y = Y'$ n.s.

Bevis. Entydighetsdelen har vi redan diskuterat, så vi koncenterar oss på existensen. Antag först att $X \in L^2(\Omega, \mathcal{F}, P)$ och sätt $\mathcal{K} = L^2(\Omega, \mathcal{G}, P)$. Enligt Sats 6.7 är $\mathcal{K}$ ett fullständigt underrum till $L^2(\mathcal{F})$ och eftersom $\mathcal{K}$ uppenbarligen är ett vektorrum följer det av Sats 6.8 att det existerar en stokastisk variabel $Y \in L^2(\mathcal{G})$ sådan att $\langle X - Y, Z\rangle = 0$ för alla $Z \in L^2(\mathcal{G})$. Genom att för en godtycklig mängd $G \in \mathcal{G}$ sätta $Z = I_G$ får vi

$$\langle X - Y, Z\rangle = \int_G (X - Y) dP = \int_G X dP - \int_G Y dP = 0$$

dvs Y är just en sådan stokastisk variabel som vi sökte.

Antag nu att X är en ickenegativ integrerbar stokastisk variabel. Sätt $X_n = X \wedge n$. Då har vi att $X_n \uparrow X$ och att $X_n \in L^2$ för alla n. Enligt vad vi nyss visade finns alltså stokastiska variabler $Y_n \in L^2(\mathcal{G})$ sådana att det för alla $G \in G$ gäller att $\int_G Y_n dP = \int_G X_n dP$. Enligt monoton-konvergens-satsen gäller att högersidan konvergerar mot $\int_G X dP$. Eftersom $X_n \leq X_{n+1}$ gäller att $\int_G Y_n dP \leq \int_G Y_{n+1} dP$ för alla $G \in \mathcal{G}$ så enligt Sats 5.10 gäller att $Y_n \leq Y_{n+1}$ n.s. Tack vare att en uppräknelig union av nollmängder är en nollmängd gäller alltså att $\{Y_n(\omega)\}$ är växande för nästan alla ω. Genom att justera Y_n:na på undantagsmängden kan vi tillse att $\{Y_n\}$ blir växande överallt. Således existerar en gränsfunktion $Y = \lim_{n\to\infty} Y_n$. Enligt monoton-konvergens-satsen gäller nu att $\int_G Y_n dP \uparrow \int_G Y dP$, dvs Y är en version av $\mathbf{E}[X|\mathcal{G}]$.

Till sist för generellt X, sätt $Y^+ = \mathbf{E}[X^+|\mathcal{G}]$ och $Y^- = \mathbf{E}[X^-|\mathcal{G}]$ och $Y = Y^+ - Y^-$. $\square$

Exempel. Fallet med kontinuerliga stokastiska variabler: Antag att (X, Z) är en tvådimensionell stokastisk variabel med täthetsfunktion $f_{X,Z}(x, z)$ och sätt

$$f_{X|Z}(x|z) = \frac{f_{X,Z}(x,z)}{f_Z(z)} I_{\{f_Z>0\}}(z).$$

Låt $h : \mathbf{R} \to \mathbf{R}$ vara en funktion sådan att $\mathbf{E}|h(X)| < \infty$. Enligt erfarenhet från grundkurser borde det gälla att

$$\mathbf{E}[h(X)|Z = z] = \int_{\mathbf{R}} h(x) f_{X|Z}(x|z) dx$$

dvs om vi betecknar högerledet $g(z)$ så borde $g(Z)$ vara en version av $\mathbf{E}[h(X)|Z]$. Låt oss kontrollera om detta är sant. Vi ska då kontrollera om $\int_{\{Z\in B\}} g(Z) dP = \int_{\{Z\in B\}} h(X) dP$ för alla $B \in \mathcal{B}$ (kom ihåg att $\sigma(Z) = \{\{Z \in B\} : B \in \mathcal{B}\}$), dvs om $\mathbf{E}[g(Z)I_B(Z)] = \mathbf{E}[h(X)I_B(Z)]$ för alla $B \in \mathcal{B}$. Detta är sant ty

$$\mathbf{E}[g(Z)I_B(Z)] = \int_{\mathbf{R}} g(z) I_B(z) f_Z(z) dz$$

$$= \int_{\mathbf{R}} (\int_{\mathbf{R}} h(x) f_{X|Z}(x|z) dx) f_Z(z) I_B(z) dz = \int_{\mathbf{R}} (\int_{\mathbf{R}} h(x) I_B(z) f_{X,Z}(x,z) dx) dz$$

$$= \int_{\mathbf{R}^2} h(x) I_B(z) f_{X,Z}(x,z) d(x,z) = \mathbf{E}[h(X)I_B(Z)]$$

där vi i de olika beräkningsleden kombinerat vad vi sett tidigare i några olika exempel och där den näst sista likheten följer av Fubinis sats.

Här följer nu en mastodontsats som stadfäster de viktigaste grundläggande egenskaperna hos betingade väntevärden.

SATS 9.2 *I samtliga av följande påståenden förutsätter vi att de stokastiska variabler som omnämns tillhör $L^1(\Omega, \mathcal{F}, P)$ och att $\mathcal{G}$ och $\mathcal{H}$ är del-σ-algebror till $\mathcal{F}$.*

(a) $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]] = \mathbf{E}[X]$.

(b) Om X är $\mathcal{G}$-mätbar är $\mathbf{E}[X|\mathcal{G}] = X$ n.s.

(c) $\mathbf{E}[(a_1X_1 + a_2X_2)|\mathcal{G}] = a_1\mathbf{E}[X_1|\mathcal{G}] + a_2\mathbf{E}[X_2|\mathcal{G}]$ *n.s.*

(d) Om $X \geq 0$ n.s. gäller att $\mathbf{E}[X|\mathcal{G}] \geq 0$ n.s.

(e) Monoton-konvergens-satsen: Om $X_n \geq 0$ n.s. för alla n och $X_n \uparrow X$ n.s. gäller att $\mathbf{E}[X_n|\mathcal{G}] \uparrow \mathbf{E}[X|\mathcal{G}]$ n.s.

(f) Fatous lemma: Om $X_n \geq 0$ n.s. för alla n gäller att $\mathbf{E}[\liminf_n X_n|\mathcal{G}] \leq \liminf_n \mathbf{E}[X_n|\mathcal{G}]$ n.s.

(g) Dominerad-konvergens-satsen: Om $X_n \to X$ n.s. och det existerar en stokastisk variabel V sådan att $|X_n| \leq V$ n.s. för alla n och $\mathbf{E}[V] < \infty$ gäller att $\mathbf{E}[X_n|\mathcal{G}] \to \mathbf{E}[X|\mathcal{G}]$ n.s.

(h) Jensens olikhet: Om $c : \mathbf{R} \to \mathbf{R}$ är en konvex funktion och $\mathbf{E}|c(X)| < \infty$ gäller att $\mathbf{E}[c(X)|\mathcal{G}] \geq c(\mathbf{E}[X|\mathcal{G}])$ n.s.

(i) Tornegenskapen: Om $\mathcal{H} \subseteq \mathcal{G}$ gäller att $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbf{E}[\mathbf{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbf{E}[X|\mathcal{H}]$.

(j) Bryt ut det som är känt: Om Z är en begränsad $\mathcal{G}$-mätbar stokastisk variabel gäller att $\mathbf{E}[ZX|\mathcal{G}] = Z\mathbf{E}[X|\mathcal{G}]$ n.s. Samma sak gäller om antingen $Z \geq 0$, $X \geq 0$ och $\mathbf{E}[ZX] < \infty$ eller om $1/p + 1/q = 1$ och $X \in L^p$ och $Z \in L^q$.

*(k) Om $\mathcal{H}$ är oberoende av $\sigma(\mathcal{G}, \sigma(X))$ gäller att $\mathbf{E}[X|\sigma(\mathcal{G}, \mathcal{H})] = \mathbf{E}[X|\mathcal{G}]$ n.s. Speciellt gäller att om $\mathcal{H}$ är oberoende av X är $\mathbf{E}[X|\mathcal{H}] = \mathbf{E}[X]$ n.s. (Sätt $\mathcal{G} = \{\emptyset, \Omega\}$.) (***Varning:*** Om $\mathcal{H}$ är oberoende av $\mathcal{G}$ och av X betyder detta **inte** att $\mathcal{H}$ är oberoende av $\sigma(\mathcal{G}, \sigma(X))$. Kom ihåg att tre stokastiska variabler inte säkert är oberoende även om de är parvis oberoende.)*

Bevis. Eftersom $\mathcal{G}$ är en σ-algebra har vi att $\Omega \in \mathcal{G}$ och således att $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]] = \int_\Omega \mathbf{E}[X|\mathcal{G}]dP = \int_\Omega XdP = \mathbf{E}[X]$ vilket bevisar (a). Del (b) är trivial och del (c) följer av det faktum att $\int_G (a_1X_1 + a_2X_2)dP = a_1\int_G X_1dP + a_2\int_G X_2dP$ för alla $G \in \mathcal{G}$. Del (d) följer av Sats 5.10 ty $\int_G \mathbf{E}[X|\mathcal{G}]dP = \int_G XdP \geq 0$ för alla $G \in \mathcal{G}$. För att visa del (e) påminner vi oss att vi i beviset av Sats 9.1 ovan visade att följden $\{\mathbf{E}[X_n|\mathcal{G}]\}$ är växande om man väljer lämpliga versioner av de betingade väntevärdena. Om vi sätter $Y = \lim_{n\to\infty} \mathbf{E}[X_n|\mathcal{G}]$ har vi enligt monoton-konvergens-satsen för alla $G \in \mathcal{G}$ att

$$\int_G YdP = \lim_{n\to\infty}\int_G \mathbf{E}[X_n|\mathcal{G}]dP = \lim_{n\to\infty}\int_G X_ndP = \int_G XdP = \int_G \mathbf{E}[X|\mathcal{G}]dP,$$

dvs $Y = \mathbf{E}[X|\mathcal{G}]$ n.s.

Med hjälp av (e) klarar vi (f) på samma sätt som man bevisar vanliga Fatous lemma med hjälp av monoton-konvergens-satsen:

$$\mathbf{E}[\liminf_n X_n|\mathcal{G}] = \lim_{n\to\infty} \mathbf{E}[\inf_{m\geq n} X_m|\mathcal{G}] \leq \liminf_n \mathbf{E}[X_n|\mathcal{G}]$$

där alla relationer gäller n.s. och där olikheten följer av det faktum att $X \leq Y \Rightarrow \mathbf{E}[X|\mathcal{G}] \leq \mathbf{E}[Y|\mathcal{G}]$ n.s. vilket i sin tur följer av att kombinera (d) och (c). För att bevisa del (g) obeserverar vi att genom att tillämpa (f) på följden $\{V - X_n\}$ och utnyttja (c) följer omvända Fatous lemma i betingad version, dvs $\mathbf{E}[\limsup_n X_n|\mathcal{G}] \geq \limsup_n \mathbf{E}[X_n|\mathcal{G}]$ n.s. Vi får då att

$$\mathbf{E}[X|\mathcal{G}] = \mathbf{E}[\liminf_n X_n|\mathcal{G}] \leq \liminf_n \mathbf{E}[X_n|\mathcal{G}]$$

$$\leq \limsup_n \mathbf{E}[X_n|\mathcal{G}] \leq \mathbf{E}[\limsup_n X_n|\mathcal{G}] = \mathbf{E}[X|\mathcal{G}]$$

(där alla relationer gäller n.s.) som önskat.

Kom nu ihåg stödlinjessatsen som sade att för den konvexa funktionen $c(x)$ gäller att det finns två följder, $\{a_n\}$ och $\{b_n\}$ av reella tal sådana att $c(x) = \sup_n(a_n x + b_n)$ för alla x. Således har vi att $\mathbf{E}[c(X)|\mathcal{G}] = \mathbf{E}[\sup_n(a_n X + b_n)|\mathcal{G}] \geq \sup_n(a_n\mathbf{E}[X|\mathcal{G}] + b_n) = c(\mathbf{E}[X|\mathcal{G}])$ n.s. (Observera att eftersom alla relationer mellan betingade väntevärden bara gäller n.s. så är det viktigt att $\{a_n\}$ och $\{b_n\}$ är uppräkneliga mängder.)

Låt oss nu gripa oss an tornegenskapen. Eftersom $\mathcal{H} \subseteq \mathcal{G}$ följer det direkt av (b) att $\mathbf{E}[\mathbf{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbf{E}[X|\mathcal{H}]$. Dessutom gäller för en godtycklig mängd $H \in \mathcal{H}$ att H också tillhör $\mathcal{G}$ och därför säger definitionen av betingat väntevärde att

$$\int_H \mathbf{E}[\mathbf{E}[X|\mathcal{G}]|\mathcal{H}]dP = \int_H \mathbf{E}[X|\mathcal{G}]dP = \int_H X dP = \int_H \mathbf{E}[X|\mathcal{H}]dP$$

och resultatet följer av Sats 5.10.

För att visa del (j) ska vi visa att $\int_G ZX dP = \int_G Z\mathbf{E}[X|\mathcal{G}]dP$ för alla $G \in \mathcal{G}$. Tack vare (c) kan vi utan att förlora generalitet anta att $X \geq 0$. Antag först att $Z = I_{G'}$ för någon mängd $G' \in \mathcal{G}$. Då gäller att

$$\int_G ZX dP = \int_{G\cap G'} X dP = \int_{G\cap G'} \mathbf{E}[X|\mathcal{G}]dP = \int_G Z\mathbf{E}[X|\mathcal{G}]dP$$

ty $G \cap G' \in \mathcal{G}$. Genom att använda linjaritet följer att resultatet gäller för alla enkla Z och genom att tillämpa monoton-konvergens-satsen följer resultatet för alla ickenegativa Z sådana att $ZX \in L^1$. Slutligen ger linjariteten igen att resultatet är sant för alla $Z \in m\mathcal{G}$ sådana att $ZX \in L^1$. Det sistnämnda kravet är uppfyllt då Z är begränsad p.g.a. det faktum att $X \in L^1$ och följer av Hölders olikhet då $X \in L^p$ och $Z \in L^q$ och $1/p + 1/q = 1$.

Låt oss slutligen visa del (k). Tack vare (c) kan vi återigen anta att $X \geq 0$. Vad vi ska visa är att $\int_F X dP = \int_F \mathbf{E}[X|\mathcal{G}]dP$ för alla $F \in \sigma(\mathcal{G}, \mathcal{H})$ för att sedan kunna dra den önskade slutsatsen av Sats 5.10. Men kom ihåg att om vi ser dessa två uttryck som funktioner av F så är de mått. För att visa att de två måtten är lika räcker det om vi kan visa att de är lika för alla mängder i π-systemet $\{G \cap H : G \in \mathcal{G}, H \in \mathcal{H}\}$. Observera nu att för en mängd $G \in \mathcal{G}$ gäller att den stokastiska variabeln XI_G är $\sigma(\mathcal{G}, \sigma(X))$-mätbar och därför oberoende av $\mathcal{H}$. Likadant är $\mathbf{E}[X|\mathcal{G}]I_G$ oberoende av $\mathcal{H}$. Därför gäller

$$\int_{G\cap H} X dP = \mathbf{E}[XI_G I_H] = \mathbf{E}[XI_G]P(H) = \mathbf{E}[\mathbf{E}[X|\mathcal{G}]I_G]P(H)$$

$$= \mathbf{E}[\mathbf{E}[X|\mathcal{G}]I_G I_H] = \int_{G\cap H} \mathbf{E}[X|\mathcal{G}]dP$$

där den tredje likheten följer av definitionen av $\mathbf{E}[X|\mathcal{G}]$. Saken är klar. $\square$

Exempel. Tillämpning av betingade Jensens olikhet: Sätt $c(x) = |x|^p$ och få att $\mathbf{E}[|X|^p|\mathcal{G}] \geq |\mathbf{E}[X|\mathcal{G}]|^p$ n.s. Tag nu väntevärde av bägge sidor och få att $\mathbf{E}[|X|^p] \geq \mathbf{E}[|\mathbf{E}[X|\mathcal{G}]|^p]$. Genom att nu ta p:te roten ur bägge sidor får vi att $||X||_p \geq ||\mathbf{E}[X|\mathcal{G}]||_p$.

Anmärkning. Det betingade väntevärdet $\mathbf{E}[X|\mathcal{G}]$ definieras genom att vi kräver att det för varje $G \in \mathcal{G}$ ska gälla att $\int_G \mathbf{E}[X|\mathcal{G}]dP = \int_G X dP$. Detta kan naturligtvis skrivas som att $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]; G] = \mathbf{E}[X; G]$ eller som $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]I_G] = \mathbf{E}[XI_G]$. Använd det skrivsätt som känns bäst!

9.1 Reguljära betingade sannolikheter

För $F \in \mathcal{F}$ har vi att $P(F) = \mathbf{E}[I_F]$. På motsvarande sätt definierar vi den betingade sannolikheten av F givet $\mathcal{G}$, $P(F|\mathcal{G})$, som $\mathbf{E}[I_F|\mathcal{G}]$. Den betingade sannolikheten $P(F|\mathcal{G})$ är en stokastisk variabel och därmed en funktion av $\omega \in \Omega$, men också en funktion av F. En naturlig fråga är om mängdfunktionen $P(\cdot|\mathcal{G})(\omega)$ är ett sannolikhetsmått för alla ω, eller om man ska vara noga för nästan alla ω eftersom det finns olika versioner av $P(\cdot|\mathcal{G})$.

För att $P(\cdot|\mathcal{G})$ ska vara ett sannolikhetsmått krävs att om $F_1, F_2, \ldots$ är en följd av disjunkta händelser i $\mathcal{F}$ så är $P(\cup_n F_n|\mathcal{G}) = \sum_n P(F_n|\mathcal{G})$. Genom att använda den betingade versionen av monoton-konvergens-satsen inser vi att detta gäller *nästan säkert*, dvs $P(\cup_n F_n|\mathcal{G})(\omega) = \sum_n P(F_n|\mathcal{G})(\omega)$ för P-n.a. ω. Men detta betyder att för just följden $\{F_n\}$ kan det finnas en undantagsmängd med sannolikhetsmåttet 0. Eftersom det utom i de enklaste fallen finns överuppräkneligt många följder av disjunkta mängder i $\mathcal{F}$ kan unionen av alla undantagsmängderna, om den överhuvudtaget är mätbar, ha ett mått som är betydligt större än 0. Det är alltså inte uppenbart att $P(\cdot|\mathcal{G})(\omega)$ är ett sannolikhetsmått för något ω alls. Om så ändå är fallet talar vi om en *reguljär* betingad sannolikhet:

DEFINITION. Antag att $P(\cdot,\cdot) : \Omega \times \mathcal{F} \to [0,1]$ är en avbildning sådan att

- för alla $F \in \mathcal{F}$ är $P(\cdot, F)$ en version av $P(F|\mathcal{G})$,
- för nästan alla $\omega \in \Omega$ är $P(\omega, \cdot)$ ett sannolikhetsmått på $\mathcal{F}$.

Då kallas $P(\cdot,\cdot)$ för en *reguljär betingad sannolikhet* givet $\mathcal{G}$.

För de flesta sannolikhetsrum $(\Omega, \mathcal{F}, P)$ som man stöter på finns det reguljära betingade sannolikheter. Exempelvis gäller detta då $\Omega = \mathbf{R}$ och $\mathcal{F} = \mathcal{B}$. Vi går inte djupare in på detta. Ett bevis kan man till exempel finna i R. B. Ash, "Real Analysis and Probability," Academic Press, 1972.

Ett specialfall av en reguljär betingad sannolikhet får vi om vi bygger ut exemplet där (X, Z) var tvådimensionellt kontinuerligt fördelad med täthet $f_{X,Z}(x, z)$. Vi visade där att om man sätter $\mathbf{E}[h(X)|Z](\omega) = \int_{\mathbf{R}} h(x) f_{X|Z}(x|Z(\omega))dx$ får man en version av $\mathbf{E}[h(X)|Z]$. Genom att sätta $h = I_B$, $B \in \mathcal{B}$ får man $P(X \in B|Z)(\omega) = \int_B f_{X|Z}(x|Z(\omega))dx$ som definierar ett mått på $\sigma(X)$ för varje ω. Således är $P(\cdot|Z)(\cdot)$ en reguljär betingad sannolikhet givet $\sigma(Z)$ definierad på $\sigma(X)$.

9.2 Ett symmetriresultat

Följande lemma är ett viktigt redskap när man bevisar stora talens lag med hjälp av martingalteori och beviset är en nyttig övning på betingade väntevärden.

LEMMA 9.3 *Låt $X_1, X_2, \ldots$ vara en följd av oberoende likafördelade stokastiska variabler med $\mathbf{E}|X_i| < \infty$. Sätt $S_n = \sum_{i=1}^n X_i$ och låt $\mathcal{G}_n = \sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n, X_{n+1}, X_{n+2}, \ldots)$. Det gäller att $\mathbf{E}[X_1|\mathcal{G}_n] = \mathbf{E}[X_1|S_n] = S_n/n$ n.s.*

Bevis. Eftersom $\sigma(X_1, S_n) \subseteq \sigma(X_1, \ldots, X_n)$ gäller att $\sigma(X_1, S_n)$ är oberoende av $\sigma(X_{n+1}, X_{n+2}, \ldots)$. Därför är Sats 9.2(k) tillämpbar och medför att $\mathbf{E}[X_1|\mathcal{G}_n] = \mathbf{E}[X_1|S_n]$.

För att visa den andra likheten skall vi visa att $\int_{\{S_n \in B\}} \mathbf{E}[X_1|S_n]dP = \int_{\{S_n \in B\}} (S_n/n)dP$ för alla $B \in \mathcal{B}$. Men

$$\int_{\{S_n \in B\}} \mathbf{E}[X_1|S_n]dP = \int_{\{S_n \in B\}} X_1 dP = \int_\Omega X_1 I_B(S_n)dP$$

$$= \int_{\mathbf{R}} \cdots \int_{\mathbf{R}} x_1 I_B(x_1 + \ldots + x_n) dP_{X_1} \ldots dP_{X_n} = \ldots = \int_{\{S_n \in B\}} \mathbf{E}[X_k|S_n]dP$$

för alla $k \in \{1, \ldots, n\}$ enligt Fubinis sats och det faktum att $P_{X_1} = \ldots = P_{X_n}$. Alltså gäller n.s. att $\mathbf{E}[X_1|S_n] = \ldots = \mathbf{E}[X_n|S_n]$ och eftersom summan av dessa är $\mathbf{E}[S_n|S_n] = S_n$ n.s. gäller att $\mathbf{E}[X_1|S_n] = S_n/n$ n.s. $\square$

10 Introduktion av martingalteori

Vi börjar med den grundläggande terminologin. Som vanligt har vi i botten sannolikhetsrummet $(\Omega, \mathcal{F}, P)$. En *filtrering* är en *växande* följd $\{\mathcal{F}_n\}_{n=0}^{\infty}$ av del-σ-algebror till $\mathcal{F}$, dvs en följd sådan att $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots \subseteq \mathcal{F}$. Med en *stokastisk process* ska vi i fortsättningen avse en stokastisk process i "diskret tid", dvs helt enkelt en följd av stokastiska variabler $\{X_n\}_{n=0}^{\infty}$. En stokastisk process $\{X_n\}$ sägs vara *adapterad* till filtreringen $\{\mathcal{F}_n\}$ om det för varje n gäller att X_n är $\mathcal{F}_n$-mätbar. Det vanligaste specialfallet av en filtrering får man genom att för en given stokastisk process $\{W_n\}$ helt enkelt sätta $\mathcal{F}_n = \sigma(W_0, \ldots, W_n)$. Processen $\{W_n\}$ blir då förstås automatiskt adapterad till denna filtrering och genom att använda en multivariat version av Proposition 3.5 följer det att en annan process $\{X_n\}$ är adapterad till denna filtrering om och endast om det för varje n finns en $\mathcal{B}(\mathbf{R}^{n+1})$-mätbar funktion sådan att $X_n = f(W_0, \ldots, W_n)$. Martingaler är speciella fall av adapterade processer:

Definition. Låt $\{\mathcal{F}_n\}_{n=0}^{\infty}$ vara en filtrering och låt $\{X_n\}_{n=0}^{\infty}$ vara en stokastisk process. Följden $\{X_n, \mathcal{F}_n\}$ sägs vara en *submartingal* om $\{X_n\}$ är adapterad till $\{\mathcal{F}_n\}$ och det för varje n gäller att $\mathbf{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$ n.s. Om istället $\mathbf{E}[X_{n+1}|\mathcal{F}_n] \leq X_n$ n.s. för alla n sägs $\{X_n, \mathcal{F}_n\}$ vara en *supermartingal.* Vi säger att följden $\{X_n, \mathcal{F}_n\}$ är en *martingal* om den är både en submartingal och en supermartingal.

Ofta, då filtreringen är underförstådd, kommer vi att rätt och slätt säga att $\{X_n\}$ är en (sub/super)martingal.

Om vi säger att $\{X_n\}$ är en (sub/super)martingal utan att överhuvudtaget ha specificerat en filtrering menar vi automatiskt att $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.

Låt oss göra ett par omedelbara observationer: Om $\{X_n\}$ är en martingal har vi för $m = 1, 2, \ldots$ att $\mathbf{E}[X_{n+m}|\mathcal{F}_n] = \mathbf{E}[\mathbf{E}[X_{n+m}|\mathcal{F}_{n+m-1}]|\mathcal{F}_n] = \mathbf{E}[X_{n+m-1}|\mathcal{F}_n] = \ldots = X_n$, med motsvarande olikhet i sub- respektive supermartingalfallen. Det gäller också att processen $\{Y_n\}$, där $Y_n = X_n - X_0$, är en (sub/super)martingal sådan att $Y_0 \equiv 0$.

Exempel. Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler sådana att $\mathbf{E}[X_k] = 0$ för alla k. Sätt $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ och $S_n = \sum_{k=1}^{n} X_k$. (Vi sätter $S_0 = 0$ och $\mathcal{F}_0 = \{\emptyset, \Omega\}$.) Då är $\{S_n, \mathcal{F}_n\}$ en martingal ty

$$\mathbf{E}[S_{n+1}|\mathcal{F}_n] = \mathbf{E}[S_n|\mathcal{F}_n] + \mathbf{E}[X_{n+1}|\mathcal{F}_n] = S_n + \mathbf{E}[X_{n+1}] = S_n \text{ n.s.}$$

Om vi istället har att $\mathbf{E}[X_k] = \mu$ för alla k blir $\{S_n\}$ en supermartingal om $\mu \leq 0$ och en submartingal om $\mu \geq 0$.

Exempel. Tag en stokastisk variabel $X \in L^1(\Omega, \mathcal{F}, P)$ och låt $\{\mathcal{F}_n\}$ vara en godtycklig filtrering. Då är följden $\{\mathbf{E}[X|\mathcal{F}_n], \mathcal{F}_n\}$ en martingal, ty $\mathbf{E}[X|\mathcal{F}_n]$ är per definition $\mathcal{F}_n$-mätbar och

$$\mathbf{E}[\mathbf{E}[X|\mathcal{F}_{n+1}]|\mathcal{F}_n] = \mathbf{E}[X|\mathcal{F}_n]$$

enligt tornegenskapen hos betingade väntevärden.

Vanliga tillämpningsområden för martingalteori är spelteori, diffusionteori, stokastiska differentialekvationer och förgreningsprocesser. Det är ofta naturligt att tänka i termer av spelteori. Då tolkar man en martingal $\{X_n\}$ som att man går in i ett rättvist spel med förmögenheten X_0 och har efter n spelomgångar en förmögenhet som uppgår till X_n. Tolkningarna är desamma för sub- och supermartingaler, men i dessa fall handlar det om ett gynnsamt respektive ett missgynnsamt spel.

I martingalteorisammanhang är det särskilt lämpligt att tänka på σ-algebror som "informationsmängder". Vi tänker oss att $\mathcal{F}_n$ innehåller den information vi har vid tidpunkten n. Om vi nu tänker i termer av spelteori infinner sig snart frågan om huruvida det är möjligt att med hjälp av en finurlig spelstrategi göra ett missgynnsamt spel till ett gynnsamt. En möjlighet man har att påverka skeendet ligger i att ändra *insatsen* i de enskilda spelomgångarna. Insatsen i spelomgång nr. n kan då naturligtvis tänkas bero på vad som skett tidigare, men rimligen inte vad som ska ske i framtiden. Med andra ord kan man säga att insatsen i spelomgång n kan vara slumpmässig men att den måste vara given av informationen i σ-algebran $\mathcal{F}_{n-1}$.

DEFINITION. En stokastisk process $\{C_n\}_{n=1}^{\infty}$ sägs vara *förutsebar* (engelska: *previsible*) (m.a.p. $\{\mathcal{F}_n\}$) om det för alla n gäller att C_n är $\mathcal{F}_{n-1}$-mätbar.

En förutsebar stokastisk process kan vi alltså tolka som följden av insatser i ett spel. Låt nu $\{C_n\}$ vara en förutsebar process och $\{X_n\}$ vara en annan stokastisk process. Definiera processen $\{Y_n\}$ genom att sätta

$$Y_n = \sum_{k=1}^{n} C_k(X_k - X_{k-1}).$$

Vi tolkar Y_n som den totala vinsten efter n spelomgångar med spelstrategin $\{C_k\}$. Följande sats säger att *det finns ingen vinnande spelstrategi.*

SATS 10.1 *(a) Om $\{X_n\}$ är en supermartingal och det finns en kontant $K < \infty$ sådan att $0 \leq C_n(\omega) \leq K$ för alla n och $\omega \in \Omega$ gäller att även $\{Y_n\}$ är en supermartingal. Motsvarande gäller om $\{X_n\}$ är en submartingal eller en martingal.*

(b) Om $\{X_n\}$ är en martingal och det finns en konstant $K < \infty$ sådan att $|C_n(\omega)| \leq K$ för alla n och ω gäller att $\{Y_n\}$ är en martingal.

(c) *Påståendena i (a) och (b) är sanna även i fallet då kravet om begränsning hos $\{C_n\}$ ersätts med kravet att $X_n \in L^2$ och $C_n \in L^2$ för alla n.*

Bevis.

(a) Vi har att $\mathbf{E}[Y_{n+1} - Y_n | \mathcal{F}_n] = \mathbf{E}[C_{n+1}(X_{n+1} - X_n)|\mathcal{F}_n] = C_{n+1}\mathbf{E}[X_{n+1} - X_n|\mathcal{F}_n] \leq 0$, där den andra olkheten följer av Sats 9.2(j) och olikheten följer av supermartingalegenskapen. De övriga fallen är analoga.

(b) Kopiera beviset i (a).

(c) Återigen kopierar vi beviset i (a) med den skillnaden att vi hänvisar till "Hölders olikhet-villkoret" i Sats 9.2(j).

□

Det gäller alltså att ingen rimlig satsningsstrategi kan göra ett missgynnsamt spel gynnsamt.

Man kan fråga sig om man istället för att använda en smart satsningsstrategi skulle kunna använda sig av en smart *stoppregel* för att vända det missgynnsamma spelet till ett gynnsamt spel. Man skulle kunna tänka sig att använda en strategi av typen "Jag sluter spela när jag har 1000 kronor eller när pengarna är slut" eller "jag slutar när jag har vunnit min första spelomgång" etc. I analysen av konsekvenserna av att använda sådana stoppregler är det, precis som när det gäller satsningsstrategier, rimligt att begränsa sig till endast sådana stoppregler som bygger på nuet och det förflutna och inte på framtiden.

Definition. En ickenegativ diskret stokastisk variabel T (som även kan anta värdet ∞) kallas för en *stopptid* m.a.p. filtreringen $\{\mathcal{F}_n\}$ om det för varje $n = 0, 1, 2, \ldots$ gäller att

$$\{\omega \in \Omega : T(\omega) \leq n\} \in \mathcal{F}_n.$$

De stoppregler vi ser som "tillåtna" är alltså sådana som säger att man ska sluta spela vid en tidpunkt T där T är någon stopptid. (Observera att definitionen av en stopptid formellt endast kräver en filtrering och inte ens att det överhuvudtaget existerar en adapterad stokastisk process. Ofta är det ju dock så att $\{\mathcal{F}_n\}$ ges av att $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ för någon process $\{X_n\}$ och då kommer villkoret för att T ska vara en stopptid automatiskt att formuleras i termer av X_n:na.)

Proposition 10.2 *Det gäller att T är en stopptid om och endast om $\{T = n\} \in \mathcal{F}_n$ för alla n.*

Bevis. Å ena sidan gäller att $\{T=n\}=\{T\leq n\}\setminus\{T\leq n-1\}\in\mathcal{F}_n\cup\mathcal{F}_{n-1}=\mathcal{F}_n$ och å andra sidan har vi att $\{T\leq n\}=\cup_{k=1}^n\{T=k\}\in\cup_{k=1}^n\mathcal{F}_k=\mathcal{F}_n$. □

Exempel. Antag att $\{X_n\}$ är en $\{\mathcal{F}_n\}$-adapterad process och låt $T=\min\{n: X_n\geq 1000 \text{ eller } X_n\leq 0\}$. (Tänk på detta som att vi startar med förmögenheten X_0 och spelar tills vi har 1000 kronor eller gjort slut på alla pengar.) Är T en stopptid? Ja, ty $\{T\leq n\}=\cup_{k=1}^n\{X_k\notin(0,1000)\}\in\mathcal{F}_n$.

Exempel. Låt $T=\max\{n\leq 100: X_n\geq 1000\}$, dvs vi vill sluta sista gången före den 101:a omgången som vår förmögenhet är minst 1000 kronor. Då är T inte en stopptid om inte processen $\{X_n\}$ är trivial.

Svaret på frågan om det är möjligt att göra ett missgynnsamt spel gynnsamt genom att designa en lämplig stopptid ligger delvis redan i Sats 10.1; att använda en stopptid är ju inget annat än att använda en satsningsstrategi där man satsar beloppen 0 eller 1 beroende på om man har slutat spela eller inte. Låt oss genomföra detta resonemang ordentligt: Antag att $\{X_n\}$ är en supermartingal och låt T vara en godtyckligt vald stopptid. Definiera nu de stokastiska variablerna $C_n^{(T)}$, $n=1,2,\ldots$ genom att sätta $C_n^{(T)}=I_{\{n\leq T\}}=I_{\{T>n-1\}}$. Då blir $\{C_n^{(T)}\}$ en förutsebar process (ty $\{T>n-1\}=\{T\leq n-1\}^c\in\mathcal{F}_{n-1}$) och om vi som förut definierar processen $\{Y_n\}$ genom att sätta

$$Y_n=\sum_{k=1}^n C_k^{(T)}(X_k-X_{k-1})$$

följer det av Sats 10.1(a) att $\{Y_n\}$ är en supermartingal. Men det gäller ju att $Y_n=X_{T\wedge n}-X_0$ och det följer att $\{X_{T\wedge n}\}$ är en supermartingal. Vi har visat:

Sats 10.3 *Om $\{X_n\}$ är en supermartingal och T är en stopptid gäller att $\{X_{T\wedge n}\}$ också är en supermartingal. Speciellt gäller att*

$$\mathbf{E}[X_{T\wedge n}]\leq\mathbf{E}[X_0].$$

Motsvarande gäller för submartingaler och martingaler.

I alla *praktiska* situationer gäller att T måste vara begränsad av något N så att för $n\geq N$ gäller att $X_{T\wedge n}=X_T$ och det följer av Sats 10.3 att $\mathbf{E}[X_T]\leq\mathbf{E}[X_0]$. I matematikens värld är vi dock inte hindrade av praktiska invändningar och det går lätt att konstruera stopptider för vilka den sistnämnda slutsatsen är falsk. Ett standardexempel får man om man låter $\{X_n\}$ vara en enkel symmetrisk slumpvandring på $\mathbf{Z}$ som startar i origo. Då är $\{X_n\}$ en martingal och därmed också en supermartingal och den stokastiska variabeln $T=\min\{n: X_n=1\}$ är en stopptid sådan att $T<\infty$ n.s. Enligt resultaten ovan gället för alla n att $\mathbf{E}[X_{T\wedge n}]=0$ men det är ändå uppenbarligen så att $\mathbf{E}[X_T]=1>0$. I ljuset av detta infinner sig förstås frågan om precis vad man måste kräva av T och $\{X_n\}$ för att det ska vara sant att $\mathbf{E}[X_T]\leq\mathbf{E}[X_0]$. Svaret ges av följande viktiga sats:

Sats 10.4 (Doobs sats om optionella stopp) *(a) Låt $\{X_n, \mathcal{F}_n\}$ vara en supermartingal och låt T vara en stopptid. Det gäller att X_T är en integrerbar stokastisk variabel och att $\mathbf{E}[X_T] \leq \mathbf{E}[X_0]$ i var och en av följande situationer:*

(i) Om T är begränsad, dvs $\exists N < \infty : T(\omega) \leq N$ för alla $\omega \in \Omega$.

(ii) Om $T < \infty$ n.s. och $\{X_n\}$ är likformigt begränsad, dvs $\exists K < \infty : |X_n(\omega)| \leq K$ för alla n och alla $\omega \in \Omega$.

(iii) Om $\mathbf{E}[T] < \infty$ och $\{X_n\}$ har **begränsade ökningar**, *dvs $\exists K < \infty : |X_n(\omega) - X_{n-1}(\omega)| \leq k$ för alla n och alla $\omega \in \Omega$.*

(b) Motsvarigheten till resultaten i (a) gäller även för submartingaler och martingaler.

(c) Om $\{C_n\}$ är en likformigt begränsad förutsebar process och $\{M_n\}$ är en martingal med begränsade ökningar gäller för processen $\{Y_n\}$ definierad av:

$$Y_n = \sum_{k=1}^{n} C_k(M_k - M_{k-1})$$

och varje stopptid T sådan att $\mathbf{E}[T] < \infty$ att $\mathbf{E}[Y_T] = 0$.

(d) Om $\{X_n\}$ är en ickenegativ supermartingal och $T < \infty$ n.s. gäller att $\mathbf{E}[X_T] \leq \mathbf{E}[X_0]$.

Innan vi gör beviset är det värt att notera ett par saker. I slumpvandringsexemplet ovan gäller (naturligtvis) att (i), (ii) och (iii) spricker. För (i) och (ii) är detta uppenbart och vad (iii) beträffar spricker det på att det i detta exempel gäller att $\mathbf{E}[T] = \infty$.

Observera också att del (d) säger att inte bara måste man ha ett oändligt långt liv för att kunna vända ett missgynnsamt spel till ett gynnsamt, man måste också ha en outtömlig förmögenhet när man startar eller möjlighet att låna hur mycket pengar som helst. I del (d) kan man också låta T anta värdet ∞ med positiv sannolikhet om man samtidigt inför konventionen $X_\infty = 0$. (Detta är åtminstone i spelteorisammanhang naturligt eftersom $T = \infty$ ju betyder att man aldrig får ut sin eventuella vinst; även om man har ett oändligt långt liv får man ju aldrig uppleva "tidpunkten ∞".)

Bevis. Vi börjar med att observera att (b) följer omedelbart av (a) eftersom det gäller att om $\{X_n\}$ är en submartingal så är $\{-X_n\}$ en supermartingal och eftersom resultatet för martingaler följer genom att slå ihop resultaten för sub- och supermartingaler. Dessutom gäller att del (c) är en direkt kombination av Sats 10.1 och (a) där man då tillämpar villkor (iii). Således följer både (a), (b) och (c) om vi kan visa (a).

Att villkor (i) är tillräckligt observerade vi redan före satsen så antag att villkor (ii) gäller. Enligt Sats 10.3 gäller att $\{X_{T\wedge n}\}$ är en supermartingal och således att $\mathbf{E}[X_{T\wedge n} - X_0] \leq 0$. Eftersom $T < \infty$ n.s. gäller att $X_{T\wedge n} \to X_T$ n.s. och eftersom $|X_{T\wedge n} - X_0| \leq 2K$ kan vi tillämpa dominerad-konvergens-satsen och få att

$$\mathbf{E}[X_T - X_0] = \lim_{n\to\infty} \mathbf{E}[X_{T\wedge n} - X_0] \leq 0$$

som önskat. Under villkor (iii) använder vi dominerad-konvergens-satsen på precis samma sätt. Att detta fungerar följer dels av att $\mathbf{E}[T] < \infty \Rightarrow T < \infty$ n.s. så att $X_{T\wedge n} \to X_T$ n.s., dels av att $|X_{T\wedge n} - X_0| \leq \sum_{k=1}^{T\wedge n} |X_k - X_{k-1}| \leq KT$ och $\mathbf{E}[KT] = K\mathbf{E}[T] < \infty$.

Nu återstår att visa del (d). Eftersom $T < \infty$ n.s. har vi att

$$\mathbf{E}[X_T] = \mathbf{E}[\liminf_n X_{T\wedge n}] \leq \liminf_n \mathbf{E}[X_{T\wedge n}] \leq \mathbf{E}[X_0]$$

där vi använder Fatous lemma till den första olikheten och Sats 10.3 till den andra. □

Vi avslutar detta kapitel med en intressant observation som ligger lite vid sidan om övriga resultat i kapitlet, men som naturligt hör hemma i ett introducerande kapitel om martingalteori som detta.

SATS 10.5 (DOOB-DEKOMPOSITION AV EN ADAPTERAD PROCESS) *Låt $\{X_n\}$ vara en stokastisk process som är adapterad till filtreringen $\{\mathcal{F}_n\}$. Då finns det en martingal $\{M_n\}$ och en stokastisk process $\{A_n\}$ som är förutsebar m.a.p. $\{\mathcal{F}_n\}$ sådana att*

$$X_n = X_0 + M_n + A_n.$$

Det gäller vidare att $M_0 = A_0 = 0$ och att om $\{M'_n\}$ och $\{A'_n\}$ är två andra sådana processer så gäller att $P(\forall n : M_n = M'_n, A_n = A'_n) = 1$. Dessutom gäller det att $\{X_n\}$ är en submartingal/supermartingal om och endast om $\{A_n\}$ är växande/avtagande.

Bevis. Låt oss först konstatera att om satsen ska kunna vara sann måste det gälla för alla n att

$$\mathbf{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] = \mathbf{E}[M_n - M_{n-1}|\mathcal{F}_{n-1}] + \mathbf{E}[A_n - A_{n-1}|\mathcal{F}_{n-1}]$$

$$= A_n - A_{n-1}$$

nästan säkert vilket medför att

$$A_n = \sum_{k=1}^n (A_k - A_{k-1}) = \sum_{k=1}^n \mathbf{E}[X_k - X_{k-1}|\mathcal{F}_{k-1}]$$

n.s. och därmed måste vi också ha att

$$M_n = \sum_{k=1}^n (X_k - \mathbf{E}[X_k|\mathcal{F}_{k-1}]).$$

Detta betyder att unikhetsdelen av satsen är klar och genom att verifiera att de givna uttrycken gör $\{A_n\}$ förutsebar och $\{M_n\}$ till en martingal följer även existensdelen. Att göra detta är en enkel övning i räkning med betingade väntevärden som överlåtes till läsaren.

Slutligen obeserverar vi att om $\{X_n\}$ är en submartingal gäller att $\mathbf{E}[X_n - X_{n-1}] \geq 0$, dvs att $A_n - A_{n-1} \geq 0$ n.s. Det omvända påståendet är uppenbart och supermartingalfallet är analogt. $\square$

11 Martingalkonvergenssatsen

Låt $\{X_n, \mathcal{F}_n\}$ vara en supermartingal. Antag att vi försöker göra detta "missgynnsamma spel" till ett gynnsamt genom att använda följande strategi. Vi inför två "gränser" a och b sådana att $a < b$ och går med i spelet direkt efter den första tidpunkt n sådan att $X_n \leq a$ och deltar sedan t.o.m. den första tidpunkt m därefter sådan att $X_m \geq b$. Efter detta väntar vi tills processen återigen understiger a varvid vi på nytt går in i spelet och spelar fram till nästa tidpunkt då processen antar ett värde som är minst b. Detta förfarande upprepas i all oändlighet. (Denna strategi är typiskt vad man kan tänkas använda sig av när man handlar med aktier; köp när kursen är låg och sälj när kursen är hög. Förhoppningsvis är dock börskursen en submartingal snarare än en supermartingal.)

Formellt kan denna strategi definieras i termer av en förutsebar process $\{C_n\}$ som ges induktivt av att $C_1 = I_{\{X_0 \leq a\}}$ och av att $C_n = I_{\{C_{n-1}=1\}} I_{\{X_{n-1}<b\}} + I_{\{C_{n-1}=0\}} I_{\{X_{n-1} \leq a\}}$. Att $\{C_n\}$ är förutsebar följer av ett enkelt induktionsargument. Vår sammanlagda vinst vid tidpunkten n ges av

$$Y_n = \sum_{k=1}^{n} C_k (X_k - X_{k-1}).$$

Enligt Sats 10.1 gäller att även $\{Y_n\}$ är en supermartingal. (Dvs inte heller denna strategi lurar systemet, men det visste vi ju som sagt redan.) Speciellt gäller för alla N att $\mathbf{E}[Y_N] \leq \mathbf{E}[Y_0]$. Definiera nu den stokastiska variabeln $U_N^{(a,b)}$ som antalet *uppkorsningar* som processen $\{X_n\}$ har gjort av intervallet (a, b) t.o.m. tidpunkten N, dvs antalet tillfällen som vi vid tiden N har hunnit att börja *och avsluta* en spelsekvens. Formellt har vi $U_N^{(a,b)} = \max\{k \in \mathbf{Z}^+ : \exists s_1, t_1, \ldots, s_k, t_k : 0 \leq s_1 < t_1 < s_2 < \ldots < t_k \leq N, X_{s_i} \leq a, X_{t_i} \geq b\}$. Det kan alltså vara så att vi vid tidpunkten N befinner oss i en oavslutad spelsekvens och denna räknas alltså inte med i antalet uppkorsningar. Uppenbarligen gäller att $Y_N \geq (b-a)U_N^{(a,b)} - (X_N - a)^-$ där den sista termen är en övre gräns för den förlust vi eventuellt kan ha gjort om vi vid tidpunkten N befinner oss i en oavslutad spelsekvens. Eftersom $\mathbf{E}[Y_N] \leq 0$ gäller att

$$\mathbf{E}[U_N^{(a,b)}] \leq \frac{\mathbf{E}[(X_N - a)^-]}{b - a}.$$

Denna olikhet är känd under namnet *Doob's uppkorsningslemma.*

Antag nu att supermartingalen $\{X_n\}$ är *begränsad i* L^1, dvs $\sup_n \mathbf{E}[|X_n|] < \infty$. Då gäller speciellt att $\sup_n \mathbf{E}[X_n^-] = K < \infty$. Låt nu $U_\infty^{(a,b)} = \lim_{n\to\infty} U_N^{(a,b)}$ vara antalet uppkorsningar av (a,b) av hela processen $\{X_n\}$. Enligt monoton-konvergens-satsen gäller att $\mathbf{E}[U_\infty^{(a,b)}] = \lim_{n\to\infty} \mathbf{E}[U_N^{(a,b)}]$ och eftersom

$$\mathbf{E}[U_N^{(a,b)}] \leq \frac{\mathbf{E}[(X_n - a)^-]}{b-a} \leq \frac{\mathbf{E}[X_n^-] + |a|}{b-a} \leq \frac{K+a}{b-a} < \infty$$

medför detta att $\mathbf{E}[U_\infty^{(a,b)}] < \infty$. Speciellt gäller att $P(U_\infty^{(a,b)} = \infty) = 0$. Detta faktum har intressanta konsekvenser ty

$$\begin{aligned}
&P(\{\omega \in \Omega : \{X_n(\omega)\} \text{ konvergerar ej}\}) \\
&= P(\bigcup_{a,b\in\mathbf{Q}} \{\omega \in \Omega : \liminf_n X_n(\omega) \leq a < b \leq \limsup_n X_n(\omega)\}) \\
&= P(\bigcup_{a,b\in\mathbf{Q}:a<b} \{\omega \in \Omega : U_\infty^{(a,b)}(\omega) = \infty\}) \leq \sum_{a,b\in\mathbf{Q}:a<b} P(U_\infty^{(a,b)} = \infty) = 0.
\end{aligned}$$

Det gäller alltså att $\{X_n\}$ konvergerar n.s. Definiera nu X_∞ genom att sätta $X_\infty(\omega) = \limsup_n X_n(\omega)$. Då har vi att $X_n \to X_\infty$ n.s. Dessutom gäller enligt Fatous lemma att $\mathbf{E}[|X_\infty|] \leq \liminf_n \mathbf{E}[|X_n|] \leq \sup_n \mathbf{E}[|X_n|] < \infty$ och således att $X_\infty \in \mathbf{R}$ n.s. Genom att justera X_∞ på en nollmängd kan vi tillse att $X_\infty(\omega) \in \mathbf{R}$ för alla $\omega \in \Omega$. Vi har visat:

SATS 11.1 (MARTINGALKONVERGENSSATSEN) *Låt $\{X_n, \mathcal{F}_n\}$ vara en supermartingal som är begränsad i L^1. Då existerar en reellvärd integrerbar stokastisk variabel X_∞ sådan att $X_n \to X_\infty$ n.s. då $n \to \infty$.*

Observera att man lätt kan konstatera att satsen gäller även för submartingaler, ty om $\{X_n\}$ är en submartingal som är begränsad i L^1 är $\{-X_n\}$ en supermartingal begränsad i L^1.

Observera också att eftersom $|X_n| = X_n^+ + X_n^- = X_n + 2X_n^-$ och supermartingalegenskapen implicerar att $\mathbf{E}[X_n] \leq \mathbf{E}[X_0]$ har vi att $\mathbf{E}[|X_n|] \leq \mathbf{E}[X_0] + 2\mathbf{E}[X_n^-]$. Därför gäller att en supermartingal $\{X_n\}$ är begränsad i L^1 om och endast om $\sup_n \mathbf{E}[X_n^-] < \infty$. På motsvarande sätt gäller att en submartingal $\{X_n\}$ är begränsad i L^1 om och endast om $\sup_n \mathbf{E}[X_n^+] < \infty$. Därför följer direkt av martingalkonvergenssatsen:

SATS 11.2 *Om $\{X_n, \mathcal{F}_n\}$ är en ickenegativ supermartingal eller en ickepositiv submartingal gäller att det finns en reellvärd stokastisk variabel X_∞ sådan att $X_n \to X_\infty$ n.s.*

Varning: Det är *inte* säkert att $X_n \to X_\infty$ i L^1, dvs att $\mathbf{E}[|X_n - X_\infty|] \to 0$. (Domineringsvillkoret i dominerad-konvergens-satsen är inte automatiskt satisfierat.)

12 Martingaler begränsade i L^2 och stora talens lag

Antag att $\{M_n, \mathcal{F}_n\}$ är en martingal. Vi säger att $\{M_n\}$ är begränsad i L^p, $p \geq 1$ om det gäller att $\sup_n \mathbf{E}[|M_n|^p] < \infty$. Detta är en direkt generalisering av definitionen av begränsning i L^1 som vi använde i förra avsnittet.

Låt nu $\{M_n, \mathcal{F}_n\}$ vara en martingal sådan att $M_n \in L^2$ för alla n. Per definition gäller att $\mathbf{E}[M_{n+1}|\mathcal{F}_n] = M_n$ n.s. för alla n och enligt sats 9.1 gäller att $\mathbf{E}[M_{n+1}|\mathcal{F}_n] = M_n$ är den ortogonala projektionen av M_{n+1} på $L^2(\mathcal{F}_n)$. Detta betyder att för alla $Y \in L^2(\mathcal{F}_n)$ gäller att $\langle M_{n+1} - M_n, Y\rangle = 0$. Om vi nu tar $k \leq n$ gäller att $M_k - M_{k-1} \in L^2(\mathcal{F}_n)$ och det följer att $\mathbf{E}[(M_{n+1} - M_n)(M_k - M_{k-1})] = 0$. Eftersom det uppenbarligen gäller att $M_n = M_0 + \sum_{k=1}^n (M_k - M_{k-1})$ följer det att $\mathbf{E}[M_n^2] = \mathbf{E}[M_0^2] + \sum_{k=1}^n \mathbf{E}[(M_k - M_{k-1})^2]$. Som en omedelbar konsekvens följer del (a) av följande sats:

SATS 12.1 *(a) Martingalen $\{M_n, \mathcal{F}_n\}$ är begränsad i L^2 om och endast om $\sum_{k=1}^\infty \mathbf{E}[(M_k - M_{k-1})^2] < \infty$.*

(b) Om martingalen $\{M_n, \mathcal{F}_n\}$ är begränsad i L^2 existerar det en stokastisk variabel $M_\infty \in L^2$ sådan att $M_n \to M_\infty$ n.s. och i L^2.

Bevis. Del (a) är som sagt redan klar. Tack vare monotonicitetn hos L^p-normer gäller att $\{M_n\}$ är begränsad i L^1 så det följer av martingalkonvergenssatsen att det finns en stokastisk variabel M_∞ sådan att $M_n \to M_\infty$ n.s. För att visa konvergensen i L^2, fixera n och observera att enligt Fatous lemma gäller att

$$\mathbf{E}[(M_\infty - M_n)^2] \leq \liminf_r \mathbf{E}[(M_{n+r} - M_n)^2]$$

$$= \liminf_r \sum_{k=n+1}^{n+r} \mathbf{E}[(M_k - M_{k-1})^2] = \sum_{k=n+1}^{\infty} \mathbf{E}[(M_k - M_{k-1})^2].$$

Låt nu $n \to \infty$. Då gäller att högerledet konvergerar mot 0 enligt (a) och saken är klar. $\square$

Sats 12.1 var steg 1 på vägen mot Kolmogorovs tre-seriers-sats och stora talens lag. Steg 2 är följande resultat om konvergens av summor av oberoende stokastiska variabler.

SATS 12.2 *Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler med väntevärde 0 och $Var(X_k) = \sigma_k^2 < \infty$.*

(a) Om $\sum_{k=1}^\infty \sigma_k < \infty$ gäller att $\sum_{k=1}^\infty X_k$ är n.s. konvergent.

(b) Om det finns en konstant $K < \infty$ sådan att $|X_k| \leq K$ för alla k och $\sum_{k=1}^\infty X_k$ är n.s. konvergent gäller att $\sum_{k=1}^\infty \sigma_k < \infty$.

Bevis.

(a) Sätt $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $M_0 = 0$ och $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ och $M_n = \sum_{k=1}^n X_k$ för $n = 1, 2, \ldots$. Då är $\{X_n, \mathcal{F}_n\}$ en martingal där $M_n \in L^2$ så det följer av Sats 12.1(a) att $\{M_n\}$ är begränsad i L^2 om och endast om $\sum_{n=1}^\infty \mathbf{E}[(M_n - M_{n-1})^2] < \infty$, vilket är sant ty $\mathbf{E}[(M_n - M_{n-1})^2] = \mathbf{E}[X_n^2] = \sigma_n^2$ och $\sum_{n=1}^\infty \sigma_n^2 < \infty$ enligt förutsättning.

(b) Med samma definition som ovan av $\{M_n, \mathcal{F}_n\}$ gäller att

$$\sigma_k^2 = \mathbf{E}[X_k^2] = \mathbf{E}[(M_k - M_{k-1})^2]$$

$$= \mathbf{E}[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}] = \mathbf{E}[M_k^2 | \mathcal{F}_{k-1}] - M_{k-1}^2$$

nästan säkert, där den tredje likheten följer av det faktum att $M_k - M_{k-1} = X_k$ är oberoende av $\mathcal{F}_{k-1}$ och där den fjärde likheten följer av grundläggande regler för räkning med betingade väntevärden. Genom att flytta över σ_k^2 till andra sidan av ekvationen får vi alltså att $\mathbf{E}[M_k^2 | \mathcal{F}_{k-1}] = M_{k-1}^2 + \sigma_k^2$ n.s. Detta betyder att om vi sätter $N_n = M_n^2 - \sum_{k=1}^n \sigma_k^2$, $n = 0, 1, 2, \ldots$ så bildar $\{N_n, \mathcal{F}_n\}$ en martingal.

Låt nu c vara en positiv konstant och definiera en stopptid T genom att sätta $T = \min\{r : |M_r| \geq c\}$. Eftersom det enligt förutsättning gäller att $\{M_r\}$ konvergerar nästan säkert kan c väljas så stort att $P(T = \infty) \geq 1/2$ och vi antar att c är vald på det sättet.

Enligt Sats 10.3 gäller att följden $\{N_{T \wedge n}\}$ är en martingal. Speciellt gäller att $0 = \mathbf{E}[N_0] = \mathbf{E}[N_{T \wedge n}] = \mathbf{E}[M_{T \wedge n}^2] - \mathbf{E}[\sum_{k=1}^{T \wedge n} \sigma_k^2]$. Om nu $T > n$ gäller att $|M_{T \wedge n}| = |M_n| \leq c \leq c + K$ och om $T \leq n$ gäller att $|M_{T \wedge n}| = |M_T| \leq |M_{T-1}| + |X_T| \leq c + K$. Således gäller att $\mathbf{E}[M_{T \wedge n}^2] \leq (c + K)^2$ och därför att $\mathbf{E}[\sum_{k=1}^{T \wedge n} \sigma_k^2] \leq (c + K)^2$. Speciellt gäller att $\mathbf{E}[\sum_{k=1}^{T \wedge n} \sigma_k^2; T = \infty] = P(T = \infty) \sum_{k=1}^n \sigma_k^2 \leq (c + K)^2$ och således att $\sum_{k=1}^n \sigma_k^2 \leq (c + K)^2 / P(T = \infty) \leq 2(c + K)^2$. Eftersom denna övre begränsning är oberoende av n följer det att $\sum_{k=1}^\infty \sigma_k^2 < \infty$ som önskat.

□

Exempel. Antag att $Z_1, Z_2, \ldots$ är oberoende stokastiska variabler sådana att $P(Z_k = \pm 1) = 1/2$ och att $a_1, a_2, \ldots$ är en följd av reella tal. Eftersom $Var(a_n Z_n) = a_n^2$ följer det av Sats 12.2 att $\sum_{n=1}^\infty a_n Z_n$ är konvergent n.s. om och endast om $\sum_{n=1}^\infty a_n^2 < \infty$.

Sats 12.2 har den inskränkningen att den förutsätter att $\mathbf{E}[X_k] = 0$. Vi kommer att behöva en mer generell version av del (b) av satsen.

SATS 12.3 *Antag att $X_1, X_2, \ldots$ är en följd av oberoende stokastiska variabler sådana att det existerar en konstant $K < \infty$ sådan att $|X_k| \leq K$ för alla k och antag att $\sum_{k=1}^\infty X_k$ är n.s. konvergent. Då gäller att $\sum_{k=1}^\infty \mathbf{E}[X_k]$ är konvergent och att $\sum_{k=1}^\infty Var(X_k) < \infty$.*

Bevis. Börja med att definiera ett nytt sannolikhetsrum $(\Omega', \mathcal{F}', P')$ och låt $Y_1, Y_2, \ldots$ vara en följd av stokastiska variabler definierade på Ω' sådana att Y_k:na är oberoende och Y_k har samma fördelning som X_k för alla k. (Det går bra att göra detta. Repetera i kapitel 4.4 eller 8.7 om du känner dig osäker.)

Låt nu $(\Omega^*, \mathcal{F}^*, P^*) = (\Omega \times \Omega', \mathcal{F} \times \mathcal{F}', P \times P')$ och definiera $X_k^*(\omega, \omega') = X_k(\omega)$, $Y_k^*(\omega, \omega') = Y_k(\omega')$. Då blir automatiskt följderna $\{X_k^*\}$ och $\{Y_k^*\}$ sinsemellan oberoende och har bägge samma fördelning som $\{X_k\}$. Låt nu $Z_k^* = X_n^* - Y_k^*$. Då blir $\{Z_k^*\}$ en följd av oberoende stokastiska variabler sådana att $|Z_k^*| \leq 2K$, $\mathbf{E}[Z_k^*] = 0$ och $Var(Z_k^*) = 2Var(X_n)$. Eftersom det enligt förutsättning gäller att $\sum_{k=1}^\infty X_k^*$ och $\sum_{k=1}^\infty Y_k^*$ är n.s. konvergenta gäller också att $\sum_{k=1}^\infty Z_k^*$ är n.s. konvergent. Sammantaget gäller att $\{Z_k^*\}$ uppfyller förutsättningarna i Sats 12.2(b) och vi får att $\sum_{k=1}^\infty Var(Z_k^*) = 2\sum_{k=1}^\infty Var(X_k) < \infty$. Men då gäller ju enligt Sats 12.2(a) att $\sum_{k=1}^\infty (X_k - \mathbf{E}[X_k])$ konvergerar n.s. vilket medför att $\sum_{k=1}^\infty \mathbf{E}[X_k]$ är n.s. konvergent. □

Sats 12.4 (Kolmogorovs tre-seriers-sats) *Låt $\{X_n\}$ vara en följd av oberoende stokastiska variabler. Det gäller att $\sum_{n=1}^\infty X_n$ är n.s. konvergent om och endast om det gäller för något (alla) $K \in (0, \infty)$ att*

(i) $\sum_{n=1}^\infty P(|X_n| > K) < \infty$,

(ii) $\sum_{n=1}^\infty \mathbf{E}[X_n^K]$ *är konvergent,*

(iii) $\sum_{n=1}^\infty Var(X_n^K) < \infty$,

där $X_n^K = X_n I_{\{|X_n| \leq K\}}$.

Bevis. Antag först att $\sum_{n=1}^\infty X_n$ är n.s. konvergent och tag ett godtyckligt $K \in (0, \infty)$. Eftersom $X_n(\omega) \to 0$ för nästan alla ω gäller för nästan alla ω att det finns ett index n_0 sådant att $n \geq n_0 \Rightarrow |X_n(\omega)| \leq K$. Med andra ord gäller för nästan alla ω att $\omega \notin \limsup_n \{|X_n| > K\}$, dvs att $P(\limsup_n \{|X_n| > K\}) = 0$. Enligt Borel-Cantelli II medför detta att $\sum_{n=1}^\infty P(|X_n| > K) < \infty$ och därmed är (i) klar. Men eftersom det n.s. gäller att $X_n = X_n^K$ för alla utom ändligt många n följer det också att $\sum_{n=1}^\infty X_n^K$ är n.s. konvergent och därför följer (ii) och (iii) av Sats 12.3.

Antag nu att (i), (ii) och (iii) är satisfierade för något K. Enligt (i) och Borel-Cantelli I gäller att $X_n = X_n^K$ för alla utom ändligt många n n.s. Därför räcker det att visa att $\sum_{n=1}^\infty X_n^K$ är n.s. konvergent. Tack vare (ii) följer detta om vi kan visa att $\sum_{n=1}^\infty (X_n^K - \mathbf{E}[X_n^K])$ är n.s. konvergent. Detta följer i sin tur av Sats 12.2(a) som är tillämpbar tack vare (iii). □

Låt oss nu attackera stora talens lag. För att klara det behöver vi veta ett par saker till. Vi börjar med två analytiska lemmor.

LEMMA 12.5 (CÉSÀROS LEMMA) *Antag att* $\{b_n\}$ *är en följd av positiva reella tal sådana att* $b_n \uparrow \infty$ *och att* $\{v_n\}$ *är en annan följd av reella tal sådan att* $v_n \to v_\infty$*. Då gäller att*

$$\frac{1}{b_n}\sum_{k=1}^{n}(b_k - b_{k-1})v_k \to v_\infty$$

då $n \to \infty$.

Notera att som specialfall av Césàros lemma följer att $\frac{1}{n}\sum_{k=1}^{n} v_k \to v_\infty$.

Bevis. Fixera ett godtyckligt $\epsilon > 0$ och välj N så stort att $n \geq N \Rightarrow v_n \geq v_\infty - \epsilon$. Det gäller för $n \geq N$ att

$$\frac{1}{b_n}\sum_{k=1}^{n}(b_k - b_{k-1})v_k \geq \frac{1}{b_n}\sum_{k=1}^{N}(b_k - b_{k-1})v_k + \frac{b_n - b_N}{b_n}(v_\infty - \epsilon)$$

och eftersom den första termen konvergerar mot 0 och den andra konvergerar mot $v_\infty - \epsilon$ då $n \to \infty$ följer det att $\liminf_n \frac{1}{b_n}\sum_{k=1}^{n}(b_k - b_{k-1})v_k \geq v_\infty$. På ett analogt sätt följer det att $\limsup_n \frac{1}{b_n}\sum_{k=1}^{n}(b_k - b_{k-1})v_k \leq v_\infty$. □

LEMMA 12.6 (KRONECKERS LEMMA) *Antag att* $\{b_n\}$ *är en följd av positiva reella tal sådana att* $b_n \uparrow \infty$ *och att* $\{x_n\}$ *är en annan följd av reella tal sådan att* $\sum_{n=1}^{\infty}(x_n/b_n)$ *är konvergent. Sätt* $s_n = \sum_{k=1}^{n} x_k$*. Det gäller då att* $s_n/b_n \to 0$.

Återigen är det värt att notera ett specialfall: Om $\{x_n\}$ är en följd sådan att $\sum_{n=1}^{\infty}(x_n/n)$ är konvergent gäller att $s_n/n \to 0$.

Bevis. Sätt $v_n = \sum_{k=1}^{n}(x_n/b_n)$. Enligt förutsättning är $\{v_n\}$ konvergent så låt oss kalla gränsen för v_∞. Vi kan nu tillämpa Césàros lemma på $\{v_n\}$. Eftersom

$$s_n = \sum_{k=1}^{n} b_k \frac{x_k}{b_k} = \sum_{k=1}^{n} b_k(v_k - v_{k-1}) = b_n v_n - \sum_{k=1}^{n}(b_k - b_{k-1})v_{k-1}$$

följer det då att

$$\frac{s_n}{b_n} = v_n - \frac{1}{b_n}\sum_{k=1}^{n}(b_k - b_{k-1})v_{k-1} \to v_\infty - v_\infty = 0.$$

□

En direkt följd av Kroneckers lemma är följande sats som är stora talens lag under variansrestriktion.

SATS 12.7 *Låt $\{W_n\}$ vara en följd av stokastiska variabler sådana att $\mathbf{E}[W_n] = 0$ och $\sum_{n=1}^{\infty}(Var(W_n)/n^2) < \infty$. Då gäller att $(\sum_{k=1}^{n} W_k)/n \to 0$ n.s.*

Bevis. Enligt Kroneckers lemma räcker det att visa att $\sum_{n=1}^{\infty}(W_n/n)$ är n.s. konvergent. Detta följer i sin tur av Sats 12.2(a) som är tillämpbar tack vare villkoret på varianserna. $\square$

LEMMA 12.8 (KOLMOGOROVS TRUNKERINGSLEMMA) *Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade stokastiska variabler i L^1 och sätt $\mu = \mathbf{E}[X_1]$. Låt $Y_1, Y_2, \ldots$ ges av $Y_n = X_n I_{\{|X_n|<n\}}$, $n = 1, 2, \ldots$. Då gäller:*

(i) $\mathbf{E}[Y_n] \to \mu$,

(ii) $P(\limsup_n \{X_n \neq Y_n\}) = 0$,

(iii) $\sum_{n=1}^{\infty}(Var(Y_n)/n^2) < \infty$.

Bevis. Kom ihåg att för ett reellt tal x skriver man $[x]$ för heltalsdelen av x, dvs $[x] = \min\{k \in \mathbf{Z} : k \leq x\}$.

Låt genomgående X vara en stokastisk variabel med samma fördelning som X_n:na. Om vi sätter $Z_n = X I_{\{|X|<n\}}$ har Z_n samma fördelning som Y_n och $Z_n \to X$ n.s. Eftersom dessutom $|Z_n| \leq |X|$ och $X \in L^1$ är dominerad-konvergens-satsen tillämpbar och vi får att $\mathbf{E}[Y_n] = \mathbf{E}[Z_n] \to \mathbf{E}[X] = \mu$. Därmed är (i) klar så låt oss nu gripa oss an (ii). Om vi kan visa att $\sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty$ följer (ii) av Borel-Cantelli I. Men

$$\sum_{n=1}^{\infty} P(X_n \neq Y_n) = \sum_{n=1}^{\infty} P(|X_n| \geq n) = \sum_{n=1}^{\infty} P(|X| \geq n) = \sum_{n=1}^{\infty} \mathbf{E}[I_{\{|X|\geq n\}}]$$

$$= \mathbf{E}[\sum_{n=1}^{\infty} I_{\{|X|\geq n\}}] = \mathbf{E}[[|X|]] \leq \mathbf{E}[|X|] < \infty.$$

För att klara (iii) krävs det lite räknande med tungan rätt i mun. Observera först att

$$\sum_{n=1}^{\infty} \frac{Var(Y_n)}{n^2} \leq \sum_{n=1}^{\infty} \frac{\mathbf{E}[Y_n^2]}{n^2} = \sum_{n=1}^{\infty} \frac{\mathbf{E}[X^2 I_{\{|X|<n\}}]}{n^2}$$

$$= \mathbf{E}[X^2 \sum_{n=1}^{\infty} \frac{I_{\{|X|<n\}}}{n^2}] = \mathbf{E}[X^2 \sum_{n=[|X|]+1}^{\infty} \frac{1}{n^2}].$$

Nu är det ju så att $1/n^2 \leq 2/(n(n+1)) = 2(1/n - 1/(n+1))$ (med likhet då $n = 1$) och således att

$$\sum_{n=[|X|]+1}^{\infty} \frac{1}{n^2} \leq 2 \sum_{n=[|X|]+1}^{\infty} (\frac{1}{n} - \frac{1}{n+1}) = \frac{2}{[|X|]+1} \leq \frac{2}{|X|}.$$

Därför får vi att

$$\sum_{n=1}^{\infty} \frac{Var(Y_n)}{n^2} \leq 2\mathbf{E}[|X|] < \infty$$

och saken är klar. □

Vi har till sist kommit till stora talens lag.

Sats 12.9 (Stora talens lag) *Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade stokastiska variabler i L^1 med $\mathbf{E}[X_1] = \mu$ och sätt $S_n = \sum_{k=1}^{n} X_k$. Det gäller att $S_n/n \to \mu$ n.s.*

Bevis. Definiera $\{Y_n\}$ som i Kolmogorovs trunkeringslemma ovan. Tack vare (iii) i Kolmogorovs trunkeringslemma kan vi använda stora talens lag under variansrestriktion ovan till att konstatera att $\frac{1}{n}\sum_{k=1}^{n}(Y_k - \mathbf{E}[Y_k]) \to 0$ n.s. Enligt (i) i trunkeringslemmat gäller att $\mathbf{E}[Y_n] \to \mu$ så enligt Césàros lemma har vi att $\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}[Y_k] \to \mu$ och således har vi att $\frac{1}{n}\sum_{k=1}^{n} Y_k \to \mu$ n.s. Slutligen gäller enligt (ii) i trunkeringslemmat n.s. att $X_n = Y_n$ för alla utom ändligt många n varför det även gäller att $\frac{1}{n}\sum_{k=1}^{n} X_k \to \mu$ n.s. som önskat. □

13 Likformig integrerbarhet

Vi ska nu införa ett begrepp som hjälper oss att tala om när martingaler inte bara är konvergenta nästan säkert utan också i L^1 och när gränsvariabeln fungerer som ett extra "sista" element i martingalen. Vi kommer också att kunna formulera en starkare version av dominerad-konvergens-satsen som är den starkast möjliga generella versionen. Begreppet i fråga kallas för "likformig integrerbarhet".

Antag att $X \in L^1(\Omega, \mathcal{F}, P)$. Man kan fråga sig om det är sant att det för varje $\epsilon > 0$ finns ett $\delta > 0$ sådant att så fort $P(F) < \delta$ gäller att $\int_F |X| dP < \epsilon$. Svaret är ja och för att bevisa det antar vi motsatsen, dvs att det finns ett $\epsilon > 0$ sådant att för varje $n = 1, 2, \ldots$ finns en mängd F_n med $P(F_n) < 2^{-n}$ och $\int_{F_n} |X| dP \geq \epsilon$. Enligt Borel-Cantelli I gäller att $P(\limsup_n F_n) = 0$. Men eftersom $I_{\limsup_n F_n} = \limsup_n I_{F_n}$ gäller att

$$\int_{\limsup_n F_n} |X| dP = \int_{\Omega} \limsup_n (|X| I_{F_n}) dP \geq \limsup_n \int_{F_n} |X| dP \geq \epsilon$$

där den första olikheten följer från den omvända versionen av Fatous lemma, och vi har fått den önskade motsägelsen.

Ett specialfall av detta faktum erhåller man om man noterar att enligt Markovs olikhet gäller att $KP(|X| > K) \leq \mathbf{E}[|X|]$ så att om man väljer K tillräckligt stort gäller att $P(|X| > K) < \delta$ varför det följer att $\mathbf{E}[|X|; |X| > K] < \epsilon$ för tillräckligt stora värden på K.

Om vi istället för att betrakta en enstaka stokastisk variabel X betraktar en hel familj $\{X_i\}$ av stokastiska variabler kallar vi denna familj likformigt integrerbar om man kan välja K så stort att $\mathbf{E}[|X_i|; |X_i| > K] < \epsilon$ för alla i *samtidigt*.

DEFINITION. En familj $\{X_i\}_{i\in I}$ av stokastiska variabler sägs vara *likformigt integrerbar* (förkortas LI) om det för varje $\epsilon > 0$ finns ett $K_\epsilon < \infty$ sådant att

$$\mathbf{E}[|X_i|; |X_i| > K_\epsilon] < \epsilon$$

för alla $i \in I$ (dvs likformigt i i).

Låt oss först göra klart för oss hur likformig integrerbarhet förhåller sig till begränsning i L^p för olika p. För det första gäller att om $\{X_i\}$ är LI har vi att

$$\mathbf{E}[|X_i|] = \mathbf{E}[|X_i|; |X_i| \leq K_1] + \mathbf{E}[|X_i|; |X_i| > K_1] \leq K_1 + 1$$

dvs $\{X_i\}$ är begränsad i L^1. Å andra sidan gäller inte omvändningen. Ett motexempel får man om man sätter $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}, Leb)$ och $X_n = nI_{[0,1/n)}$, $n = 1, 2, \ldots$. Då är $\mathbf{E}[X_n] = 1$ för alla n så $\{X_n\}$ är begränsad i L^1, men för vilket K man än väljer har vi att $\mathbf{E}[|X_n|; |X_n| > K] = 1$ så fort $n > K$, dvs $\{X_n\}$ är inte LI.

Däremot blir situationen annorlunda om vi betraktar relationen mellan LI och begränsning i L^p för $p > 1$. Antag först att $\{X_i\}$ är begränsad i L^p. Utsagan $|X_i| > K$ är ekvivalent med $|X_i|^{1-p} < K^{1-p}$ vilket i sin tur är samma sak som att $|X_i| < |X_i|^p K^{1-p}$. Därför gäller att $\mathbf{E}[|X_i|; |X_i| > K] \leq K^{1-p}\mathbf{E}[|X_i|^p; |X_i| > K] \leq K^{1-p}\mathbf{E}[|X_i|^p] \leq K^{1-p}A$, där $A = \sup_i \mathbf{E}[|X_i|^p] < \infty$ enligt förutsättning och eftersom $1-p < 0$ gäller att $K^{1-p} \to 0$ då $K \uparrow \infty$, dvs $\{X_i\}$ är LI. Å andra sidan är inte en likformigt integrerbar familj automatiskt begränsad i L^p. För att konstruera ett motexempel, sätt $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}, Leb)$ och tag ett $c \in (1/p, 1)$ och sätt $X_n = n^c I_{[0,1/n)}$. Då gäller att $\mathbf{E}[|X_n|; |X_n| > K]$ är 0 då $n \leq K^{1/c}$ och n^{c-1} då $n > K^{1/c}$. Välj nu bara K så stort att $K^{(c-1)/c} < \epsilon$ och det följer att $\{X_n\}$ är LI. Dock gäller det att $\mathbf{E}[|X_n|^p] = n^{pc-1} \uparrow \infty$, dvs $\{X_n\}$ är inte begränsad i L^p.

Sammanfattningsvis gäller att likformig integrerbarhet är starkare än begränsning i L^1, men svagare än begränsning i L^p för $p > 1$.

Ett tillräckligt villkor för likformig integrerbarhet av en annan typ är följande:

Proposition 13.1 *Om $\{X_i\}_{i \in I}$ är en familj av stokastiska variabler sådana att det finns en stokastisk variabel Y med $\mathbf{E}[Y] < \infty$ sådan att $|X_i| \le Y$ för alla i gäller att $\{X_i\}$ är likformigt integrerbar.*

Bevis. Tag $\epsilon > 0$. Det gäller att $\mathbf{E}[|X_i|; |X_i| > K] \le \mathbf{E}[Y; |X_i| > K] \le \mathbf{E}[Y; Y > K]$ och eftersom $\mathbf{E}[Y; Y > K] < \epsilon$ för tillräckligt stort K enligt tidigare observationer är saken därmed klar. $\square$

Följande sats är intressant eftersom den bland annat visar att om $\{\mathcal{F}_n\}$ är en filtrering och $X \in L^1$ gäller att $\{\mathbf{E}[X|\mathcal{F}_n], \mathcal{F}_n\}$ är en likformigt integrerbar martingal. (Martingalegenskapen observerade vi redan i kapitel 10.)

Sats 13.2 *Låt $\mathbf{G}$ vara familjen av alla del-σ-algebror till $\mathcal{F}$ och låt $X \in L^1(\Omega, \mathcal{F}, P)$. Det gäller att familjen $\mathcal{C} = \{\mathbf{E}[X|\mathcal{G}] : \mathcal{G} \in \mathbf{G}\}$ är likformigt integrerbar.*

Bevis. Tag $\epsilon > 0$ och låt δ vara sådant att $P(F) < \delta \Rightarrow \mathbf{E}[|X|; F] < \epsilon$.

Vad vi ska visa är att man kan välja K så stort att $\mathbf{E}[|Y|; |Y| > K] < \epsilon$ för alla $Y \in \mathcal{C}$. Men enligt den betingade versionen av Jensens olikhet gäller för $Y = \mathbf{E}[X|\mathcal{G}]$ att $|Y| \le \mathbf{E}[|X||\mathcal{G}]$ n.s. så att $\mathbf{E}[|Y|; |Y| > K] \le \mathbf{E}[\mathbf{E}[|X||\mathcal{G}]; |Y| > K] = \mathbf{E}[|X|; |Y| > K]$ enligt definitionen av betingat väntevärde, ty $\{|Y| > K\} \in \mathcal{G}$. Detta betyder att om vi kan visa att K kan väljas så att $P(|Y| > K) < \delta$ för alla $Y \in \mathcal{C}$ så är vi klara. Men enligt Markovs olikhet gäller att $KP(|Y| > K) \le \mathbf{E}[|Y|]$ och eftersom $\mathbf{E}[|Y|] \le \mathbf{E}[\mathbf{E}[|X||\mathcal{G}]] = \mathbf{E}[|X|]$ gäller att $P(|Y| > K) \le \mathbf{E}[|X|]/K < \delta$ för alla $Y \in \mathcal{C}$ så fort K är stort nog. $\square$

Vi ska nu ägna resten av detta kapitel till att förbättra dominerad-konvergens-satsen. Som vi nu känner den säger den att om $X_n \to X$ n.s. och det finns en integrerbar stokastisk variabel Y sådan att $|X_n| \le Y$ för alla n, gäller att $\mathbf{E}[|X_n - X|] \to 0$, dvs $X_n \to X$ i L^1. Den optimala versionen som vi nu ska bevisa säger att om $X_n \xrightarrow{P} X$ och $\{X_n\}$ är LI gäller att $\mathbf{E}[|X_n - X|] \to 0$. (Kom ihåg definitionen av konvergens i sannolikhet; att $X_n \xrightarrow{P} X$ betyder att för alla $\epsilon > 0$ gäller att $P(|X_n - X| > \epsilon) \to 0$. Det är uppenbart att detta är svagare än att $X_n \to X$ n.s.)

Vi börjar med att bevisa följande lemma som du antagligen redan känner till.

Lemma 13.3 *Antag att $\{X_n\}$ är en följd av stokastiska variabler sådan att $X_n \xrightarrow{P} X$. Då finns det en delföljd $\{X_{n_k}\}_{k=1}^{\infty}$ sådan att $X_{n_k} \to X$ n.s. då $k \to \infty$.*

Bevis. Eftersom $X_n \xrightarrow{P} X$ kan vi för varje $k = 1, 2, \ldots$ välja ett index n_k sådant att $P(|X_{n_k} - X| > 1/k) \le 2^{-k}$. Det följer då från Borel-Cantelli I att $P(\limsup_k \{|X_{n_k} - X| > 1/k\}) = 0$, dvs för nästan alla ω gäller att $|X_{n_k}(\omega) - X(\omega)| \le 1/k$ för alla utom ändligt många k, dvs $X_{n_k}(\omega) \to X(\omega)$. $\square$

Följande lemma är en utvidgning av dominerad-konvergens-satsen för begränsade följder till konvergens i sannolikhet.

Lemma 13.4 *Antag att* $X_n \xrightarrow{P} X$ *och att det finns en konstant* $K < \infty$ *sådan att* $|X_n| \leq K$ *för alla* n*. Då gäller att* $\mathbf{E}[|X_n - X|] \to 0$.

Bevis. Låt $\{X_{n_k}\}$ vara vilken delföljd som helst till $\{X_n\}$. Till denna delföljd finns enligt lemmat ovan i sin tur en delföljd $\{X_{n_{k_m}}\}$ sådan att $X_{n_{k_m}} \to X$ n.s. Därför följer det av den "vanliga" formen av dominerad-konvergens-satsen att $\mathbf{E}[|X_{n_{k_m}} - X|] \to 0$. Vi har alltså visat att till varje delföljd finns en delföljd för vilken vi har konvergens och det följer av grundläggande kunskaper om konvergens att det då följer att $\mathbf{E}[|X_n - X|] \to 0$. $\square$

Nu griper vi oss an huvudresultatet.

Sats 13.5 *Antag att* $X, X_1, X_2, \ldots$ *är stokastiska variabler. Det gäller att* $\{X_n\}$ *är likformigt integrerbar och* $X_n \xrightarrow{P} X$ *om och endast om* $\mathbf{E}[|X_n - X|] \to 0$.

Bevis. Tanken är att betrakta X_n:na som begränsade och att använda den likformiga integrerbarheten till att visa att detta endast leder till försumbara fel.

Vi börjar med framåtimplikationen. Definiera för $K < \infty$ funktionen $\phi_K : \mathbf{R} \to [-K, K]$ genom att låta

$$\phi_K(x) = \begin{cases} -K & \text{då } x < -K \\ x & \text{då } x \in [-K, K] \\ K & \text{då } x > K. \end{cases}$$

Fixera $\epsilon > 0$ och fixera K så att $\mathbf{E}[|X_n - \phi_K(X_n)|] = \mathbf{E}[|X_n| - K; |X_n| > K] \leq \mathbf{E}[|X_n|; |X_n| > K] < \epsilon/3$ för alla n och $\mathbf{E}[|X - \phi_K(X)|] < \epsilon/3$. Att ett sådant K existerar följer av den likformiga integrerbarheten.

Eftersom $|\phi_K(x) - \phi_K(y)| \leq |x - y|$ för alla x och y gäller att $\phi_K(X_n) \xrightarrow{P} \phi_K(X)$. Enligt Lemma 13.4 gäller alltså att $\mathbf{E}[|\phi_K(X_n) - \phi_K(X)|] \to 0$. Vi kan alltså välja N så att $n \geq N \Rightarrow \mathbf{E}[|\phi_K(X_n) - \phi_K(X)|] < \epsilon/3$. Nu följer det av triangelolikheten att $\mathbf{E}[|X_n - X|] < \epsilon$ då $n \geq N$ och framåtimplikationen är klar.

Antag nu att $X_n \to X$ i L^1, dvs att $\mathbf{E}[|X_n - X|] \to 0$. Fixera $\epsilon > 0$. Enligt Markovs olikhet gäller att $\epsilon P(|X_n - X| \geq \epsilon) \leq \mathbf{E}[|X_n - X|] \to 0$ och därmed att $P(|X_n - X| \geq \epsilon) \to 0$, dvs att $X_n \xrightarrow{P} X$. Därmed återstår det att visa den likformiga integrerbarheten.

Välj $\delta > 0$ så att $P(F) < \delta \Rightarrow \mathbf{E}[|X|; F] < \epsilon/2$. Eftersom $\mathbf{E}[X_n] \to \mathbf{E}[X]$ gäller att följden $\{X_n\}$ är begränsad i L^1 varför vi kan välja K så stort att $P(|X_n| > K) \leq$

$\mathbf{E}[|X_n|]/K \leq \sup_n \mathbf{E}[|X_n|]/K < \delta$ för alla n. Sammantaget är alltså K vald så att $\mathbf{E}[|X|;|X_n| > K] < \epsilon/2$ för alla n. Dessutom gäller att $\mathbf{E}[|X_n|;|X_n| > K] \leq \mathbf{E}[|X|;|X_n| > K] + \mathbf{E}[|X_n - X|;|X_n| > K] \leq \epsilon/2 + \mathbf{E}[|X_n - X|]$. Eftersom $\mathbf{E}[|X_n - X|] \to 0$ finns det ett N så att $n \geq N \Rightarrow \mathbf{E}[|X_n - X|] < \epsilon/2$ och det följer att $\mathbf{E}[|X_n|;|X_n| > K] < \epsilon$ för $n \geq N$. Nu vill vi ju noga taget att detta ska gälla för alla n, men detta ordnar vi nu genom att om det behövs justera upp K så att $\mathbf{E}[|X_n|;|X_n| > K] < \epsilon$ även för $n < N$. □

14 Likformigt integrerbara martingaler

14.1 Sista element och Lévys uppåtsats

Vi börjar detta kapitel med en sats som säger att till en likformigt integrerbar martingal finns det ett s.k. *sista element.*

SATS 14.1 *Antag att $\{M_n, \mathcal{F}_n\}$ är en LI martingal. Då finns det en stokastisk variabel $M_\infty \in L^1(\Omega, \mathcal{F}, P)$ sådan att $M_n \to M_\infty$ n.s. och i L^1. Dessutom gäller för alla n att $\mathbf{E}[M_\infty|\mathcal{F}_n] = M_n$ n.s.*

Det är det sista faktumet som föranleder oss att kalla M_∞ för att sista element, eftersom martingalegenskapen fortfarande gäller om vi betraktar följden $\{M_1, M_2, \ldots, M_\infty\}$.

Bevis. Eftersom $\{M_n\}$ är LI gäller att $\{M_n\}$ är begränsad i L^1 och det följer av martingalkonvergenssatsen att det existerar ett M_∞ sådant att $M_n \to M_\infty$ n.s. Att $M_n \to M_\infty$ i L^1 är en direkt konsekvens av den generella versionen av dominerad-konvergenssatsen i förra kapitlet, Sats 13.5. Enligt samma sats gäller dessutom för alla $F \in \mathcal{F}_n$ att

$$\int_F M_\infty dP = \lim_{r\to\infty} \int_F M_r dP = \int_F M_n dP$$

där den andra olikheten följer av martingalegenskapen. Enligt definitionen av betingat väntevärde gäller alltså att $\mathbf{E}[M_\infty|\mathcal{F}_n] = M_n$ n.s. □

En snabb blick på beviset övertygar oss om att satsen är sann även för likformigt integrerbara sub- och supermartingaler med motsvarande formulering.

SATS 14.2 (LÉVYS UPPÅTSATS) *Tag $X \in L^1(\Omega, \mathcal{F}, P)$ och låt $\{\mathcal{F}_n\}$ vara en filtrering. Sätt $\mathcal{F}_\infty = \sigma(\cup_{n=1}^\infty \mathcal{F}_n)$ och låt $M_n = \mathbf{E}[X|\mathcal{F}_n]$. Då är $\{M_n\}$ en LI martingal och $M_n \to \mathbf{E}[X|\mathcal{F}_\infty]$ n.s.*

Bevis. Att $\{M_n\}$ är en martingal har vi redan sett (kapitel 10) och att $\{M_n\}$ är LI följer av Sats 13.2. Därför följer det av Sats 14.1 att det finns ett M_∞ sådant att $M_n \to$

M_∞ n.s. och i L^1. Det återstår att visa att $M_\infty = \mathbf{E}[X|\mathcal{F}_\infty]$ n.s. Sätt $Y = \mathbf{E}[X|\mathcal{F}_\infty]$. För varje n har vi att $\mathbf{E}[Y|\mathcal{F}_n] = \mathbf{E}[\mathbf{E}[X|\mathcal{F}_\infty]|\mathcal{F}_n] = \mathbf{E}[X|\mathcal{F}_n] = M_n = \mathbf{E}[M_\infty|\mathcal{F}_n]$ n.s. där den andra likheten följer av tornegenskapen och den sista likheten följer av sista-element-egenskapen hos M_∞. Således har vi att $\int_F Y dP = \int_F M_\infty dP$ för alla $F \in \mathcal{F}_n$ för alla n, dvs för alla $F \in \cup_n \mathcal{F}_n$. Om vi nu antar att X är ickenegativ följer det av unicitetssatsen att $\int_F Y dP = \int_F M_\infty dP$ för alla $F \in \mathcal{F}_\infty$, dvs att $Y = M_\infty$ n.s. som önskat. I det generella fallet delar vi upp X i X^+ och X^- och tillämpar vad vi just visat på dessa: Vi har att $\mathbf{E}[X^+|\mathcal{F}_n] \to \mathbf{E}[X^+|\mathcal{F}_\infty]$ n.s. och i L^1 och motsvarande för X^-. Kombinera nu dessa fakta med linjariteten hos betingade väntevärden och saken är klar. $\square$

Lévys uppåtsats ger oss möjlighet att ge ett kort och elegant bevis av Kolmogorovs 0-1-lag: Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler och låt som vanligt $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$ och $\mathcal{T} = \cap_n \mathcal{T}_n$. Sätt $X = I_F$ för en händelse $F \in \mathcal{T}$. Enligt Lévys uppåtsats och det faktum att X är $\mathcal{F}_\infty$-mätbar gäller att $\mathbf{E}[X|\mathcal{F}_n] \to \mathbf{E}[X|\mathcal{F}_\infty] = X$ n.s. Men eftersom $\mathcal{T}$ är oberoende av $\mathcal{F}_n$ för alla n gäller ju att $\mathbf{E}[X|\mathcal{F}_n] = \mathbf{E}[X] = P(F)$ n.s. för alla n. Sammantaget har vi att $I_F = P(F)$ n.s. och därmed att $P(F) \in \{0, 1\}$, dvs vi har bevisat Kolmogorovs 0-1-lag.

14.2 Bakåtmartingaler och stora talens lag

Antag att $\{\mathcal{F}_n\}$ är en *avtagande* följd av del-σ-algebror till $\mathcal{F}$ (dvs $\{\mathcal{F}_n\}$ är "motsatsen" till en filtrering). Antag vidare att $\{X_n\}$ är en stokastisk process. Vi säger att $\{X_n, \mathcal{F}_n\}$ är en *bakåtmartingal* om det gäller för alla n att X_n är $\mathcal{F}_n$-mätbar och $\mathbf{E}[X_n|\mathcal{F}_{n+1}] = X_{n+1}$ n.s.

Antag nu att $\{X_n\}$ är en bakåtmartingal begränsad i L^1. Om vi fixerar N och ser på följden $\{X_N, X_{N-1}, \ldots, X_0, X_0, X_0, \ldots\}$ med tillhörande σ-algebror så är denna följd en martingal begränsad i L^1. Vi påminner oss nu uppkorsningsargumentet från beviset av martingalkonvergenssatsen; om vi låter $U_N^{(a,b)}$ vara antalet uppkorsningar som denna martingal gör av intervallet (a, b) gäller att $\mathbf{E}[U_N^{(a,b)}] \leq A$, där $A = (\sup_n \mathbf{E}[|X_n|] + a^-)/(b-a)$, dvs A beror inte av n. Eftersom $U_N^{(a,b)}$ också är antalet *nedkorsningar* av intervallet (a, b) av den ursprungliga delföljden $\{X_0, \ldots, X_N\}$ får vi om vi sätter $U_\infty^{(a,b)} = \lim_{N\to\infty} U_N^{(a,b)}$ att denna stokastiska variabel är antalet nedkorsningar av intervallet (a, b) av hela bakåtmartingalen $\{X_n\}$. Eftersom $\mathbf{E}[U_\infty^{(a,b)}] \leq A < \infty$ enligt monoton-konvergens-satsen gäller att $U_\infty^{(a,b)} < \infty$ n.s. och det följer av samma argument som ledde fram till martingalkonvergenssatsen att $X_n \to X_\infty$ n.s. för någon integrerbar stokastisk variabel X_∞. Om man dessutom antar att $\{X_n\}$ är LI tillkommer slutsatsen att $X_n \to X_\infty$ även i L^1 enligt Sats 13.5. Vi har visat:

SATS 14.3 (KONVERGENSSATSEN FÖR BAKÅTMARTINGALER) *Låt $\{X_n, \mathcal{F}_n\}$ vara en bakåtmartingal som är begränsad i L^1. Då gäller att det finns en integrerbar stokastisk variabel X_∞ sådan att $X_n \to X_\infty$ n.s. Om $\{X_n\}$ är likformigt integrerbar gäller dessutom att $X_n \to X_\infty$ i L^1.*

Ett specialfall av konvergenssatsen är *Lévys nedåtsats*: Tag $X \in L^1(\Omega, \mathcal{F}, P)$ och antag att $\{\mathcal{F}_n\}$ är en avtagande följd av σ-algebror. Sätt $M_n = \mathbf{E}[X|\mathcal{F}_n]$. Då gäller enligt nu välkända argument att $\{X_n\}$ är en LI bakåtmartingal. Således finns det ett M_∞ sådant att $M_n \to M_\infty$ n.s. och i L^1. Låt nu $\mathcal{G}_\infty = \cap_n \mathcal{F}_n$. Eftersom M_∞ är mätbar m.a.p. $\mathcal{F}_n$ för alla n gäller att M_∞ är $\mathcal{G}_\infty$-mätbar. Dessutom gäller för alla $G \in \mathcal{G}_\infty$ att $\int_G \mathbf{E}[X|\mathcal{G}_\infty]dP = \int_G X dP = \int_G M_r dP$ för alla r, där den sista likheten följer av det faktum att $G \in \mathcal{F}_r$ och definitionen av betingat väntevärde. Låt nu $r \to \infty$ och använd Sats 13.5 till att inse att $\int_G M_r dP \to \int_G M_\infty dP$. Det följer att $M_\infty = \mathbf{E}[X|\mathcal{G}_\infty]$ n.s. dvs vi har visat att $\mathbf{E}[X|\mathcal{F}_n] \to \mathbf{E}[X|\mathcal{G}_\infty]$ n.s. och i L^1.

Nytt bevis av stora talens lag: Vi ska nu demonstrera hur kraftfull martingalteorin kan vara genom att göra ett kort och klart bevis av stora talens lag. Om man samtidigt drar sig till minnes hur mycket möda som det klassiska trunkeringsbeviset krävde framgår styrkan i den nya tekniken med all önskvärd tydlighet.

Antag att $X_1, X_2, \ldots$ är oberoende stokastiska variabler med väntevärde μ. Sätt som vanligt $S_n = \sum_{k=1}^n X_k$. Enligt Sats 9.3 gäller att $S_n/n = \mathbf{E}[X_1|\mathcal{F}_n]$ där $\mathcal{F}_n = \sigma(S_n, S_{n+1}, \ldots)$. Enligt Lévys nedåtsats gäller alltså att $S_n \to L$ n.s. och i L^1, där L är en integrerbar stokastisk variabel. (I själva verket vet vi att $L = \mathbf{E}[X_1| \cap_n \mathcal{F}_n]$ men den informationen har vi ingen nytta av här.) För att vara specifika kan vi låta $L(\omega) = \limsup_n (\sum_{k=1}^n X_k(\omega))/n$. Men det gäller ju då också för godtyckligt m att $L(\omega) = \limsup_n (\sum_{k=m+1}^{m+n} X_k(\omega))/n$, dvs att L är mätbar m.a.p. $\sigma(X_{m+1}, X_{m+2}, \ldots)$ för alla m, dvs L är en svansfunktion. Enligt Kolmogorovs 0-1-lag gäller alltså att $L = c$ n.s. för någon konstant c. Men eftersom $S_n/n \to c$ i L^1 har vi att $\mu = \mathbf{E}[S_n/n] \to c$, dvs $c = \mu$. Saken är klar.

14.3 Doobs L^p-olikhet

Antag att $\{X_n\}$ är en ickenegativ submartingal som är begränsad i L^p för något $p > 1$. (Kom ihåg att detta är ett starkare antagande än att $\{X_n\}$ är LI.) Vi ska nu ägna en del utrymme åt att bevisa Doob's L^p-olikhet som säger att då är $\{X_n\}$ faktiskt dominerad av en stokastisk variabel Z^* som är sådan att $Z^* \in L^p$. Enligt martingalkonvergenssatsen gäller att $\{X_n\}$ konvergerar n.s. mot en stokastisk variabel X_∞. En direkt konsekvens av Doobs L^p-olikhet och dominerad-konvergens-satsen blir att $X_n \to X_\infty$ i L^p, ty $|X_n - X_\infty|^p \leq 2^p |Z^*|^p$. Observera skillnaden mot fallet $p = 1$; vi vet ju att en martingal begränsad i L^1 inte alltid är konvergent i L^1.

Innan vi ger oss i kast med vår uppgift är det värt att notera att om $\{M_n, \mathcal{F}_n\}$ är en martingal och $c : \mathbf{R} \to \mathbf{R}$ är en konvex funktion sådan att $\mathbf{E}[|c(M_n)|] < \infty$ för alla n så är följden $\{c(M_n), \mathcal{F}_n\}$ en submartingal. Detta följer av den betingade versionen av Jensens olikhet ty $\mathbf{E}[c(M_{n+1})|\mathcal{F}_n] \geq c(\mathbf{E}[M_{n+1}|\mathcal{F}_n]) = c(M_n)$. Ett specialfall av denna observation är att om $\{M_n\}$ är en martingal begränsad i L^p så är även $\{|M_n|\}$ en martingal begränsad i L^p.

Följande sats påminner mycket om Markovs olikhet:

SATS 14.4 (DOOBS SUBMARTINGALOLIKHET) *Låt $\{Z_n\}$ vara en ickenegativ submartingal. Det gäller för alla $c > 0$ att*

$$cP(\sup_{k\leq n} Z_k \geq c) \leq \mathbf{E}[Z_n; \sup_{k\leq n} Z_k \geq c] \leq \mathbf{E}[Z_n].$$

Observera att om man tillämpar Markovs olikhet på $\sup_{k\leq n} Z_k$ så får man en olikhet som ser ut som den som ges av Doobs submartingalolikhet men där Z_n ersatts med $\sup_{k\leq n} Z_k$, dvs ett svagare resultat. Om man vill kan man säga att Doobs submartingalolikhet tillåter att vi i Markovs olikhet "låtsas att Z_n är störst av $Z_1, \ldots, Z_n$".

Bevis. Vi kan skriva händelsen $\{\sup_{k\leq n} Z_k \geq c\}$ som $F_0 \cup F_1 \cup \ldots \cup F_n$ där $F_0 = \{Z_0 \geq c\}$, $F_1 = \{Z_0 < c, Z_1 \geq c\}$, ..., $F_n = \{Z_0 < c, Z_1 < c, \ldots, Z_{n-1} < c, Z_n \geq c\}$. Då är $F_0, \ldots, F_n$ disjunkta händelser och i ord gäller att F_k är händelsen att Z_k är det första elementet i följden $Z_0, \ldots, Z_n$ som överstiger c. Eftersom $F_k \in \mathcal{F}_k$ följer det av submartingalegenskapen att $\mathbf{E}[Z_n; F_k] \geq \mathbf{E}[Z_k; F_k] \geq cP(F_k)$ så att eftersom Z_n är ickenegativ och F_k:na är disjunkta gäller att

$$\mathbf{E}[Z_n; \cup_{k=1}^n F_k] = \sum_{k=1}^n \mathbf{E}[Z_n; F_k] \geq c \sum_{k=1}^n P(F_k) = cP(\cup_{k=1}^n F_k).$$

Detta bevisar den första olikheten och den andra följer trivialt av ickenegativiteten hos Z_n. $\square$

Nästa resultat påminner om Chebyshevs olikhet:

SATS 14.5 (KOLMOGOROVS OLIKHET) *Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler med väntevärde 0 och $Var(X_k) = \sigma_k^2 < \infty$ och låt $S_n = \sum_{k=1}^n X_k$. Det gäller att*

$$c^2 P(\sup_{k\geq n} |S_k| \leq c) \leq Var(S_n) = \sum_{k=1}^n \sigma_k^2.$$

Bevis. Låt $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Då är $\{S_n, \mathcal{F}_n\}$ en martingal och således är $\{S_n^2, \mathcal{F}_n\}$ är en ickenegativ submartingal. Etersom $\{\sup_{k\leq n} |S_k| \geq c\} = \{\sup_{k\leq n} S_k^2 \geq c^2\}$ följer satsen genom att tillämpa Doobs submartingalolikhet på $\{S_n^2\}$. $\square$

Innan vi är redo att gripa oss an beviset av Doobs submartingalolikhet behöver vi följande Hölderliknande lemma:

LEMMA 14.6 *Låt X och Y vara två ickenegativa stokastiska variabler som satisfierar*

$$cP(X \geq c) \leq \mathbf{E}[Y; X \geq c]$$

för alla $c > 0$. För $p > 1$ och q sådant att $1/p + 1/q = 1$ gäller att $||X||_p \leq q||Y||_p$.

Observera att villkoret i lemmat är uppfyllt enligt Markovs olikhet om $X \leq Y$. Observera också att om $\{Z_n\}$ är en ickenegativ submartingal och vi sätter $X = \sup_{k \leq n} Z_k$ och $Y = Z_n$ så följer det av Doobs submartingalolikhet att villkoret är uppfyllt.

Bevis. Det följer direkt av antagandet att

$$\int_0^\infty pc^{p-1} P(X \geq c) dc \leq \int_0^\infty pc^{p-2} \mathbf{E}[Y; X \geq c] dc.$$

Genom att tillämpa Fubinis sats följer att vänsterledet är

$$\int_0^\infty \left(\int_\Omega I_{\{X \geq c\}}(\omega) P(d\omega)\right) pc^{p-1} dc = \int_\Omega \left(\int_0^{X(\omega)} pc^{p-1} dc\right) P(d\omega)$$

$$= \int_\Omega X(\omega)^p P(d\omega) = \mathbf{E}[X^p].$$

Högerledet å sin sida kan skrivas skrivas som

$$\int_0^\infty \left(\int_\Omega Y(\omega) I_{\{X(\omega) \geq c\}} P(d\omega)\right) pc^{p-2} dc = \int_\Omega Y(\omega) \left(\int_0^{X(\omega)} pc^{p-2} dc\right) P(d\omega)$$

$$= \int_\Omega \frac{p}{p-1} Y(\omega) X(\omega)^{p-1} P(d\omega) = q\mathbf{E}[X^{p-1} Y] \leq q||X^{p-1}||_q ||Y||_p,$$

där olikheten följer av Hölders olikhet. Vi har visat att

$$\mathbf{E}[X^p] \leq q||Y||_p ||X^{p-1}||_q. \tag{1}$$

Om $||Y||_p = \infty$ har vi inget att visa så antag att $||Y||_p < \infty$. Eftersom $||X^{p-1}||_q = \mathbf{E}[X^{(p-1)/p}]^{1/q} = \mathbf{E}[X^p]^{1/q}$ följer satsen av (1) genom division av bägge sidor med $\mathbf{E}[X^p]^{1/q}$ förutsatt att $\mathbf{E}[X^p] < \infty$ så att denna division är tillåten. Eftersom vi inte vet om det sistnämnda gäller observerar vi att om vi ersätter X med $X \wedge n$ för något n gäller alla förutsättningar fortfarande och dessutom att $\mathbf{E}[(X \wedge n)^p] < \infty$ och vi får att $||X \wedge n||_p \leq q||Y||_p$. Låt nu $n \to \infty$ och använd monoton-konvergens-satsen för att se att $||X||_p \leq q||Y||_p$ som önskat. $\Box$

Sats 14.7 (Doobs L^p-olikhet) *Låt $p > 1$ och tag q så att $1/p + 1/q = 1$. Låt $\{Z_n, \mathcal{F}_n\}$ vara en ickenegativ submartingal som är begränsad i L^p och sätt $Z^* = \sup_n Z_n$. Då gäller att $||Z^*||_p \leq q \sup_n ||Z_n||_p < \infty$. Speciellt gäller att $Z^* \in L^p$.*

Vidare gäller att det finns en stokastisk variabel $Z_\infty \in L^p$ sådan att $Z_\infty \in L^p$ och $Z_n \to Z_\infty$ n.s. och i L^p. Dessutom gäller att $||Z_n||_p \uparrow ||Z_\infty||_p$.

Bevis. Låt oss ta sista biten först. Att det finns en Z_∞ så att $Z_n \to Z_\infty$ n.s. följer av martingalkonvergenssatsen och att $Z_n \to Z_\infty$ i L^p argumenterade vi för redan i inledningen av detta avsnitt. Dessutom gäller att eftersom avbildningen $x \to x^p$ är konvex och eftersom Z_n är ickenegativ följer det av den betingade versionen av Jensens olikhet att $\{Z_n^p\}$ är en submartingal. Därför har vi att $\{\mathbf{E}[Z_n^p]\}$ är en växande följd. Eftersom dominerad-konvergens-satsen är tillämpbar med Z^* som dominerande element följer det att $\mathbf{E}[Z_n^p] \uparrow \mathbf{E}[Z_\infty^p]$, dvs att $||Z_n||_p \uparrow ||Z_\infty||_p$.

Låt oss nu koncentera oss på den första delen. Låt $Z_n^* = \sup_{k \leq n} Z_k$. Då gäller uppenbarligen att $Z_n^* \uparrow Z^*$ så enligt monoton-konvergens-satsen följer att $||Z^*||_p = \lim_{n\to\infty} ||Z_n^*||_p$. Nu är ju (som vi obseverade ovan) förutsättningarna i Lemma 14.6 satisfierade med $X = Z_n^*$ och $Y = Z_n$. Detta följer av Doobs submartingalolikhet. Vi får att $||Z_n^*||_p \leq q||Z_n||_p$ och genom att gå i gräns får vi $||Z^*||_p \leq \sup_n ||Z_n||_p$ som önskat. $\square$

En följd av Doobs submartingalolikhet är att om $\{M_n\}$ är en martingal som är begränsad i L^p gäller att $M_n \to M_\infty$ n.s. och i L^p. Detta är en av övningarna bak i kompendiet.

14.4 Ett martingalteoretiskt bevis av Radon-Nikodyms sats

Kom ihåg Radon-Nikodyms sats: Om γ och μ är två σ-ändliga mått på $(\Omega, \mathcal{F})$ och $\gamma << \mu$ finns det en funktion $X \in (m\mathcal{F})^+$ sådan att $\gamma(F) = \int_F X d\mu$ för alla $F \in \mathcal{F}$. Det är enkelt att inse att för att bevisa detta räcker det att klara fallet då γ är ett ändligt mått Q och μ är ett sannolikhetsmått P. (Största delen av denna insikt skaffade vi oss redan i kapitel 5.) Vi ska med hjälp av martingalteori konstruera ett bevis av detta i specialfallet då σ-algebran $\mathcal{F}$ är *uppräkneligt genererad*, dvs då det existerar en följd $\{F_n\}$ av händelser i $\mathcal{F}$ sådan att $\mathcal{F} = \sigma(F_n : n = 1, 2, \ldots)$. Det är värt att notera att $\mathcal{B}(\mathbf{R})$ är uppräkneligt genererad; låt $\{F_n\}$ vara en uppräkning av alla intervall med rationella ändpunkter. Däremot är $\mathcal{L}(\mathbf{R})$ inte uppräkneligt genererad.

Vi börjar med att observera att för ett godtyckligt $\epsilon > 0$ finns ett $\delta > 0$ sådant att $Q(F) < \epsilon$ så fort $P(F) < \delta$. Detta följer av ett motsägelseargument: Antag att det för varje $k = 1, 2, \ldots$ finns en mängd F_k med $P(F_k) < 2^{-k}$ sådan att $Q(F_k) \geq \epsilon$. Enligt Borel-Cantelli I gäller att $P(\limsup_k F_k) = 0$ och därmed att $Q(\limsup_k F_k) = 0$. Men enligt omvända Fatous lemma för mängder gäller att $Q(\limsup_k F_k) \geq \limsup_k Q(F_k) \geq \epsilon$, en motsägelse. (Det går bra att tillämpa omvända Fatous lemma här eftersom Q är ett ändligt mått.)

Låt nu för varje n σ-algebran $\mathcal{F}_n$ ges av $\sigma(F_1, \ldots, F_n)$. Eftersom $\mathcal{F}_n$ är en ändlig σ-algebra kan $\mathcal{F}_n$ också skrivas som $\sigma(A_{n1}, \ldots, A_{nr(n)})$ för någon partition $\{A_{n1}, \ldots, A_{nr(n)}\}$ av Ω. Definiera nu den stokastiska variabeln X_n genom att sätta:

$$X_n(\omega) = \sum_{i: A_{ni} > 0} \frac{Q(A_{ni})}{P(A_{ni})}.$$

Då gäller det uppenbarligen för alla $F \in \mathcal{F}_n$ att $\int_F X_n dP = Q(F)$. Eftersom $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$

följer det av detta att $\int_F X_{n+1}dP = Q(F) = \int_F X_n dP$ för alla $F \in \mathcal{F}_n$, dvs $\{X_n, \mathcal{F}_n\}$ är en martingal. Eftersom $\{X_n\}$ är ickenegativ följer det nu av martingalkonvergenssatsen att det finns en stokastisk variabel X_∞ sådan att $X_n \to X_\infty$ n.s.

Är $\{X_n\}$ LI? Svaret är ja: Fixera $\epsilon > 0$ och välj δ så att $P(F) < \delta \Rightarrow Q(F) < \epsilon$. Enligt Markovs olikhet gäller att $P(X_n > K) \leq \mathbf{E}[X_n]/K = Q(\Omega)/K < \delta$ likformigt i n om bara K väljs stort nog. Således gäller för sådant K att $Q(X_n > K) < \epsilon$ likformigt i n och eftersom $\mathbf{E}[X_n; X_n > K] = \int_{\{X_n > K\}} X_n dP = Q(X_n > K)$ (ty $\{X_n > K\} \in \mathcal{F}_n$) följer det att $\{X_n\}$ är LI. Detta medför att $X_n \to X_\infty$ i L^1. Speciellt gäller för alla $F \in \cup_n \mathcal{F}_n$ att $\int_F X_\infty dP = \lim_{n\to\infty} \int_F X_n dP = Q(F)$. Enligt unicitetssatsen gäller nu denna likhet för alla $F \in \sigma(\cup_n F_n) = \mathcal{F}$, dvs X_∞ är den sökta Radon-Nikodym-derivatan.

Det är möjligt att genom ett abstrakt matematiskt resonemang utvidga detta bevis till godtyckliga mätbara rum, men eftersom detta argument är rent ickeprobabilistiskt och vi redan har givit ett fullständigt bevis av Radon-Nikodyms sats hoppar vi över detta.

14.5 Produktmartingaler

Vi ska nu arbeta ett tag med följande modell: Antag att $X_1, X_2, \ldots$ är oberoende ickenegativa stokastiska variabler som alla har väntevärde 1. Sätt $M_n = \prod_{k=1}^n X_k$ och $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n = 1, 2, \ldots$. Eftersom $\mathbf{E}[M_{n+1}|\mathcal{F}_n] = \mathbf{E}[X_1 \ldots X_n X_{n+1}|\mathcal{F}_n] = X_1 \ldots X_n \mathbf{E}[X_{n+1}|\mathcal{F}_n] = X_1 \ldots X_n = M_n$ gäller det tydligen att $\{M_n, \mathcal{F}_n\}$ är en ickenegativ martingal och således följer det av martingalkonvergenssatsen att $M_n \to M_\infty$ n.s. för någon integrerbar stokastisk variabel M_∞. För $\{M_n\}$ och M_∞ gäller följande:

SATS 14.8 (KAKUTANIS SATS) *Låt $a_n = \mathbf{E}[X_n^{1/2}]$, $n = 1, 2, \ldots$. Följande fem påståenden är ekvivalenta:*

(i) $\mathbf{E}[M_\infty] = 1$,

(ii) $M_n \to M_\infty$ *i* L^1.

(iii) $\{M_n\}$ *är likformigt integrerbar,*

(iv) $\prod_{n=1}^\infty a_n > 0$,

(v) $\sum_{n=1}^\infty (1 - a_n) < \infty$.

Dessutom gäller om dessa påståenden ej är satisfierade att $M_\infty = 0$ n.s.

Bevis. Att (iv) och (v) är ekvivalenta är ett känt faktum från algebran, så låt oss koncentera oss på de övriga ekvivalenserna. Sätt

$$N_n = \frac{X_1^{1/2} \ldots X_n^{1/2}}{a_1 \ldots a_n}$$

för $n = 1, 2, \ldots$. Eftersom $\mathbf{E}[X_k^{1/2}/a_k] = 1$ gäller det av samma skäl som för $\{M_n\}$ att $\{N_n\}$ är en ickenegativ martingal och således konvergerar N_n n.s. mot någon integrerbar stokastisk variabel. Om det nu är så att (iv) inte är satisfierad, dvs att $\prod_{n=1}^\infty a_n = 0$, måste det alltså gälla att $\prod_{n=1}^\infty X_n^{1/2} = 0$ n.s. och därmed även att $\prod_{n=1}^\infty X_n = 0$ n.s., dvs att $M_\infty = 0$ n.s. Detta bevisar satsens avslutande del och även att (i) och (ii) spricker. Därmed spricker även (iii) eftersom (iii) implicerar (ii) och (i); en LI martingal konvergerar ju i L^1. Å andra sidan, om $\prod_{n=1}^\infty a_n > 0$, gäller att

$$\mathbf{E}[N_n^2] = \frac{1}{a_1^2 \ldots a_n^2} \leq \frac{1}{(\prod_{n=1}^\infty a_n)^2} < \infty,$$

där olikheten följer av att $a_n \leq 1$ för alla n, vilket i sin tur följer av att $a_n = ||X_n^{1/2}||_1 \leq ||X_n^{1/2}||_2 = \mathbf{E}[X_n]^{1/2} = 1$. Med andra ord gäller att $\{N_n\}$ är begränsad i L^2. Eftersom $a_n \leq 1$ gäller att $M_n \leq N_n^2$ och vi har att

$$\mathbf{E}[\sup_n M_n] \leq \mathbf{E}[\sup_n N_n^2] = ||\sup_n N_n||_2^2 \leq (2 \sup_n ||N_n||_2)^2$$

$$= 4 \sup_n ||N_n||_2^2 = 4 \sup_n \mathbf{E}[N_n^2] < \infty,$$

där den andra olikheten följer av Doobs submartingalolikhet. Vi har visat att $\sup_n M_n \in L^1$, dvs att $\{M_n\}$ är dominerad av en stokastisk variabel i L^1 och således är $\{M_n\}$ LI. Alltså gäller (iii) och därmed även (i) och (ii) och vi är klara. $\Box$

Vi avslutar kursen med att göra en statistikteoretisk tillämpning av Kakutanis sats. Antag att X är en kontinuerlig stokastisk variabel. Antag också att vi vet att X antingen har täthet f eller täthet g och vill via en observation av X uttala oss om vilket. I grundläggande statistikkurser lär man sig att använda LR-testet (LR för engelskans "likelihood ratio"), dvs om kvoten

$$\frac{g(X)}{f(X)}$$

blir hög tror man på g medan om den blir låg gissar man på f.

Låt oss för enkelhets skull anta att f och g är sådana att $f(x) > 0$ och $g(x) > 0$ för alla $x \in \mathbf{R}$. Låt P och Q vara fördelningarna som svarar mot f respektive g, dvs $P(B) = \int_B f(x)dx$ och $Q(B) = \int_B g(x)dx$, $B \in \mathcal{B}(\mathbf{R})$. För en godtycklig mängd $B \in \mathcal{B}$ har vi att $\int_B (g(x)/f(x))P(dx) = \int_B (g(x)/f(x))f(x)dx = \int_B g(x)dx = Q(B)$, dvs kvoten $g(x)/f(x)$ är ingenting annat än en version av Radon-Nikodym-derivatan $\frac{dQ}{dP}(x)$. På samma sätt gäller förstås att $f(x)/g(x) = \frac{dP}{dQ}(x)$. Detta inbjuder till en generalisering: Allmänt definierar vi LR mellan två sannolikhetsmått P och Q på samma mätbara rum som $\frac{dQ}{dP}$. En sådan definition kräver förstås att $Q << P$. Om P och Q är sådana att $Q << P$ och $P << Q$ säger vi att P och Q är *ekvivalenta* sannolikhetsmått. Vi kan i det läget även definiera $\frac{dP}{dQ}$. Man kan visa att om P och Q är ekvivalenta gäller att $P(\frac{dQ}{dP} > 0) = Q(\frac{dP}{dQ} > 0) = 1$ och att att $\frac{dP}{dQ} = (\frac{dQ}{dP})^{-1}$. Tag det gärna som en övning att göra detta.

LEMMA 14.9 *Antag att P och Q är två sannolikhetsmått på ett mätbart rum $(\Omega, \mathcal{F})$ sådana att $Q << P$. Sätt $X = \frac{dQ}{dP}$ och låt $\mathcal{G}$ vara en del-σ-algebra till $\mathcal{F}$. Låt $P_{\mathcal{G}}$ och $Q_{\mathcal{G}}$ vara restriktionerna av måtten P och Q till $\mathcal{G}$. Det gäller att $\frac{dQ_{\mathcal{G}}}{dP_{\mathcal{G}}} = \mathbf{E}_P[X|\mathcal{G}]$ n.s.*

Bevis. Det gäller att $\frac{dQ_{\mathcal{G}}}{dP_{\mathcal{G}}}$ och $\mathbf{E}_P[X|\mathcal{G}]$ bägge är $\mathcal{G}$-mätbara funktioner och för en godtycklig mängd $G \in \mathcal{G}$ gäller att

$$\int_G \mathbf{E}_P[X|\mathcal{G}]dP = \int_G X dP = Q(G) = \int_G \frac{dQ_{\mathcal{G}}}{dP_{\mathcal{G}}} dP.$$

□

En direkt konsekvens av lemmat ovan är att om $\{\mathcal{F}_n\}$ är en filtrering och vi låter $P_n = P_{\mathcal{F}_n}$ och $Q_n = Q_{\mathcal{F}_n}$ gäller att $\{\frac{dQ_n}{dP_n}\}$ är en LI P-martingal.

Låt oss nu se på en följd av LR-test, där vi i test nr. k testar tätheten f_k mot tätheten g_k. Detta betyder att test nr. k testar sannolikhetsmåtten P_k och Q_k, som är definierade på $(\Omega_k, \mathcal{F}_k) = (\mathbf{R}, \mathcal{B})$, mot varandra. Vi antar att för alla k gäller att f_k och g_k är strikt positiva tätheter, vilket medför att kvoterna $f_k(\omega)/g_k(\omega)$ är väldefinierade. Låt nu $(\Omega, \mathcal{F}) = (\prod_{k=1}^{\infty} \Omega_k, \prod_{k=1}^{\infty} \mathcal{F}_k) = (\mathbf{R}^{\infty}, \mathcal{B}^{\infty})$ och låt $\mathcal{F}_n = \sigma(\rho_1, \ldots, \rho_n)$, där $\rho_1, \rho_2, \ldots$ är koordinatavbildningarna. Låt vidare $P = \prod_{k=1}^{\infty} P_k$ med en analog definition av Q. Låt också $P_{\leq n}$ vara restriktionen av P till $\mathcal{F}_n$, dvs $P_{\leq n} = \prod_{k=1}^{n} P_k$, och definiera $Q_{\leq n}$ analogt. Definiera nu de stokastiska variablerna Y_k, $k = 1, 2, \ldots$ genom att sätta $Y_k(\omega_1, \omega_2, \ldots) = g_k(\omega_k)/f_k(\omega_k)$. Per konstruktion gäller att Y_k:na är oberoende och om vi sätter $M_n = Y_1 \ldots Y_n$ gäller att $M_n = \frac{dQ_{\leq n}}{dP_{\leq n}}$. Detta följer av att

$$\int_F Y_1 \ldots Y_n dP_{\leq n} = \int_F g_1(\omega_1) \ldots g_n(\omega_n) d\omega_1 \ldots d\omega_n = Q_{\leq n}(F)$$

enligt Fubinis sats för alla $F \in \mathcal{F}_n$ av typen $A_1 \times \ldots \times A_n$, $A_k \in \mathcal{F}_k$, och därmed för alla $F \in \mathcal{F}_n$ enligt unicitetssatsen.

Eftersom $\mathbf{E}_P[Y_k] = \int_{\mathbf{R}} (g_k(\omega_k)/f_k(\omega_k)) f_k(\omega_k) d\omega_k = \int_{\mathbf{R}} g_k(\omega_k) d\omega_k = 1$ så är M_n en P-martingal av den produkttyp som behandlas i Kakutanis sats. Antag nu att $Q << P$. (Detta är inte självklart. Vi vet att $P_{\leq n}$ och $Q_{\leq n}$ är ekvivalenta men vi vet inte vad som händer i gräns.) Då gäller enligt lemmat ovan att $\{M_n\}$ är LI.

Om vi å andra sidan antar att $\{M_n\}$ är LI gäller att $M_n \to M_{\infty}$ n.s. och i $L^1(P)$ och att $\mathbf{E}_P[M_{\infty}|\mathcal{F}_n]dP = M_n$ n.s. för alla n. Därför gäller att $\int_F M_{\infty} dP = Q(F)$ för alla $F \in \cup_n \mathcal{F}_n$ och därmed enligt unicitetssatsen för alla $F \in \mathcal{F}$. Med andra ord är M_{∞} en version av $\frac{dQ}{dP}$ och speciellt gäller att $Q << P$. Vi har visat att $Q << P$ är ekvivalent med att $\{M_n\}$ är LI.

Enligt Kakutanis sats gäller att $\{M_n\}$ är LI om och endast om $\prod_{k=1}^{\infty} \mathbf{E}[Y_k^{1/2}] = \prod_{k=1}^{\infty} \int_{\mathbf{R}} \sqrt{f_k(x) g_k(x)} dx > 0$. I specialfallet då $f_k = f$ och $g_k = g$ för alla k inträffar detta om och endast om $f = g$ n.ö. och då är $Q = P$. Vi ser att i detta specialfall gäller att Q är ekvivalent med P om och endast om $Q = P$ vilket också är ekvivalent med att $f = g$ n.ö. och att $\{M_n\}$ är LI och dessutom säger Kakutanis sats att om $Leb\{f \neq g\} > 0$ gäller att $M_n \to 0$ P-n.s. Detta betyder att om f och g verkligen är olika tätheter i den meningen att de representerar olika fördelningar och vi har ett stickprov om storlek n av den stokastiska variabeln X gäller att

$$\frac{g(X_1)}{f(X_1)} \frac{g(X_2)}{f(X_2)} \cdots \frac{g(X_n)}{f(X_n)} \to 0$$

n.s. då $n \to \infty$ om f är den korrekta tätheten och vice versa. Det är detta som man brukar referera till som konsistens hos LR-test; asymptotiskt kan vi med 100-procentig säkerhet avgöra vilken som är den korrekta tätheten.

Övningar

Samtliga övningar som återfinns nedan är sådana som inte redan finns i Williams. Övningarna är sorterade efter kapitel och att en övning står under kapitel n betyder att man behöver ha läst t.o.m. kapitel n för att kunna lösa uppgiften.

Kapitel 1 & 2

1. Visa att om Σ är en algebra sådan att $A_1, A_2, \ldots \in \Sigma$, $A_1 \subseteq A_2 \subseteq \ldots \Rightarrow \cup_{n=1}^{\infty} A_n \in \Sigma$ så är Σ en σ-algebra.

2. Antag att μ är ett mått på (S, Σ) och fixera $E \in \Sigma$. Definiera μ_E genom att sätta $\mu_E(A) = \mu(E \cap A)$, $A \in \Sigma$. Visa att μ_E är ett mått.

3. Ett mått μ på (S, Σ) sägs vara *semiändligt* om det gäller att $\mu(A) = \infty \Rightarrow \exists B \subset A : 0 < \mu(B) < \infty$. Visa att varje σ-ändligt mått är ett semiändligt mått och att om μ är semiändligt, $\mu(A) = \infty$ och $C < \infty$ finns det en mängd $B \subset A$ sådan att $C < \mu(B) < \infty$.

4. Låt μ vara Lebesguemåttet på $\mathcal{B}[0,1]$. Bevisa eller ge motexempel till följande två påståenden:

   (a) Om $\mu(A) > 0$ gäller att A innehåller ett öppet intervall.

   (b) Om A är sådan att $\mu(A^c) = 0$ så är A tät.

   (Att A är tät betyder att för alla $x \in [0,1]$ och alla $\delta > 0$ gäller att $(x-\delta, x+\delta) \cap A \neq \emptyset$.)

5. Ett måttrum (S, Σ, μ) kallas *fullständigt* om $A \subseteq B, B \in \Sigma, \mu(B) = 0 \Rightarrow A \in \Sigma$. Exempelvis är $(\mathbf{R}, \mathcal{L}(\mathbf{R}), Leb)$ ett fullständigt måttrum till skillnad från $(\mathbf{R}, \mathcal{B}(\mathbf{R}), Leb)$. Antag att (S, Σ, μ) är ett fullständigt måttrum och låt $\mathcal{J} = \{A \in \Sigma : \mu(A^c) = 0\}$. Visa att $\mathcal{J}$ är en topologi på S. Visa också i specialfallet $(S, \Sigma, \mu) = (\mathbf{R}, \mathcal{L}(\mathbf{R}), Leb)$ att $\mathcal{J}$ är T_1 men inte Hausdorff.

Kapitel 3

1. Antag att X är en stokastisk variabel sådan att $P(X = -\infty) = P(X = \infty) = 0$. Använd kontinuitet hos mått till att visa att $\lim_{x \downarrow -\infty} F(x) = 0$ och att $\lim_{x \uparrow \infty} F(x) = 1$.

2. Visa att om $f : \mathbf{R} \to \mathbf{R}$ är monoton gäller att f är en Borelfunktion.

3. Visa att om $f : S \to \overline{\mathbf{R}}$ och $f^{-1}([q, \infty]) \in \Sigma$ för alla rationella tal q så är f en Σ-mätbar funktion.

4. Låt $\mathcal{B}(\mathbf{R})$ vara Borel-σ-algebran på $\mathbf{R}$. För $a \in \mathbf{R}$ och $A \subseteq \mathbf{R}$, definiera $a + A = \{a + x : x \in A\}$. Visa att $a \in \mathbf{R}, A \in \mathcal{B}(\mathbf{R}) \Rightarrow a + A \in \mathcal{B}(\mathbf{R})$.

Kapitel 4

1. Borel-Cantellis andra lemma kräver oberoende stokastiska variabler. Konstruera ett exempel som visar att lemmat inte är generellt sant om man utelämnar kravet på oberoende.

2. Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler sådana att $P(X_n = -3^n) = 2^{-n}$ och $P(X_n = 1) = 1 - 2^{-n}$. Sätt $S_n = \sum_{k=1}^n X_k$. Visa att $\mathbf{E}[S_n] \to -\infty$ men att $S_n \to +\infty$ n.s.

3. Perkolation på $\mathbf{Z}^2$: Betrakta heltalsgittret i två dimensioner (d.v.s. den oändliga graf som fås genom att sätta en kant mellan varje två punkter, (x_1, x_2) och (y_1, y_2), i $\mathbf{Z}^2$ sådana att $|x_1 - y_1| + |x_2 - y_2| = 1$). Låt E vara mängden av kanter och låt $\{U_e\}_{e \in E}$ vara en uppsättning av oberoende stokastiska variabler likformigt fördelade på $[0, 1]$. Fixera $p \in [0, 1]$ och sätt för varje $e \in E$

   $$X_e = I_{[0,p]}(U_e).$$

   Låt C_p vara händelsen att det finns en oändlig sammanhängande komponent av kanter sådana att $X_e = 1$. (Vi bekymrar oss inte om konstruktion av sannolikhetsrum och mätbarhet av C_p här, men som vanligt är våra stokastiska variabler definierade på ett och samma sannolikhetsrum $(\Omega, \mathcal{F}, P)$.)

   (a) Visa att $P(C_p)$ är 0 eller 1.

   (b) Det är känt att $P(C_p) = 0$ då $p \leq 1/2$ och att $P(C_p) = 1$ då $p > 1/2$. Således gäller att $P(C_{1/2}) = 0$ medan $P(C_{1/2+\epsilon}) = 1$ för alla $\epsilon > 0$. Följande modell ligger "mittemellan": Sätt

   $$Y_e = I_{[0,1/2+\epsilon_e]}(U_e), e \in E$$

   där $\epsilon_e > 0$ för alla e och $\sum_{e \in E} \epsilon_e < \infty$. Låt $\tilde{C}$ vara händelsen att det finns en oändlig sammanhängande komponent av kanter sådana att $Y_e = 1$. Visa att $P(\tilde{C}) = 0$.

Kapitel 5

1. Låt ϕ och ψ vara enkla funktioner och visa att $\int_S (\phi + \psi) d\mu = \int_S \phi d\mu + \int_S \psi d\mu$.

2. Låt (S, Σ, μ) vara ett måttrum och tag $f \in (m\Sigma)^+$. Definiera måttet $\gamma(A) = \int_A f d\mu$, $A \in \Sigma$. Visa att om $h \in L^1(S, \Sigma, \gamma)$ gäller att $\int_S h d\gamma = \int_S h f d\mu$.

3. Visa att

   (a) $\int_1^2 e^{-t} \ln t dt = \lim_{n \to \infty} \int_1^2 (1 - t/n)^n \ln t dt$,

   (b) $\int_0^1 e^{-t} \ln t dt = \lim_{n \to \infty} \int_0^1 (1 - t/n)^n \ln t dt$.

4. Bevisa Scheffés lemma som säger att om f_n, $n = 1, 2, \ldots$ och f är ickenegativa Σ-mätbara integerbara funktioner sådana att $f_n \to f$ gäller att $\int_S |f_n - f| d\mu \to 0$ om och endast om $\int_S f d\mu \to \int_S f d\mu$.

Kapitel 6

1. Med hjälp av Hölders olikhet kan man bevisa monotoniciteten hos L^p-normer (dvs att för en stokastisk variabel X gäller att $1 \leq p \leq r < \infty \Rightarrow ||X||_p \leq ||X||_r$) på en rad. Gör detta.

2. Låt X vara en ickenegativ diskret stokastisk variabel i L^1och visa att

   (a) $P(X > 0) \leq \mathbf{E}[X]$,

   (b) $P(X > 0) \geq \frac{\mathbf{E}[X]^2}{\mathbf{E}[X^2]}$.

   Fotnot: Dessa två olikheter går under namnen "förstamomentsmetoden" respektive "andramomentsmetoden".

3. Låt X vara en stokastisk variabel i $L^2(\Omega, \mathcal{F}, P)$ och låt $\mu = \mathbf{E}[X]$. Sätt $Std(X) = Var(X)^{1/2}$. $Std(X)$ är ett naturligt mått på hur mycket man förväntar sig att X skall avvika från μ. Ett alternativ är att använda $\mathbf{E}[|X - \mu|]$. Hur förhåller sig $\mathbf{E}[|X - \mu|]$ och $Std(X)$ till varandra? Motivera!

4. Låt $T, X_1, X_2, X_3, \ldots$ vara ickenegativa stokastiska variabler sådana att T är diskret och $\sup_n \mathbf{E}[X_n^2] < \infty$.

   (a) Använd Schwarz olikhet till att visa att $\sum_{k=1}^{\infty} \sqrt{P(T \geq k)} < \infty \Rightarrow \mathbf{E}[\sum_{k=1}^{T} X_k] < \infty$.

   (b) Använd (a) till att visa att $T \in L^p \Rightarrow \mathbf{E}[\sum_{k=1}^{T} X_k] < \infty$ om $p > 2$.

   (Tips: I del (b) har man stor nytta av att veta att för en följd $\{a_n\}$ av ickenegativa reella tal gäller att $\sum_{n=1}^{\infty} a_n < \infty$ om och endast om det existerar en följd $\{b_n\}$ sådan att $\sum_{n=1}^{\infty} b_n^{-1} < \infty$ och $\sum_{n=1}^{\infty} b_n a_n^2 < \infty$.)

Kapitel 8

1. Låt $\Sigma = \mathcal{B}[0,1] \times \mathcal{B}[0,1]$ och låt $F = \{(x,y) \in [0,1] \times [0,1] : x = y\}$. Visa att $F \in \Sigma$.

2. Antag att X och Y är stokastiska variabler med fördelningarna P_X respektive P_Y. Visa att följande tre påståenden är ekvivalenta:

   (a) X och Y är oberoende,

   (b) $P_{X,Y} = P_X \times P_Y$,

   (c) $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ för alla $x, y \in \overline{\mathbf{R}}^2$,

   och att om $P_{X,Y} << Leb \times Leb$ kan man dessutom tillföra

   (d) $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ för nästan alla $x, y \in \mathbf{R}^2$.

3. Låt $S_1 = S_2 = \mathbf{N} = \{1, 2, \ldots\}$, låt $\Sigma_1 = \Sigma_2 = \mathcal{P}(\mathbf{N})$ och låt μ och ν vara räknemått på Σ_1 respektive Σ_2. Definiera funktionen $f : S_1 \times S_2 \to R$ så att $f(m,n) = 1$ då $m = n$, $f(m,n) = -1$ då $m = n+1$ och $f(m,n) = 0$ i övriga fall. Visa att både $\int_{S_1}(\int_{S_2} f d\nu)d\mu$ och $\int_{S_2}(\int_{S_1} f d\mu)d\nu$ existerar men är olika. Vad är det som gör att Fubinis sats spricker?

Kapitel 9

1. Visa att om $P(\cdot|\mathcal{G})(\cdot)$ är en reguljär betingad sannolikhet gäller för P-n.a. $\omega_0 \in \Omega$ att

$$\mathbf{E}[X|\mathcal{G}](\omega_0) = \int_\Omega X(\omega) P(d\omega|\mathcal{G})(\omega_0).$$

2. Låt $X_1, \ldots, X_r$ vara oberoende stokastiska variabler och låt $P_{X_1}, \ldots, P_{X_r}$ vara deras respektive fördelningar. Antag att $h \in L^1(\overline{\mathbf{R}}^r, \mathcal{B}^r, P_{X_1} \times \ldots \times P_{X_r})$ och sätt $\gamma^h(x_1) = \mathbf{E}[h(x_1, X_2, \ldots, X_r)]$, $x_1 \in \overline{\mathbf{R}}$. Använd Fubinis sats till att bevisa att $\gamma^h(X_1) = \mathbf{E}[h(X_1, \ldots, X_r)|X_1]$ n.s.

Kapitel 10

1. Walds sats: Låt $X_1, X_2, \ldots$ vara oberoende ickenegativa stokastiska variabler sådana att $\mathbf{E}[X_k] = \mu$ ($\mu \in (-\infty, \infty)$) för alla k. Sätt $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ och låt T vara en stopptid m.a.p. $\{\mathcal{F}_n\}$. Om $\mathbf{E}[T] < \infty$ och vi låter $S_T = \sum_{k=1}^{T} X_k = \sum_{k=1}^{\infty} X_k I_{\{T \geq k\}}$ gäller att

$$\mathbf{E}[S_T] = \mu \mathbf{E}[T].$$

Din uppgift är att bevisa detta.

2. Antag att man genererar en följd av slumptal, $X_1, X_2, \ldots$ som är oberoende och likformigt fördelade på $\{0, 1, \ldots, 9\}$. Låt T vara den första tidpunkt då kombinationen 8318 visar sig för första gången.

   (a) Beräkna $\mathbf{E}[T]$.

   (b) Beräkna sannolikhetsgenerande funktion för T, d.v.s. $G(s) = \mathbf{E}[s^T]$ för $s \in [0, 1]$.

3. Ovidkommande information förstör inte martingalegenskapen: Låt $\{X_n, \mathcal{F}_n\}$ vara en martingal. I definitionen av en martingalegenskapen är det tillåtet att ha att $\mathcal{F}_n$ är en strikt större σ-algebra än $\sigma(X_0, \ldots, X_n)$. Det är inte bara en matematisk teknikalitet att tillåta detta. Om vi tänker i termer av spelteori skall det ju naturligtvis vara "tillåtet" inför spelomgång $n+1$ att förutom att använda information given av $X_0, \ldots, X_n$ också använda annan ovidkommande information som står till buds, t.ex. att ta råd av en spågumma eller en astrolog. Detta förstör

dock inte martingalegenskapen. (Dvs att ty sig till skrock kan inte heller göra ett missgynnsamt spel gynnsamt.):

Låt $\mathcal{F}_\infty = \sigma(\cup_{n=1}^{\infty} \mathcal{F}_n)$ och antag att $\mathcal{H} \subseteq \mathcal{F}$ är en σ-algebra som är oberoende av $\mathcal{F}_\infty$. Definiera $\mathcal{G}_n = \sigma(\mathcal{F}_n, \mathcal{H})$. Visa att $\{X_n, \mathcal{G}_n\}$ är en martingal.

Anmärkning: Den som tror på spågummor och/eller astrologer kan ju alltid hävda att $\mathcal{H}$ inte är oberoende av $\mathcal{F}_\infty$.

Kapitel 12

1. Låt $\{X_n\}_{n=0}^{\infty}$ vara en Markovkedja med $(0,1) \cap Q$ som tillståndsrum. Antag att $0 < b \leq a < 1$ och att a och b är rationella. Definiera övergångar genom att låta

$$P(X_{n+1} = bx | X_n = x) = 1 - x$$

$$P(X_{n+1} = bx + 1 - a | X_n = x) = x$$

för alla $x \in (0,1) \cap Q$.

(a) Visa att om $a = b$ så är $\{X_n\}$ en martingal och att $X_n \to X_\infty$ n.s., där $P(X_\infty = 1) = 1 - P(X_\infty = 0) = \mathbf{E}[X_0]$.

(b) Visa att om $a > b$ så är $\{X_n\}$ en supermartingal och $X_n \to 0$ n.s.

Kapitel 13

1. Man kan bevisa följande generalisering av den optimala versionen av dominerad-konvergens-satsen (Sats 13.5): Tag $p \in [1, \infty)$ och låt $X, X_1, X_2, \ldots$ vara stokastiska variabler. Det gäller att $X_n \to X$ i L^p om och endast om $X_n \xrightarrow{P} X$ och $\{|X_n|^p\}$ är likformigt integrerbar. Gör detta!

Kapitel 14

1. Visa att om $\{M_n, \mathcal{F}_n\}$ är en martingal begränsad i L^p för något $p > 1$ gäller att det finns en stokastisk variabel M_∞ sådan att $M_n \to M_\infty$ n.s. och i L^p.

2. Ergodsatsen (d.v.s. "Stora talens lag") för utbytbara stokastiska variabler: Låt $X_1, X_2, \ldots$ vara en följd av *utbytbara* stokastiska variabler, d.v.s. det gäller att om $(k_n)_{n=1}^{\infty}$ är en ändlig permutation av $(1, 2, 3, \ldots)$ så gäller att

$$(X_{k_1}, X_{k_2}, \ldots) \stackrel{d}{=} (X_1, X_2, \ldots).$$

(Att (k_n) är en ändlig permutation betyder att för något $m < \infty$ gäller att $(k_1, \ldots, k_m)$ är en permutation av $(1, \ldots, m)$ och att $k_n = n$ för $n > m$.)

Antag också att X_k:na är integerbara så att $\mathbf{E}[X_k] = \mu$ existerar.

(a) Sätt $S_n = \sum_{k=1}^n X_k$ och $\mathcal{F}_n = \sigma(S_n, S_{n+1}, \ldots)$. Visa att $\mathbf{E}[X_1|\mathcal{F}_n] = S_n/n$ n.s.

(b) Enligt (a) tillsammans med kända resultat om konvergens av bakåtmartingaler gäller att

$$\frac{S_n}{n} \to L = \mathbf{E}[X_1|\mathcal{F}_\infty]$$

n.s. där $\mathcal{F}_\infty = \bigcap_n \mathcal{F}_n$. Uppenbarligen är gränsvariabeln L $\mathcal{F}_\infty$-mätbar. Det är också sant men men följer inte av ovanstående argument att L är mätbar m.a.p. svans-σ-algebran, $\mathcal{T}$, till X_k:na. Bevisa detta.

(c) I fallet då $X_1, X_2, \ldots$ är iid gäller att L är n.s. konstant, nämligen att $L = \mu$ n.s. Detta är dock inte generellt sant för utbytbara stokastiska variabler. Ge ett exempel sådant att L ej är n.s. konstant.

Referenser

1. R. B. Ash, "Real Analysis and Probability," Academic Press, 1972.
2. G. B. Folland, "Real Analysis," Wiley, 1984.
3. D. Williams, "Probability with Martingales," Cambridge University Press, 1991.