# SANNOLIKHETSTEORINS GRUNDER

**Olle Häggström** **Torgny Lindvall**

# Innehåll

# 0 INNAN VI BÖRJAR...

Detta kompendium är avsett som kursmaterial för en avancerad kurs i sannolikhetsteori på universitet eller teknisk högskola. Lämpliga förkunskaper har du som tidigare läst minst en inledande kurs i sannolikhetsteori eller matematisk statistik.

Meningen med kursen och kompendiet är att ge dig en förtrogenhet med sannolikhetsteorins matematiska fundament, och med dess viktigaste gränsvärdesresultat. Sannolikhetsteorin är idag ett brett forskningsområde inom matematiken, med omfattande tillämpningar inom teknik och naturvetenskap. Dessutom utgör den grunden för den matematiska statistiken. Jämfört med många andra grenar av matematiken är sannolikhetsteorin relativt ung (och vital), och till skillnad från de flesta matematikkurser du läst så behandlar denna nästan uteslutande 1900-talsmaterial. Om kursen lyckas med sitt syfte, så kommer du vid dess slut att ha fått en god inblick i hur den moderna teorin arbeter med sannolikheter och väntevärden, och i de frågeställningar, begrepp och resultat som ger sannolikhetsteorin dess egenart.

I början av kursen kommer stor kraft att läggas på att introducera och noggrant definiera ett antal grundläggande begrepp, av vilka du säkert stött på de flesta (t.ex. stokastisk variabel och väntevärde) tidigare. Det kan tyckas tradigt att behöva gå igenom allt detta en gång till. Det är emellertid mödan värt, då vi denna gång gör det på en matematisk nivå som kan göra betydligt större anspråk på att kallas ”stringent” eller ”rigorös”, jämfört med de argument och definitioner som du stötte på i din första sannolikhetskurs, som i många fall var alltför snäva/intuitiva/naiva.

Du minns förmodligen att väntevärdet $\mathbb{E}[X]$ för en stokastisk variabel $X$ ges av

$$\mathbb{E}[X] = \sum_x x\mathbb{P}(X = x) \tag{1}$$

om $X$ är diskret, och

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx \tag{2}$$

(där $f(x)$ är täthetsfunktionen till $X$) om $X$ är kontinuerlig. Men vad säger du om följande?

**Exempel 0.1.** I en vägkorsning med trafikljus lyser det omväxlande rött och grönt i 30-sekundersintervall. Antag att du anländer vid en ”på måfå” vald tidpunkt, och låt den stokastiska variabeln $X$ vara din väntetid vid ljuset, mätt i sekunder. Med sannolikhet $\frac{1}{2}$ är det grönt när du anländer, varvid $X = 0$. Med resterande sannolikhet $\frac{1}{2}$ är det rött, varvid den betingade fördelningen för $X$ är likformig på intervallet $[0, 30]$. Vad är $\mathbb{E}[X]$? Observera att $X$ varken är diskret eller kontinuerlig, så att vare sig (1) eller (2) är tillämplig.

Vi skall i Kapitel 3 komma fram till en definition av väntevärde som täcker in alla typer av stokastiska variabler.

Tonvikten i kompendiet är på den allmänna teorin, snarare än på dess tillämpningar. Vi bör dock ha i bakhuvudet att teorin inte finns till för sin egen skull, utan för att vara behjälplig i den konkreta tillämpningen. Meningen med nedanstående övningar 0.1–0.5 är för det första att framhålla att det finns ett liv bortom den abstrakta teorin, och för det andra att påminna dig om en del av den sannolikhetsteori du tidigare lärt dig. Vi tror att de kunskaper du inhämtat i exempelvis Ross [R] eller Rudemo & Råde [RR] är fullt tillräckliga för att du skall kunna lösa flertalet av övningarna (dock inte utan en viss ansträngning).

Du kommer förmodligen att finna att detta kompendium är en aning kortfattat på sina ställen, och vissa av de mer avancerade satserna ges utan bevis. Den ambitiöse läsaren uppmanas därför att ta en titt i Kapitel 9, där ett antal förslag till bredvidläsning ges.

Låt oss slutligen betona att det är *du själv* som skall lära dig sannolikhetsteorins grunder – ingen annan kan göra det åt dig! Endast ett enträget räknande och härledande kan ge en god förtrogenhet i ämnet. Därför förekommer många viktiga resultat endast som övningar, och därför bör du lösa alla eller nästan alla övningar i de kommande kapitlen.

**Övningar.**

0.1 Jeff och Joe spelar bordtennis. De är båda exakt lika bra, och de är såpass skickliga servemottagare att det inte innebär någon fördel att ha serven. Nu är de efter hård kamp framme i avgörande set, och Jeff leder med 18–15. Vad är sannolikheten att han vinner?

0.2 **Födelsedagsproblemet.** Hur stor behöver en skolklass vara för att sannolikheten skall vara minst $\frac{1}{2}$ att minst två av eleverna har samma födelsedag? Vi antar att elevernas födelsedagar är oberoende och likformigt fördelade över årets 365 dagar.

0.3 **Opålitlig informationsöverföring.** Ett enkelt ja/nej-meddelande skall framföras från en avsändare $A$ till en mottagare $M$. Detta sker via $n$ budbärare $B_1, \ldots, B_n$, och går till på så sätt att $A$ överför meddelandet till $B_1$, som i sin tur överför det till $B_2$, och så vidare, tills $B_n$ till slut överför det till $M$. Dock är budbärarna inte helt pålitliga. Var och en vidarebefordrar det budskap han uppfattat endast med sannolikhet $p \in (0,1)$; med resterande sannolikhet $q = 1 - p$ ger han motsatt budskap. Vad är sannolikheten att $M$ mottar rätt meddelande? Vad händer då $n \to \infty$?

0.4 Tre punkter väljes oberoende enligt likformig fördelning på en cirkel. Vad är sannolikheten att cirkeln kan delas upp i två halvcirklar på ett sådant sätt att alla tre punkterna hamnar på samma sida? Kan du generalisera detta till fallet med $n$ punkter?

0.5 Fundera på trafikljusexemplet ovan, och föreslå (utan att snegla på Kapitel 3) vilket värde på $\mathbb{E}[X]$ som en rimlig generell teori för väntevärden bör ge.

## 0.1 Definitioner och beteckningar

I detta korta avsnitt påminner vi om en del definitioner och beteckningar som fortsättningsvis kommer att antas kända.

Med en **fördelningsfunktion** $F$, menas en funktion $F : \mathbb{R} \to [0, 1]$ sådan att

(i) $\lim_{x\to-\infty} F(x) = 0$,

(ii) $\lim_{x\to\infty} F(x) = 1$,

(iii) $F$ är icke-avtagande, dvs om $x < y$ så är $F(x) \leq F(y)$, och

(iv) $F$ är högerkontinuerlig, dvs för varje $x$ gäller ett $\lim_{h\downarrow 0} F(x + h) = F(x)$.

(En del av vitsen med denna definition är naturligtvis att om $X$ är en stokastisk variabel, så är $F(x) = \mathbb{P}(X \leq x)$ en fördelningsfunktion.)

Med en **täthetsfunktion** $f$ menar vi en funktion $f : \mathbb{R} \to [0, \infty)$ sådan att

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \,.$$

Sambandet mellan täthetsfunktion och fördelningsfunktion ges i Övning 0.6 nedan.

□ □ □ □

Den grundläggande mängdläran är omöjlig att klara sig utan i sannolikhetsteorin. Här skall vi bara ge mängdlärans notation; om du känner dig osäker på hur man arbetar med skärningar, unioner etc, så bör du gå tillbaka till lämplig lärobok och repetera detta.

Låt $A$, $B$, och $A_1, A_2, \ldots$ vara delmängder till en mängd $\Omega$ som vi betraktar som ”universum”. Med $\omega \in A$ menar vi att elementet $\omega$ tillhör $A$. Om $A$ är en delmängd av $B$ skriver vi $A \subseteq B$, och om denna inklusion är strikt (dvs om $B$ innehåller minst ett element som inte ingår i $A$) så skriver vi även $A \subset B$. En sekvens $A_1, A_2, \ldots$ av mängder kallas växande om

$$A_1 \subseteq A_2 \subseteq \cdots$$

och avtagande om

$$A_1 \supseteq A_2 \supseteq \cdots \,.$$

$A \cap B$ och $A \cup B$ betecknar skärning och union. Vi skriver också

$$\bigcap_{i=1}^{n} A_i$$

för skärningen av $A_1, \ldots, A_n$, dvs för mängden av alla element $\omega \in \Omega$ som tillhör var och en av mängderna $A_1, \ldots, A_n$. På samma sätt skriver vi

$$\bigcup_{i=1}^{n} A_i$$

för unionen av $A_1, \ldots, A_n$, dvs för mängden av alla $\omega \in \Omega$ som tillhör minst en av mängderna $A_1, \ldots, A_n$. Vidare låter vi $A^c$ beteckna komplementet till $A$ i $\Omega$, dvs $A^c$ är mängden av alla $\omega \in \Omega$ som *inte* tillhör $A$. $A \setminus B$ betecknar mängdsubtraktion, dvs $A \setminus B = A \cap B^c$. Den tomma mängden betecknas $\emptyset$. Två mängder $A$ och $B$ kallas disjunkta om $A \cap B = \emptyset$.

**Övningar.**

0.6 Visa att om $f$ är en täthetsfunktion och $F(x) = \int_{-\infty}^{x} f(y)\, dy$, så är $F$ en fördelningsfunktion.

0.7 Visa de Morgans lagar, dvs

(a) $$\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{\infty} A_i^c \text{, och}$$

(b) $$\left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{\infty} A_i^c \,.$$

# 1 SANNOLIKHETSRUM

Det grundläggande matematiska objektet i sannolikhetsteorin är ett så kallat **sannolikhetsrum** $(\Omega, \mathcal{F}, \mathbb{P})$:

- $\Omega$ är mängden av alla tänkbara **utfall** i ett givet experiment. $\Omega$ kallas **utfallsrummet**.
- $\mathcal{F}$ är en familj av delmängder av $\Omega$. Dessa delmängder kallas **händelser**.
- $\mathbb{P}$ är en funktion från $\mathcal{F}$ till $[0,1]$ som tilldelar varje händelse $A \in \mathcal{F}$ en **sannolikhet** $\mathbb{P}(A)$.

$\mathcal{F}$ och $\mathbb{P}$ måste uppfylla vissa villkor, som vi strax kommer att ange. Vi börjar med att diskutera $\Omega$ och $\mathcal{F}$ i Avsnitt 1.1, för att sedan införa sannolikheter i Avsnitt 1.2.

## 1.1 Mätbara rum

Det är standard i sannolikhetsteorin att låta $\Omega$ beteckna det utfallsrum där Fru Fortuna väljer ett element $\omega$. $\Omega$ kan strängt taget vara vilken mängd som helst.

$\mathcal{F}$ är som sagt en familj av delmängder av $\Omega$. För att vi skall kunna räkna med sannolikheter på önskat sätt måste vi ställa vissa krav på $\mathcal{F}$.

**Definition 1.1** *Låt $\mathcal{F}$ vara en familj av delmängder av $\Omega$. Antag att $\mathcal{F}$ uppfyller*

(i) $\Omega \in \mathcal{F}$,

(ii) *om $A_1, A_2, \ldots \in \mathcal{F}$ så gäller även $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, och*

(iii) *om $A \in \mathcal{F}$ så gäller även $A^c \in \mathcal{F}$.*

*Då kallar vi $\mathcal{F}$ för en* $\sigma$-**algebra** *på $\Omega$.*

För att $(\Omega, \mathcal{F}, \mathbb{P})$ skall räknas som ett sannolikhetsrum kommer vi att kräva att $\mathcal{F}$ är en $\sigma$-algebra.

**Exempel 1.1.** Den *minsta* $\sigma$-algebran till ett utfallsrum $\Omega$ är $\mathcal{F} = \{\emptyset, \Omega\}$ (denna $\sigma$-algebra är dock tämligen ointressant).

**Exempel 1.2.** Låt $A$ vara en godtycklig delmängd av $\Omega$. $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ är då en $\sigma$-algebra.

**Exempel 1.3.** Den *största* $\sigma$-algebran till ett utfallsrum $\Omega$ fås genom att helt enkelt låta $\mathcal{F}$ utgöra familjen av alla delmängder till $\Omega$. Denna familj brukar betecknas $\{0, 1\}^\Omega$.

Ett par $(\Omega, \mathcal{F})$ där $\mathcal{F}$ är en $\sigma$-algebra på $\Omega$ kallas ett **mätbart rum**. Mängderna i $\mathcal{F}$ kallas **mätbara**. Det är till varje mätbar mängd vi så småningom skall ge ett mått, speciellt en sannolikhet.

Här kanske du frågar dig: Varför väljer vi inte helt enkelt alltid $\mathcal{F} = \{0, 1\}^\Omega$, så att alla delmängder till $\Omega$ blir mätbara? När utfallsrummet $\Omega$ är ändligt eller uppräkneligt är det mycket riktigt så man brukar göra. Om däremot $\Omega$ är överuppräkneligt (exempelvis $\Omega = \mathbb{R}$) visar det sig att $\{0, 1\}^\Omega$ är en alltför stor $\sigma$-algebra – i många situationer finns det helt enkelt inte någon tillordning av ”sannolikhetsmassa” åt alla dessa mängder som uppfyller kraven på ett mått[1] (vi återkommer i Avsnitt 1.2 till vilka dessa krav är).

Låt $\mathcal{A}$ vara en godtycklig klass av delmängder till $\Omega$. Det finns då en minsta $\sigma$-algebra som innehåller $\mathcal{A}$. Denna kallar vi för **den av $\mathcal{A}$ genererade $\sigma$-algebran**, och vi betecknar den med $\sigma(\mathcal{A})$. $\mathcal{A}$ erhålles genom att ta skärningen av alla $\sigma$-algebror som innehåller $\mathcal{A}$; se Övningarna 1.4 och 1.5 nedan.

Exemplet $\Omega = \mathbb{R}$ är av fundamental betydelse. Tyvärr är $\sigma$-algebran $\{0, 1\}^\mathbb{R}$ av alla delmängder för stor. För konstruktionen av en något mindre $\sigma$-algebra som fungerar, låt $\mathcal{G}$ vara klassen av alla öppna delmängder till $\mathbb{R}$. Vi påminner om att en delmängd $A \subseteq \mathbb{R}$ kallas öppen om det för varje $x \in A$ existerar något $\varepsilon > 0$ sådant att $(x - \varepsilon, x + \varepsilon) \subseteq A$.

**Definition 1.2** *Med* **Borelmängderna** *på* $\mathbb{R}$ *menar vi de som ingår i* $\sigma(\mathcal{G})$. *Denna klass betecknas med* $\mathcal{B}$.

## Övningar.

1.1 Visa att de tre $\sigma$-algebrorna i Exempel 1.1–1.3 ovan verkligen är $\sigma$-algebror.

[1]Den som önskar se ett exempel på vad som kan gå snett med valet $\mathcal{F} = \{0, 1\}^\Omega$ då $\Omega$ är överuppräkneligt, kan exempelvis vända sig till Appendix A1.1 i Williams [W].

1.2 Visa att om $\mathcal{F}$ är en icketom klass av delmängder till $\Omega$ som uppfyller (ii) och (iii) i Definition 1.1, så är $\mathcal{F}$ en $\sigma$-algebra.

1.3 Fixera en Borelmängd $A \in \mathcal{B}$, och låt $\mathcal{B}_A = \{B : B \in \mathcal{B}, B \subseteq A\}$. Visa att $\mathcal{B}_A$ är en $\sigma$-algebra av delmängder till $A$. Vi kallar $\mathcal{B}_A$ för "Borelmängderna i $A$"; ett viktigt exempel är $A = [0, 1]$.

1.4 Låt $\mathcal{F}$ och $\mathcal{H}$ vara $\sigma$-algebror på mängden $\Omega$. Visa att $\mathcal{F} \cap \mathcal{H}$ är en $\sigma$-algebra, men att $\mathcal{F} \cup \mathcal{H}$ inte behöver vara det.

1.5 Visa att skärningen av en godtycklig klass av $\sigma$-algebror på $\Omega$ åter är en $\sigma$-algebra.

1.6 En delmängd $A$ av $\mathbb{R}$ kallas sluten om dess komplement $A^c$ är öppen. Visa att $\mathcal{B} = \sigma$(de slutna mängderna på $\mathbb{R}$).

1.7 Visa att $\mathcal{B}$ innehåller alla uppräkneliga delmängder till $\mathbb{R}$.

1.8 Visa att $\mathcal{G} \subseteq \sigma$(de öppna intervallen). Utnyttja sedan detta till att visa att $\mathcal{B} = \sigma$(de öppna intervallen).

1.9 Visa att $\mathcal{B} = \sigma$(intervallen av typ $(-\infty, x]$, $x \in \mathbb{R}$).

1.10 Låt $(\Omega, \mathcal{F})$ vara ett mätbart rum, och betrakta en sekvens $A_1, A_2, \ldots$ av händelser i $\mathcal{F}$. Definiera

$$B_n = \bigcup_{m=n}^{\infty} A_m \quad \text{och} \quad C_n = \bigcap_{m=n}^{\infty} A_m$$

och observera att $C_n \subseteq A_n \subseteq B_n$ för varje $n$. Vidare är sekvensen $B_1, B_2, \ldots$ avtagande med "gränsvärde"

$$B = \lim_{n\to\infty} B_n = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \,,$$

medan sekvensen $C_1, C_2, \ldots$ är växande med motsvarande gräns

$$C = \lim_{n\to\infty} C_n = \bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m \,.$$

$B$ och $C$ kallas $\limsup_{n\to\infty} A_n$ och $\liminf_{n\to\infty} A_n$. Visa att

(a) $B$ och $C$ är mätbara, dvs $B, C \in \mathcal{F}$,

(b) $B = \{\omega \in \Omega : \omega \in A_n$ för oändligt många $n\}$,

(c) $C = \{\omega \in \Omega : \omega \in A_n$ för alla utom högst ändligt många $n\}$.

## 1.2 Mått och sannolikheter

**Definition 1.3** *Ett* **mått** $\mu$ *på ett mätbart rum* $(\Omega, \mathcal{F})$ *är en avbildning* $\mu : \mathcal{F} \to [0, \infty]$ *som uppfyller*

(i) $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$ *om mängderna* $A_1, A_2, \ldots$ *är disjunkta, och*

(ii) $\mu(\emptyset) = 0$.

Måttet för en mängd måste alltså vara ickenegativt, men kan mycket väl anta värdet $\infty$. Egenskaperna (i) och (ii) gör det naturligt att tänka sig $\mu$ som en fördelning av massa på $\Omega$. Om totalmassan är 1, dvs om $\mu(\Omega) = 1$, kallas $\mu$ ett **sannolikhetsmått**. En trippel $(\Omega, \mathcal{F}, \mu)$ kallas ett **måttrum**. Om $\mu$ är ett sannolikhetsmått så talar vi om ett **sannolikhetsrum**, och betecknar i så fall oftast måttet med $\mathbb{P}$ istället för $\mu$ ($\mathbb{P}$ som i "probabilité").

**Exempel 1.4.** Låt $\Omega$ vara en uppräknelig mängd $\{\omega_i : i \in \mathbb{N}\}$, och låt $\mathcal{F} = \{0,1\}^\Omega$. Låt $\{p_i : i \in \mathbb{N}\}$ vara ickenegativa tal sådana att $\sum_{i=1}^\infty p_i = 1$. För varje $A \in \mathcal{F}$, låt $\mathbb{P}(A) = \sum_{\omega_i \in A} p_i$. Då är $(\Omega, \mathcal{F}, \mathbb{P})$ ett sannolikhetsrum.

För nästa, mycket viktiga, exempel krävs ett djupt resultat, som vi inte har någon möjlighet att bevisa här.

**Sats 1.1** *Det existerar ett (unikt) mått $m$ på $(\mathbb{R}, \mathcal{B})$, kallat* **Lebesguemåttet**, *med egenskapen att*

$$m([a,b]) = b - a$$

*för alla $a, b \in \mathbb{R}$ sådana att $a < b$.*

Lebesguemåttet för ett slutet intevall är alltså lika med dess längd. Detsamma gäller öppna och halvöppna intervall (se Övning 1.13).

**Exempel 1.5.** Låt $(\Omega, \mathcal{F}) = ([0,1], \mathcal{B}_{[0,1]})$ där $\mathcal{B}_{[0,1]}$ är Borelmängderna i $[0,1]$, och låt $\mathbb{P}(A) = m(A)$ för $A \in \mathcal{B}_{[0,1]}$. Detta sannolikhetsmått kallar vi **den likformiga fördelningen på** $[0,1]$.

Nästa sats (vars fullständiga bevis vi åter måste utelämna) säger oss att det finns rikligt med sannolikhetsmått på $(\mathbb{R}, \mathcal{B})$; se Avsnitt 0.1 för definitionen av fördelningsfunktion.

**Sats 1.2** *Till varje fördelningsfunktion $F$ hör ett unikt sannolikhetsmått $\mathbb{P}$ på $(\mathbb{R}, \mathcal{B})$ som satisfierar*

$$\mathbb{P}((-\infty, x]) = F(x)$$

*för alla $x \in \mathbb{R}$. Omvänt: ett sannolikhetsmått $\mathbb{P}$ på $(\mathbb{R}, \mathcal{B})$ bestämmer en fördelningsfunktion $F$ genom denna relation.*

Den nämnda omvändningen skall vi visa i Övning 1.20. Första halvan är ett betydligt djupare resultat.

**Exempel 1.6.** Låt $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B})$ och låt $f \geq 0$ vara en (Riemann-integrerbar) funktion sådan att $\int_{-\infty}^{\infty} f(x)dx = 1$. Vi såg i Övning 0.6 att funktionen $F$, definierad genom

$$F(x) = \int_{-\infty}^{x} f(y)dy\,,$$

är en fördelningsfunktion. Sats 1.2 ger nu att $f$ definierar ett sannolikhetsmått $\mathbb{P}$ på $(\mathbb{R}, \mathcal{B})$ sådant att $\mu(-\infty, x] = F(x)$ för $x \in \mathbb{R}$. Måttet $\mathbb{P}(A)$ för $A \in \mathcal{B}$ skrivs ofta som

$$\int_A f(x)dx$$

och $\mathbb{P}$ sägs ha **täthetsfunktionen** $f$. Men vad menar vi egentligen med integration över en godtycklig Borelmängd $A$? Vi kommer att besvara den frågan i Kapitel 3. Observera att den vanliga Riemann-integrationen inte klarar av denna integral för andra fall än då $A$ är ett intervall eller liknande.

**Exempel 1.7.** Det går att konstruera ett sannolikhetsmått $\mathbb{P}$ på $([0,1], \mathcal{B}_{[0,1]})$ och en sluten mängd $C \in \mathcal{B}_{[0,1]}$ sådana att $\mathbb{P}(C) = 1$, Lebesguemåttet $m(C) = 0$ *och den till* $\mathbb{P}$ *hörande fördelningsfunktionen är kontinuerlig.* Ett sådant $\mathbb{P}$ har inte någon täthetsfunktion. Som $C$ kan tas den så kallade Cantormängden. Vi återkommer till detta exempel i Avsnitt 8.1, där vi skall se att det ger oss ett mer subtilt exempel än det i Kapitel 0 på en stokastisk variabel som varken är diskret eller kontinuerlig i vanlig mening.

## Övningar.

1.11 Bevisa Booles olikheter, dvs

(a) $$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)\,, \text{ och}$$

(b) $$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) \geq 1 - \sum_{i=1}^{n} \mathbb{P}(A_i^c)\,.$$

1.12 Låt $A$ och $B$ vara händelser på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Visa att om $A \subseteq B$ så är $\mathbb{P}(A) \leq \mathbb{P}(B)$.

1.13 Visa för $a, b \in \mathbb{R}$ sådana att $a < b$, att

$$m((a,b)) = m((a,b]) = m([a,b)) = a - b\,.$$

1.14 Vad bör menas med den likformiga fördelningen på ett intervall $[a,b]$?

1.15 Elementet $\omega \in \Omega$ sägs vara en **atom** för ett sannolikhetsmått $\mathbb{P}$ om $\mathbb{P}(\{\omega\}) > 0$. Visa att $\mathbb{P}$ kan ha högst uppräkneligt många atomer.

1.16 Låt $\mathbb{P}$ vara den likformiga fördelningen på $([0,1], \mathcal{B}_{[0,1]})$, och låt $\mathbb{Q}_{[0,1]}$ beteckna de rationella talen i $[0,1]$. Visa att $\mathbb{P}(\mathbb{Q}_{[0,1]}) = 0$.

1.17 Punkten $x \in \mathbb{R}$ sägs vara en stödpunkt till sannolikhetsmåttet $\mathbb{P}$ på $(\mathbb{R}, \mathcal{B})$ om $\mathbb{P}(x - \varepsilon, x + \varepsilon) > 0$ för alla $\varepsilon > 0$. Mängden av stödpunkter kallas **stödet** för $\mathbb{P}$. Visa att stödet är en sluten mängd. På vad sätt kan du rättfärdiga påståendet att stödet är komplementet till den största öppna mängden med $\mathbb{P}$-måttet 0?

1.18 (a) Låt $A_1, A_2, \ldots$ vara en växande följd av händelser, och låt $A = \bigcup_{k=1}^{\infty} A_k$. Visa att $\mathbb{P}(A) = \lim_{k\to\infty} P(A_k)$.

(b) Låt istället $A_1, A_2, \ldots$ vara en avtagande följd av händelser, och $A = \bigcap_{k=1}^{\infty} A_k$. Visa att $\mathbb{P}(A) = \lim_{k\to\infty} P(A_k)$.

1.19 Vi återknyter till Övning 1.10. Om $\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$ så betecknar vi det gemensamma värdet med $\lim_{n\to\infty} A_n$. Visa att om $\lim_{n\to\infty} A_n = A$ så är $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

1.20 Visa den lättare halvan av Sats 1.2, dvs att om $\mathbb{P}$ är ett sannolikhetsmått på $(\mathbb{R}, \mathcal{B})$ så är funktionen $F$, definierad av

$$F(x) = \mathbb{P}((-\infty, x]) \quad \text{för alla } x \in \mathbb{R},$$

en fördelningsfunktion.

1.21 **Betingad sannolikhet.** Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum, och låt $A \in \mathcal{F}$ vara sådan att $\mathbb{P}(A) > 0$. För varje $B \in \mathcal{F}$, definiera $\mu(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$. Visa att $\mu$ är ett sannolikhetsmått. (Som du troligen vet, så brukar $\mu(B)$ betecknas $\mathbb{P}(A|B)$, och kallas den betingade sannolikheten för $B$ givet $A$.)

# 2 STOKASTISKA VARIABLER

Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum och $X$ en avbildning från $\Omega$ till $\mathbb{R}$. Det torde vara överflödigt att påpeka att man i sannolikhetsteorin ofta ställs inför uppgiften att bestämma $\mathbb{P}(X \in A)$, där $A \subset \mathbb{R}$. Men för att kunna göra det måste ju

$$\{X \in A\} = \{\omega : X(\omega) \in A\} = X^{-1}A$$

tillhöra $\mathcal{F}$, vilket ju är definitionsområdet för $\mathbb{P}$ ($\mathcal{F}$ är de delmängder av $\Omega$ vilkas sannolikhet låter sig mätas).

Har du någonsin behövt beräkna en sannolikhet $\mathbb{P}(X \in A)$ där $A$ inte är en Borelmängd? Troligen är svaret nej. Å andra sidan är det av flera skäl önskvärt att $\{X \in A\} \in \mathcal{F}$ för alla $A \in \mathcal{B}$.

**Definition 2.1** *Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum, och låt $X$ vara en avbildning $\Omega \to \mathbb{R}$. Avbildningen sägs vara en* **stokastisk variabel** *om*

$$\{X \in A\} = \{\omega : X(\omega) \in A\} = X^{-1}A \in \mathcal{F} \quad \text{för alla} \quad A \in \mathcal{B}.$$

Detta är ett specialfall av följande.

**Definition 2.2** *Låt $(\Omega_1, \mathcal{F}_1)$ och $(\Omega_2, \mathcal{F}_2)$ vara mätbara rum. En avbildning $h : \Omega_1 \to \Omega_2$ sägs vara* **mätbar** *om $h^{-1}B = \{x \in \Omega_1 : h(x) \in B\} \in \mathcal{F}_1$ för alla $B \in \mathcal{F}_2$.*

En mätbar avbildning kallas för stokastisk variabel endast i närvaro av ett sannolikhetsmått.

Finns det ett någorlunda gripbart kriterium för när en avbildning $X : \Omega \to \mathbb{R}$ är en stokastisk variabel? Svaret finns i följande sats.

**Sats 2.1** *Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum. En avbildning $X : \Omega \to \mathbb{R}$ är en stokastisk variabel om och endast om*

$$\{X \le x\} = \{\omega : X(\omega) \le x\} = X^{-1}(-\infty, x] \in \mathcal{F} \quad \text{för alla} \quad x \in \mathbb{R}.$$

**Bevis.** "Endast om"-implikationen av satsen är klar eftersom intervallen $(-\infty, x]$ tillhör $\mathcal{B}$. Återstår alltså att visa "om"-implikationen. Antag att

$$X^{-1}(-\infty, x] \in \mathcal{F} \tag{3}$$

för alla $x \in \mathbb{R}$. Vi behöver visa att $X^{-1}A \in \mathcal{F}$ för alla $A \in \mathcal{B}$. Låt

$$\mathcal{K} = \{A : A \in \mathcal{B}, X^{-1}A \in \mathcal{F}\}. \tag{4}$$

I Övning 2.1 visas att $\mathcal{K}$ är en $\sigma$-algebra. Men pga antagandet (3) har vi att $(-\infty, x] \in \mathcal{K}$ för alla $x \in \mathbb{R}$ vilket ger

$$\sigma\{(-\infty, x] : x \in \mathbb{R}\} \subseteq \mathcal{K}.$$

Enligt Övning 1.9 är $\sigma\{(-\infty, x] : x \in \mathbb{R}\} = \mathcal{B}$, så vi har alltså att $\mathcal{B} \subseteq \mathcal{K}$. Detta är samma sak som att säga att $X^{-1}A \in \mathcal{F}$ för varje $A \in \mathcal{B}$, och beviset är klart. □

Vitsen med Sats 2.1[2] är att klassen av mängder $A$ för vilka man behöver kontrollera att $X^{-1}A \in \mathcal{F}$ blir betydligt mindre, och mer hanterbar: istället för att visa att $X^{-1}A \in \mathcal{F}$ för alla $A \in \mathcal{F}$ räcker det att visa att $X^{-1}(-\infty, x] \in \mathcal{F}$ för alla $x \in \mathbb{R}$.

Följande resultat är mycket nyttigt för hanteringen av stokastiska variabler. För två funktioner $h : \Omega_1 \to \Omega_2$ och $g : \Omega_2 \to \Omega_3$, definierar vi funktionen $g \circ h$ genom att sätta $(g \circ h)(x) = g(h(x))$ för varje $x \in \Omega_1$. Vi kallar $g \circ h$ för kompositionen av $g$ och $h$.

**Sats 2.2** *Låt $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ och $(\Omega_3, \mathcal{F}_3)$ vara mätbara rum. Antag $g : \Omega_2 \to \Omega_3$ och $h : \Omega_1 \to \Omega_2$ är mätbara avbildningar. Då är $g \circ h$ en mätbar avbildning från $\Omega_1$ till $\Omega_3$.*

**Bevis.** Vi skall visa att $(g \circ h)^{-1}A \in \mathcal{F}_1$ för alla $A \in \mathcal{F}_3$. Observera att

$$(g \circ h)^{-1}A = h^{-1}(g^{-1}A).$$

Eftersom $g$ är mätbar har vi $g^{-1}A \in \mathcal{F}_2$, och eftersom $h$ är mätbar har vi $h^{-1}(g^{-1}A) \in \mathcal{F}_1$. □

**Korollarium 2.1** *Låt $X$ vara en stokastisk variabel och antag att funktionen $f : \mathbb{R} \to \mathbb{R}$ är kontinuerlig. Då är $f(X)$ en stokastisk variabel.*

**Bevis.** Med stöd av Sats 2.2 är det tillräckligt att $f$ är mätbar; satsens $(\Omega_2, \mathcal{F}_2)$ och $(\Omega_3, \mathcal{F}_3)$ är båda lika med $(\mathbb{R}, \mathcal{B})$. Enligt Sats 2.1 (nu med andra beteckningar) räcker det att visa, att $f^{-1}(-\infty, x] \in \mathcal{B}$ för alla $x$. Men när $f$ är kontinuerlig så är $f^{-1}(-\infty, x]$ sluten eftersom $(-\infty, x]$ är det, och de slutna mängderna är Borelmängder. □

När vi talar om mätbara funktioner $f : \mathbb{R} \to \mathbb{R}$ är det oftast underförstått att vi har det mätbara rummet $(\mathbb{R}, \mathcal{B})$ i åtanke, så $f$ är mätbar om och endast om $f^{-1}A \in \mathcal{B}$ för alla $A \in \mathcal{B}$.

[2] Vissa författare föredrar att använda Sats 2.1 som definition av vad en stokastiska variabel är; då blir istället vår Definition 2.1 en sats.

En särkskild klass av stokastiska variabler som vi kommer att ha stor nytta av, bland annat vid utvecklandet av teorin för väntevärden i Kapitel 3, är de så kallade **indikatorvariablerna**:

**Definition 2.3** *Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum. Till varje händelse $A \in \mathcal{F}$ kan vi definiera en stokastisk variabel $I_A$, kallad* **indikatorvariabeln för** *$A$, genom*

$$I_A(\omega) = \begin{cases} 1 & \textit{för } \omega \in A\,, \\ 0 & \textit{för } \omega \notin A\,. \end{cases}$$

□ □ □ □

Hur skall man beskriva det sätt på vilket en stokastisk variabel sprider ut sina värden på $\mathbb{R}$? Fördelningsfunktionen är en möjlighet. Ett alternativ, som vi kommer att ha stor användning för, är att beskriva det som ett sannolikhetsmått på $\mathbb{R}$:

**Sats 2.3** *Låt $X$ vara en stokastisk variabel på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Låt*

$$\mu(A) = \mathbb{P}(X \in A)$$

*för varje Borelmängd $A$. Då är $\mu$ ett sannolikhetsmått på $(\mathbb{R}, \mathcal{B})$.*

**Bevis.** Vi har naturligtvis $\mu(\emptyset) = 0$ och $\mu(\mathbb{R}) = 1$. Låt nu $A_1, A_2, \ldots$, vara disjunkta Borelmängder. Vi får

$$\left\{ X \in \bigcup_{i=1}^{\infty} A_i \right\} = X^{-1}\left( \bigcup_{i=1}^{\infty} A_i \right) = \bigcup_{i=1}^{\infty} X^{-1} A_i = \bigcup_{i=1}^{\infty} \{X \in A_i\}\,,$$

och eftersom mängderna $\{X \in A_i\}$ är disjunkta (se Övning 2.3) följer det att

$$\begin{aligned} \mu\left( \bigcup_{i=1}^{\infty} A_i \right) &= \mathbb{P}\left( X \in \bigcup_{i=1}^{\infty} A_i \right) = \mathbb{P}\left( \bigcup_{i=1}^{\infty} \{X \in A_i\} \right) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(X \in A_i) = \sum_{i=1}^{\infty} \mu(A_i)\,. \end{aligned}$$

□

Sannolikhetsmåttet $\mu$ i Sats 2.3 kallas **fördelningen för den stokastiska variablen** $X$. Ofta är benämningen $\mu_X$ lämplig för detta mått.

**Fördelningsfunktionen för** $X$ definieras genom

$$F_X(x) = \mathbb{P}(X \leq x) = \mu_X(-\infty, x] \quad \text{för} \quad x \in \mathbb{R}\,.$$

Vi har sett i Sats 1.2 och Övning 1.20 att funktionen $F_X$ definierad på detta sätt verkligen blir en fördelningsfunktion.

Ett sannolikhetsmått $\mathbb{P}$ på ett mätbart rum sägs vara **diskret** om sannolikhetsmassan är koncentrerad till uppräkneligt många punkter, d.v.s. $\exists\, \omega_1, \omega_2, \ldots$, så att $\sum_1^\infty \mathbb{P}(\{\omega_i\}) = 1$. En stokastisk variabel kallas diskret om dess fördelning på $(\mathbb{R}, \mathcal{B})$ är det.

Från din första kurs i sannolikhetsteori är du antagligen van vid att kalla en stokastisk variabel $X$ för **kontinuerlig** om och endast om det existerar en ickenegativ funktion $f$ sådan att

$$F_X(x) = \int_{-\infty}^{x} f(y)dy \quad \text{för} \quad x \in \mathbb{R}\,. \tag{5}$$

Benämningen ”de kontinuerliga stokastiska variablerna” för denna klass av stokastiska variabler är något olycklig (ty det existerar stokastiska variabler $X$ som inte har täthet trots att $F_X$ är kontinuerlig; se Exempel 1.7), men den är så etablerad att vi ändå accepterar den[3].

Med en diskret (resp. kontinuerlig) fördelning menar vi fördelningen för en diskret (resp. kontinuerlig) stokastisk variabel.

De flesta (men långt ifrån alla) stokastiska variabler man stöter på i praktiken är antingen diskreta eller kontinuerliga. Om du inte redan gjort det bör du nu plocka fram kursboken från din första kurs i sannolikhetsteori, och repetera de viktigaste diskreta och kontinuerliga fördelningarna: binomial, Poisson, normal, likformig, exponential...

## Övningar.

2.1 Fullborda beviset av Sats 2.1 genom att visa att $\mathcal{K}$ (definierad i (4)) är en $\sigma$-algebra på $\mathbb{R}$.

2.2 Visa att följande funktioner $f : \mathbb{R} \to \mathbb{R}$ är mätbara:

(a) $f_1(x) = \begin{cases} 1 & \text{om} \quad x \quad \text{rationellt} \\ 0 & \text{om} \quad x \quad \text{irrationellt} \end{cases}$

(b) $f_2(x) = \begin{cases} 0 & \text{om} \quad x \leq 0 \\ \sin(1/x) & \text{om} \quad x > 0 \end{cases}$

(c) $f_3(x) = \begin{cases} 1 & \text{om} \quad x \in A, \quad \text{där } A \text{ är en given Borelmängd} \\ 0 & \text{om} \quad x \notin A. \end{cases}$

[3]En bättre benämning på klassen av sådana variabler vore måhända ”de stokastiska variablerna med täthet”. I mer avancerade texter kallas de ofta ”absolutkontinuerliga”.

2.3 Låt $X$ vara en stokastisk variablel på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och låt $A_1$ och $A_2$ vara Borelmängder. Visa att om $A_1$ och $A_2$ är disjunkta, så är händelserna $\{X \in A_1\}$ och $\{X \in A_2\}$ disjunkta. Ge också ett motexempel som visar att den omvända implikationen inte gäller.

2.4 Låt $X$ vara en stokastisk variabel med fördelningsfunktion $F_X$. Visa att $F_X(x)$ är kontinuerlig i punkten $x$ om och endast om $\mathbb{P}(X = x) = 0$.

2.5 Låt $X$ vara en stokastisk variabel med fördelningsfunktion $F_X$ och fördelning $\mu_X$. Fyra besläktade stokastiska variabler ges av

$$\begin{cases} X^+ = \max(0, X) \\ X^- = -\min(0, X) \\ |X| = X^+ + X^- \\ -X \end{cases}$$

(a) Uttryck fördelningsfunktionerna $F_{X^+}$, $F_{X^-}$, $F_{|X|}$ och $F_{-X}$ i termer av $F_X$.

(b) Uttryck fördelningarna $\mu_{X^+}$, $\mu_{X^-}$, $\mu_{|X|}$ och $\mu_{-X}$ i termer av $\mu_X$. (För att göra detta kan det vara lämpligt att införa beteckningen $-A = \{x \in \mathbb{R} : -x \in A\}$ för $A \in \mathcal{B}$.)

# 3 VÄNTEVÄRDEN OCH INTEGRATION

## 3.1 Väntevärden av stokastiska variabler

Som vi redan framhållit, så är de "vardagliga" definitionerna av väntevärde (se ekvationerna (1) och (2) i Kapitel 0) otillräckliga. Det är nu dags att, med stöd av den måtteori du lärt dig i Kapitel 1 och 2, bygga upp en teori för väntevärden av stokastiska variabler definierade på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Detta är en del av den så kallade integrationsteorin, som behandlas mer allmänt av Folland [F]; en utsökt bok med tonvikt på sannolikhetsteori är Williams [W].

Låt oss vara överens om två önskvärda egenskaper hos väntevärdesoperatorn $\mathbb{E}[\cdot]$:

(i) $\mathbb{E}[I_A] = \mathbb{P}(A) \quad \text{för} \quad A \in \mathcal{F}$.

(ii) För två godtyckliga stokastiska variabler $X$ och $Y$, och konstanter $a$ och $b$, skall gälla $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

Önskemålet (ii) säger att $\mathbb{E}[\cdot]$ skall vara en **lineär funktional**. Med hjälp av induktion i $n$ ser vi (Övning 3.1) att (ii) implicerar

(ii') $\mathbb{E}[a_1X_1 + \cdots + a_nX_n] = a_1\mathbb{E}[X_1] + \cdots + a_n\mathbb{E}[X_n]$

för $a_1, \ldots, a_n \in \mathbb{R}$ och stokastiska variabler $X_1, \ldots, X_n$.

Det kommer strax att visa sig, att så snart vi anammat (i) och (ii) har vi också bundit upp oss ganska ordentligt för hur väntevärdet av en stokastisk variabel skall definieras.

Den stokastiska variabeln $X$ kallas **enkel** om den antar endast ändligt många värden. Låt $\mathcal{S}_+$ beteckna klassen av ickenegativa enkla stokastiska variabler. En stokastisk variabel $X \in \mathcal{S}_+$ med värdemängd $\{x_1, \ldots, x_n\}$ kan skrivas som

$$X = \sum_{k=1}^{n} x_k I_{\{X=x_k\}} .$$

Varje term i högerledets summa är i sig en stokastisk variabel, så om vi accepterar (ii') och (i) ovan, ser vi att

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{k=1}^{n} x_k I_{\{X=x_k\}}\right] = \sum_{k=1}^{n} x_k \mathbb{E}[I_{\{X=x_k\}}] = \sum_{k=1}^{n} x_k \mathbb{P}(X = x_k)$$

och vi gör därför följande definition.

**Definition 3.1** *För en enkel ickenegativ stokastisk variabel $X \in \mathcal{S}_+$ med värdemängd $\{x_1, \ldots, x_n\}$ definierar vi väntevärdet*

$$\mathbb{E}[X] = \sum_{k=1}^{n} x_k \mathbb{P}(X = x_k)\,.$$

(Jämför med ekvation (1) i Kapitel 0.)

I Lemma 3.2–3.6 nedan, skall vi se att $\mathbb{E}[X]$ definierad som ovan har ett antal naturliga och användbara egenskaper. Först visar vi ett preliminärt resultat (Lemma 3.1) som är bekvämt att ta till för att bevisa Lemma 3.2 och 3.3.

Om $A_1, \ldots, A_n$ är disjunkta mängder i $\mathcal{F}$ sådana att $\bigcup_1^n A_i = \Omega$, så kallar vi $(A_1, \ldots, A_n)$ för en **sönderläggning** eller **partition** av $\Omega$. Ett exempel på en partition av $\Omega$ är $(\{X = x_1\}, \ldots, \{X = x_n\})$, där $X$ är en enkel ickenegativ stokastisk variabel, och $\{x_1, \ldots, x_n\}$ är värdemängden till $X$.

**Lemma 3.1** *Låt $X$ vara en enkel ickenegativ stokastisk variabel. Antag att $(A_1, \ldots, A_n)$ är en partition av $\Omega$, och att $a_1, \ldots, a_n$ är ickenegativa tal sådana att $X(\omega) = a_i$ för varje $\omega \in A_i$. Då är*

$$\sum_{i=1}^{n} a_i \mathbb{P}(A_i) = \mathbb{E}[X]\,. \tag{6}$$

**Bevis.** Låt $\{x_1, \ldots, x_m\}$ beteckna värdemängden för $X$. En sortering av termerna i vänsterledet till (6) ger

$$\begin{aligned}
\sum_{i=1}^{n} a_i \mathbb{P}(A_i) &= \sum_{k=1}^{m} \sum_{i:\, a_i = x_k} a_i \mathbb{P}(A_i) = \sum_{k=1}^{m} \sum_{i:\, a_i = x_k} x_k \mathbb{P}(A_i) \\
&= \sum_{k=1}^{m} x_k \sum_{i:\, a_i = x_k} \mathbb{P}(A_i) = \sum_{k=1}^{m} x_k \mathbb{P}(X = x_k) = \mathbb{E}[X]\,.
\end{aligned}$$

$\square$

Med hjälp härav kan vi visa följande två lemmata: det första låter oss jämföra väntevärdena hos två stokastiska variabler, och det andra visar att Definition 3.1 ger upphov till ett väntevärdesbegrepp som uppfyller den önskade egenskapen (ii) på sidan 18.

**Lemma 3.2** *Antag att $X, Y \in \mathcal{S}_+$ är sådana att $X \leq Y$ (dvs $X(\omega) \leq Y(\omega)$ för varje $\omega \in \Omega$). Då gäller*
$$\mathbb{E}[X] \leq \mathbb{E}[Y] .$$

**Bevis.** Se Övning 3.2. □

**Lemma 3.3** *För $X_1, \ldots, X_n \in \mathcal{S}_+$ och $a_1, \ldots, a_n \geq 0$ gäller att*
$$\mathbb{E}[a_1 X_1 + \cdots + a_n X_n] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n] .$$

**Bevis.** Fallet $n = 2$ bevisas i Övning 3.3, och det allmänna fallet följer sedan av Övning 3.1. □

Ytterligare ett användbart lemma är följande.

**Lemma 3.4** *Låt $X_1, X_2, \ldots$ vara stokastiska variabler sådana att $X_n \in \mathcal{S}_+$ för varje $n$, och antag att $X_n(\omega) \downarrow 0$ för varje $\omega \in \Omega$. Då gäller*
$$\mathbb{E}[X_n] \downarrow 0 \quad \textit{då} \quad n \to \infty.$$

**Bevis.** Enligt Lemma 3.2 gäller
$$\mathbb{E}[X_1] \geq \mathbb{E}[X_2] \geq \cdots$$
så det räcker att visa att $\lim_{n\to\infty} \mathbb{E}[X_n] = 0$. Låt $K = \sup_\omega X_1(\omega)$, och välj ett $\varepsilon > 0$. Vi får
$$\begin{aligned}
\mathbb{E}[X_n] &= \mathbb{E}[X_n I_{\{X_n > \varepsilon\}}] + \mathbb{E}[X_n I_{\{X_n \leq \varepsilon\}}] \\
&\leq \mathbb{E}[X_1 I_{\{X_n > \varepsilon\}}] + \varepsilon \\
&\leq K \mathbb{E}[I_{\{X_n > \varepsilon\}}] + \varepsilon \\
&= K \mathbb{P}(X_n > \varepsilon) + \varepsilon
\end{aligned}$$
(de båda olikheterna följer av Lemma 3.2). Men $\mathbb{P}(X_n > \varepsilon) \to 0$ då $n \to \infty$ (se Övning 1.18 (b)) och eftersom $\varepsilon > 0$ var godtyckligt är saken klar. □

En enkel men viktig konsekvens av Lemma 3.4 är följande.

**Lemma 3.5** *För $X, X_1, X_2, \ldots \in \mathcal{S}_+$, antag att $X_n \uparrow X$ punktvis (dvs att $X_n(\omega) \uparrow X(\omega)$ för varje $\omega \in \Omega$) då $n \to \infty$. Då gäller att*
$$\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$$
*då $n \to \infty$.*

**Bevis.** $\mathbb{E}[X] = \mathbb{E}[X_n] + \mathbb{E}[X - X_n]$, och $\mathbb{E}[X - X_n] \downarrow 0$ enligt Lemma 3.4. □

**Lemma 3.6** *Låt $X'_n, n = 1, 2, \ldots$, och $X_n$, $n = 1, 2, \ldots$, vara icke avtagande sekvenser av stokastiska variabler i $\mathcal{S}_+$ sådana att $\lim_{n\to\infty} X'_n \leq \lim_{n\to\infty} X_n$. Då gäller*

$$\lim_{n\to\infty} \mathbb{E}[X'_n] \leq \lim_{n\to\infty} \mathbb{E}[X_n]. \tag{7}$$

**Bevis.** Fixera ett $m \in \mathbb{N}$. Vi erhåller

$$\lim_{n\to\infty} \mathbb{E}[X_n] \geq \lim_{n\to\infty} \mathbb{E}[\min(X'_m, X_n)] = \mathbb{E}[X'_m]\,,$$

där olikheten följer av Lemma 3.2 och likheten följer av Lemma 3.5, ty $\min(X'_m, X_n) \uparrow X'_m \in \mathcal{S}_+$ då $n \to \infty$. Om vi nu låter $m \to \infty$ erhålles (7). □

□ □ □ □

Låt $\mathcal{M}_+$ beteckna klassen av ickenegativa stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. Vi är mogna för att definiera väntevärdet för $X \in \mathcal{M}_+$.

**Definition 3.2** *För $X \in \mathcal{M}_+$ definieras $\mathbb{E}[X]$ genom*

$$\mathbb{E}[X] = \sup_{\substack{Y \leq X \\ Y \in \mathcal{S}_+}} \mathbb{E}[Y]\,.$$

För att motivera definitionen, observera att till $X \in \mathcal{M}_+$ hör en sekvens av variabler $Z_n \in \mathcal{S}_+$, $n \in \mathbb{N}$, sådan att $Z_n(\omega) \uparrow X(\omega)$ för varje $\omega \in \Omega$. Låt nämligen $Z_n$ vara definierad genom

$$Z_n(\omega) = \begin{cases} \frac{i-1}{2^n} & \text{om} \quad \frac{i-1}{2^n} \leq X(\omega) < \frac{i}{2^n}, \ \text{för} \quad 1 \leq i \leq n2^n\,, \\ 0 & \text{om} \quad X(\omega) \geq n\,. \end{cases} \tag{8}$$

Med andra ord: $Z_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} I_{\{(i-1)/2^n \leq X < i/2^n\}}$.

Observera att $Z_1 \leq Z_2 \leq \ldots \leq X$, och att $0 \leq X(\omega) - Z_n(\omega) \leq 1/2^n$ för $\omega \in \{X < n\}$. Notera också att vi inte uteslutit möjligheten att $\mathbb{E}[X] = \infty$.

Vi önskar nu utvidga våra resultat rörande stokastiska variabler i $\mathcal{S}_+$, till stokastiska variabler i $\mathcal{M}_+$. Följande lemma är den länk mellan $\mathcal{S}_+$ och $\mathcal{M}_+$ som tillåter oss att göra detta.

**Lemma 3.7** *För $X \in \mathcal{M}_+$ och $X_n \in \mathcal{S}_+$, $n \in \mathbb{N}$, antag att $X_n \uparrow X$ punktvis då $n \to \infty$. Då gäller $\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$ när $n \to \infty$.*

**Bevis.** Lemma 3.2 implicarar att $\mathbb{E}[X_1] \leq \mathbb{E}[X_2] \leq \cdots$, och Definition 3.2 ger att $\mathbb{E}[X_n] \leq \mathbb{E}[X]$, så det återstår alltså bara att visa att $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

En annan konsekvens av definitionen av $\mathbb{E}[X]$ är att det finns en sekvens $Y_n$, $n \in \mathbb{N}$, sådan att $Y_n \in \mathcal{S}_+$, $Y_n \leq X$ och $\mathbb{E}[Y_n] \to \mathbb{E}[X]$ när $n \to \infty$. Låt $X'_n = \max(Y_1, \ldots, Y_n)$, vilket gör $X'_n$, $n \in \mathbb{N}$, till en växande sekvens. Att $X'_n \leq X$ ger nu

$$\lim_{n\to\infty} X'_n \leq X = \lim_{n\to\infty} X_n \,,$$

och Lemma 3.6 medför

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[Y_n] \leq \lim_{n\to\infty} \mathbb{E}[X'_n] \leq \lim_{n\to\infty} \mathbb{E}[X_n] \leq \mathbb{E}[X].$$

□

**Sats 3.1** *Låt $X, Y \in \mathcal{M}_+$ vara ickenegativa stokastiska variabler.*

(i) *För $a, b \geq 0$ gäller*
$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y] \,.$$

(ii) *Om $X \leq Y$ så är*
$$\mathbb{E}[X] \leq \mathbb{E}[Y].$$

**Bevis.** Låt $X_n \in \mathcal{S}_+$, $n \in \mathbb{N}$, vara en sekvens av enkla ickenegativa stokastiska variabler sådan att $X_n \uparrow X$ då $n \to \infty$; observera genom konstruktionen (8) ovan att vi alltid kan hitta en sådan sekvens. Låt även $Y_n \in \mathcal{S}_+$, $n \in \mathbb{N}$, vara en motsvarande sekvens sådan att $Y_n \uparrow Y$ då $n \to \infty$. Vi erhåller

$$\mathbb{E}[aX + bY] = \lim_{n\to\infty} \mathbb{E}[aX_n + bY_n] = \lim_{n\to\infty} (a\mathbb{E}[X_n] + b\mathbb{E}[Y_n]) = a\mathbb{E}[X] + b\mathbb{E}[Y] \,,$$

där första och tredje likheten följer av Lemma 3.7, och den andra följer av Lemma 3.3. Därmed är (i) bevisad. För att visa (ii) noterar vi att

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[X_n] \leq \lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[Y]$$

där de båda likheterna följer av Lemma 3.7, och olikheten följer av Lemma 3.6. Därmed är även (ii) bevisad. □

Sats 3.1 (i) förutsätter konventionen att $c \cdot \infty = \infty$ för $c > 0$, och $0 \cdot \infty = 0$, eftersom vi inte antagit att $\mathbb{E}[X] < \infty$.

Det är nu dags för de två huvudsatserna för $\mathcal{M}_+$.

**Sats 3.2 (Satsen om monoton konvergens, ”Monotona konvergenssatsen”)** *För $X, X_1, X_2, \ldots \in \mathcal{M}_+$, antag att $X_n \uparrow X$ punktvis när $n \to \infty$. Då gäller att*

$$\mathbb{E}[X_n] \uparrow \mathbb{E}[X] \quad \text{då } n \to \infty.$$

**Bevis.** Enligt Sats 3.1 (ii) är $\mathbb{E}[X_1] \leq \mathbb{E}[X_2] \leq \cdots \leq \mathbb{E}[X]$, så det räcker alltså att visa att $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$. Tag en sekvens $Y_n' \in \mathcal{S}_+, n \in \mathbb{N}$, sådan att $Y_n' \leq X_n$ och $Y_n' \to X$ punktvis då $n \to \infty$. Att en sådan sekvens finns följer av konstruktionen (8); se Övning 3.5. Låt $Y_n = \max(Y_1', \ldots, Y_n')$. Då gäller $Y_n \in \mathcal{S}_+$, $Y_n \uparrow X$, och Lemma 3.7 implicerar att $\mathbb{E}[Y_n] \uparrow \mathbb{E}[X]$ när $n \to \infty$. Men $Y_n \leq X_n$ vilket ger

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[Y_n] \leq \lim_{n\to\infty} \mathbb{E}[X_n] \leq \mathbb{E}[X].$$

□

**Sats 3.3 (Fatous lemma)** *För $X_1, X_2, \ldots \in \mathcal{M}_+$ gäller att*

$$\mathbb{E}[\liminf_{n\to\infty} X_n] \leq \liminf_{n\to\infty} \mathbb{E}[X_n].$$

**Bevis.** Låt

$$Y_n = \inf_{k\geq n} \; X_k \,.$$

Då gäller att $Y_n \uparrow \liminf_{m\to\infty} X_m$, och Monotona konvergenssatsen ger

$$\mathbb{E}[\liminf_{m\to\infty} X_m] = \lim_{n\to\infty} \mathbb{E}[Y_n] = \liminf_{n\to\infty} \mathbb{E}[Y_n] \leq \liminf_{n\to\infty} \mathbb{E}[X_n]\,.$$

□

Ett viktigt specialfall är om $X, X_1, X_2, \ldots \in \mathcal{M}_+$ är sådana att $X_n \to X$ punktvis. Då ger Fatous lemma att

$$\mathbb{E}[X] \leq \liminf_{n\to\infty} \mathbb{E}[X_n]\,.$$

Man skulle kunna tro att denna olikhet kunde skärpas till en likhet, men så är inte fallet. Se Övning 3.6.

□ □ □ □

Vi kan nu gå till klassen av alla stokastiska variabler definierade på $(\Omega, \mathcal{F}, \mathbb{P})$; vi tillåter alltså nu att de även antar negativa värden. För en stokastisk variabel $X$ definierar vi (som i Övning 2.5) $X^+ = \max(0, X)$ och $X^- = -\min(0, X)$, så att $X = X^+ - X^-$.

**Definition 3.3** *En stokastisk variabel $X$ sägs ha* **ändligt väntevärde** *om $\mathbb{E}[|X|] < \infty$, och väntevärdet definieras då genom*

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

Vi börjar med att visa två satser som generaliserar Sats 3.1 till det allmänna fallet.

**Sats 3.4** *Låt $X, Y$ vara stokastiska variabler med ändligt väntevärde, och $a, b \in \mathbb{R}$. Då är*

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

**Bevis.** Det räcker att bevisa att

(i) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, och

(ii) $\mathbb{E}[cX] = c\mathbb{E}[X]$ för $c \in \mathbb{R}$,

ty när detta är gjort får vi $\mathbb{E}[aX + bY] = \mathbb{E}[aX] + \mathbb{E}[bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$. För att visa (i), notera först att $X + Y$ kan skrivas både som $(X + Y)^+ - (X + Y)^-$ och alternativt som $(X^+ + Y^+) - (X^- + Y^-)$. Alltså gäller det att

$$(X + Y)^+ + (X^- + Y^-) = (X^+ + Y^+) + (X + Y)^-,$$

och med hjälp av Sats 3.1 (i) får vi att

$$\mathbb{E}[(X + Y)^+] + \mathbb{E}[X^- + Y^-] = \mathbb{E}[X^+ + Y^+] + \mathbb{E}[(X + Y)^-].$$

Detta medför att

$$\begin{aligned} \mathbb{E}[X + Y] &= \mathbb{E}[(X + Y)^+] - \mathbb{E}[(X + Y)^-] \\ &= \mathbb{E}[X^+ + Y^+] - \mathbb{E}[X^- + Y^-] \\ &= \mathbb{E}[X^+] - \mathbb{E}[X^-] + \mathbb{E}[Y^+] - \mathbb{E}[Y^-] \\ &= \mathbb{E}[X] + \mathbb{E}[Y], \end{aligned}$$

och (i) är bevisad. Vi går vidare med beviset av (ii). Antag att $c \geq 0$. Då är $\mathbb{E}[(cX)^+] = c\mathbb{E}[X^+]$ och $\mathbb{E}[(cX)^-] = c\mathbb{E}[X^-]$, på grund av Sats 3.1 (i). Om istället $c < 0$ får vi $\mathbb{E}[(cX)^+] = -c\mathbb{E}[X^-]$ och $\mathbb{E}[(cX)^-] = -c\mathbb{E}[X^+]$. I båda fallen får vi

$$\begin{aligned} \mathbb{E}[cX] &= \mathbb{E}[(cX)^+] - \mathbb{E}[(cX)^-] \\ &= c\mathbb{E}[X^+] - c\mathbb{E}[X^-] \\ &= c(\mathbb{E}[X^+] - \mathbb{E}[X^-]) \\ &= c\mathbb{E}[X] \end{aligned}$$

där alltså den andra olikheten fås genom att behandla fallen $c \geq 0$ och $c < 0$ separat. Därmed är även (ii) bevisad. $\square$

**Sats 3.5** *Låt $X$ och $Y$ vara stokastiska variabler som uppfyller $X \leq Y$ och som har ändliga väntevärden. Då gäller*

$$\mathbb{E}[X] \leq \mathbb{E}[Y] .$$

**Bevis.** Om $X \leq Y$ gäller att $X^+ \leq Y^+$ och $X^- \geq Y^-$. Definition 3.3 och Sats 3.1 ger därför

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-] \leq \mathbb{E}[Y^+] - \mathbb{E}[Y^-] = \mathbb{E}[Y] .$$

□

Nästa sats är, jämte Monotona konvergenssatsen, ett av de viktigaste grundläggande redskapen i sannolikhetsteorin.

**Sats 3.6 ("Dominerade konvergenssatsen")** *Låt $X, X_1, X_2, \ldots$ och $Y$ vara stokastiska variabler sådana att $X_n \to X$ punktvis, $|X_n| \leq Y$ för alla $n$, och $\mathbb{E}[|Y|] < \infty$. Då gäller*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X] . \tag{9}$$

**Bevis.** Vi har $X_n + Y \geq 0$ och $X + Y \geq 0$. Fatous lemma ger

$$\begin{aligned}
\mathbb{E}[X] + \mathbb{E}[Y] &= \mathbb{E}[X + Y] \\
&\leq \liminf_{n\to\infty} \mathbb{E}[X_n + Y] \\
&= \liminf_{n\to\infty} (\mathbb{E}[X_n] + \mathbb{E}[Y]) \\
&= \liminf_{n\to\infty} \mathbb{E}[X_n] + \mathbb{E}[Y],
\end{aligned}$$

vilket (tack vare att $\mathbb{E}[Y] < \infty$) medför

$$\mathbb{E}[X] \leq \liminf_{n\to\infty} \mathbb{E}[X_n] . \tag{10}$$

Men vi har också $Y - X_n \geq 0$ och $Y - X \geq 0$. Fatou hjälper oss igen:

$$\begin{aligned}
\mathbb{E}[Y] - \mathbb{E}[X] &= \mathbb{E}[Y - X] \\
&\leq \liminf_{n\to\infty} \mathbb{E}[Y - X_n] \\
&= \liminf_{n\to\infty} (\mathbb{E}[Y] - \mathbb{E}[X_n]) \\
&= \mathbb{E}[Y] - \limsup_{n\to\infty} \mathbb{E}[X_n] ,
\end{aligned}$$

så att $\mathbb{E}[X] \geq \limsup_{n\to\infty} \mathbb{E}[X_n]$, vilket i kombination med (10) medför (9). □

Ett ofta använt specialfall är följande:

**Korollarium 3.1 ("Begränsade konvergenssatsen")** *Låt $X, X_1, X_2, \ldots$ vara stokastiska variabler sådana att $|X_n| \leq K$ för någon fix konstant $K$ och alla $n$. Om $X_n \to X$ punktvis då $n \to \infty$, så gäller även $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.*

**Bevis.** Definiera den stokastiska variabeln $Y$ genom att låta $Y(\omega) = K$ för varje $\omega \in \Omega$, och tillämpa Dominerade konvergenssatsen. □

## Övningar.

3.1 Visa att postulaten (ii) och (ii') på sidan 18 är ekvivalenta.

3.2 Bevisa Lemma 3.2. (Ledning: Använd Lemma 3.1, och sönderlägg $\Omega$ genom att låta $A_{i,j} = \{\omega : X(\omega) = x_i, Y(\omega) = y_j\}$.)

3.3 Bevisa Lemma 3.3 i fallet $n = 2$. (Ledning: Använd Lemma 3.1 på liknande sätt som i Övning 3.2.)

3.4 Låt $X$ vara en ickenegativ stokastisk variabel. Visa att $\mathbb{E}[X] < \infty$ om och endast om $\sum_0^\infty \mathbb{P}(X > n) < \infty$.

3.5 Låt $X, X_1, X_2, \ldots$ vara ickenegativa stokastiska variabler sådana att $X_n \uparrow X$ punktvis då $n \to \infty$. Fyll i luckan i beviset av Sats 3.2 genom att visa att $Y_n' \uparrow X$, där $Y_n'$ definieras som

$$Y_n'(\omega) = \begin{cases} \frac{i-1}{2^n} & \text{om} \quad \frac{i-1}{2^n} \leq X_n(\omega) < \frac{i}{2^n}, \quad \text{för} \quad 1 \leq i \leq n2^n\,, \\ 0 & \text{om} \quad X_n(\omega) \geq n\,. \end{cases}$$

3.6 Ge ett (eller gärna flera) exempel på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$ och ickenegativa stokastiska variabler $X$, $X_n$ definierade på $(\Omega, \mathcal{F}, \mathbb{P})$ sådana att $X_n \to X$ punktvis, men $\mathbb{E}[X_n] \to \mathbb{E}[X]$ inte är sant. Det kan till och med gälla, att $\mathbb{E}[X_n] = \infty$ för alla $n$ och $X = 0$.

3.7 Antag att $X_1, X_2, \ldots$ är stokastiska variabler som uppfyller

$$\mathbb{E}\left[\sum_{k=1}^\infty |X_k|\right] < \infty\,.$$

Visa att

$$\mathbb{E}\left[\sum_{k=1}^\infty X_k\right] = \sum_{k=1}^\infty \mathbb{E}[X_k]\,.$$

3.8 Låt $X_1, X_2, \ldots$ och $Y$ vara stokastiska variabler sådana att $|X_n| \leq Y$ för alla $n$, och $\mathbb{E}[Y] < \infty$. Visa att

$$\mathbb{E}\left[\limsup_{n\to\infty} X_n\right] \geq \limsup_{n\to\infty} \mathbb{E}[X_n]\,.$$

3.9 **Betingat väntevärde.** Låt $X$ vara en stokastisk variabel med ändligt väntevärde på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och låt $A \in \mathcal{F}$ vara sådan att $\mathbb{P}(A) > 0$. Vi definierar **det betingade väntevärdet för $X$ givet $A$**, som

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[XI_A]}{\mathbb{P}(A)}\,.$$

Visa att

(a) om såväl $X$ som $Y$ har ändliga väntevärden, så är $\mathbb{E}[aX + bY|A] = a\mathbb{E}[X|A] + b\mathbb{E}[Y|A]$, och

(b) om $(A_1, \ldots, A_k)$ är en partition av $\Omega$ sådan att $\mathbb{P}(A_i) > 0, \forall i$, så är

$$\mathbb{E}[X] = \sum_{i=1}^k \mathbb{E}[X|A_i]\mathbb{P}(A_i)\,.$$

## 3.2 Mer allmän integrationsteori

En person med kunskaper i mått- och integrationsteori skulle här säga att vi byggt upp en teori för integration för måttrum med måttet 1 för hela rummet och påpeka, att de stokastiska variablerna med ändligt väntevärde är inget annat än klassen av **integrabla** funktioner på $(\Omega, \mathcal{F}, \mathbb{P})$. I den allmänna integrationsteorins tappning kallas $X$ integrabel om $\mathbb{E}[|X|] < \infty$, och $\mathbb{E}[X]$ skrivs omväxlande som

$$\int X(\omega)\mathbb{P}(d\omega), \int X(\omega)d\mathbb{P}(\omega), \int X d\mathbb{P}$$

eller, om ingen tvekan råder, som $\int X$, och kallas **integralen** av $X$.

Ett mått $\mu$ på ett mätbart rum $(\Omega, \mathcal{F})$ kallas **$\sigma$-ändligt** om det existerar mängder $A_1, A_2, \ldots \in \mathcal{F}$, sådana att $\mu(A_i) < \infty$ för alla $i$, och

$$\bigcup_{i=1}^{\infty} A_i = \Omega\,.$$

Integrationsteorin för klassen av mätbara funktioner på ett måttrum $(\Omega, \mathcal{F}, \mu)$ där $\mu$ är $\sigma$-ändligt är föga mera komplicerad än fallet med $\mu(\Omega) = 1$. Det finns motsvarigheter för $\sigma$-ändliga mått till alla de viktigaste resultaten i Avsnitt 3.1 (detta gäller däremot inte mått som inte är $\sigma$-ändliga).

Lebesguemåttet $m$ på $(\mathbb{R}, \mathcal{B})$ är $\sigma$-ändligt (se Övning 3.10). För en integrabel funktion $f$ på $(\mathbb{R}, \mathcal{B}, m)$ kallas $\int f(x)dm(x)$ för **Lebesgueintegralen** av $f$; traditionen bjuder att skriva denna som

$$\int f(x)dx \quad \text{eller} \quad \int f dx,$$

i stället för $\int f(x)dm(x)$ eller dylikt.

Det går att visa att för godartade funktioner (exempelvis de styckvis kontinuerliga) är Riemann- och Lebesgueintegralerna lika. Men se på $g$ definierad på $[0, 1]$ genom

$$g(x) = \begin{cases} 1, & \text{om } x \text{ irrationellt} \\ 0, & \text{om } x \text{ rationellt.} \end{cases}$$

Funktionen $g$ är inte Riemannintegrabel, däremot Lebesgueintegrabel.

Med hjälp av resultaten i Avsnitt 3.1 behärskar vi integration av funktioner definierade på intervall av längden 1, t.ex. $[k, k+1)$ för ett heltal $k$, eftersom $([k, k+1), \mathcal{B}_{[k,k+1)}, m_{[k,k+1)})$ är ett sannolikhetsrum (här betecknar $m_{[k,k+1)}$ naturligtvis Lebesguemåttet inskränkt till Borelmängderna på $[k, k+1)$). En mätbar funktion $f$ på $(\mathbb{R}, \mathcal{B})$ är Lebesgueintegrabel om

$$\sum_{k=-\infty}^{\infty} \int_{[k,k+1)} |f(x)|dx < \infty.$$

Med hjälp av formeln

$$\int f(x)dx = \sum_{k=-\infty}^{\infty} \int_{[k,k+1)} f(x)dx.$$

kan vi utvidga teorin för väntevärden till en teori för Lebesgueintegration.

**Övningar.**

3.10 Visa att Lebesguemåttet på $(\mathbb{R}, \mathcal{B})$ är $\sigma$-ändligt.

## 3.3 Kontinuerliga stokastiska variabler

Du har en gång lärt dig (se formeln (2) i Kapitel 0) att väntevärdet för en kontinuerlig stokastisk variabel $x$ är lika med

$$\int x f_X(x)dx,$$

där $f_X$ är täthetsfunktionen för $x$ och vi förutsatt att $\int |x| f_X(x)dx < \infty$. Nu är det dags att inse hur detta väntevärdesbegrepp hänger samman med det vi lärt oss i Avsnitt 3.1. Det väsentliga för den insikten är följande variabelsubstitutionsresultat. Vi påminner oss från Kapitel 2 (Sats 2.3) att **fördelningen** för $\mu_X$ för en stokastisk variabel $X$ är det sannolikhetsmått på $(\mathbb{R}, \mathcal{B})$ som uppfyller $\mu_X(A) = \mathbb{P}(X \in A)$ för alla $A \in \mathcal{B}$.

**Sats 3.7** *Låt den stokastiska variabeln $X$ var definierad på sannolikhetsrummet $(\Omega, \mathcal{F}, \mathbb{P})$, låt $g : \mathbb{R} \to \mathbb{R}$ vara en mätbar funktion och antag att $g(X)$ har ett ändligt väntevärde. Då gäller att*

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) d\mu_X(x)$$

*där $\mu_X$ är fördelningen för $X$ på $(\mathbb{R}, \mathcal{B})$.*

**Bevis.** Vi använder den fyrstegsmetod som Williams [W] kallar **standardmaskinen**, nämligen:

0. Visa att resultatet är sant då $g$ är en indikatorfunktion.
1. Utvidga till fallet då $g$ är enkel och ickenegativ.
2. Utvidga till fallet då $g$ bara antas mätbar och ickenegativ.
3. Utvidga till fallet då $g$ bara antas mätbar.

(Vi har valt att numrera stegen 0–3 (snarare än 1–4), så att stegen 1–3 skall svara mot de tre steg i vilka vi definierade väntevärde i Avsnitt 3.1.

*Steg 0:* Resultatet är sant för $g = I_A$, $A \in \mathcal{B}$, ty

$$\mathbb{E}[I_A(X)] = \mathbb{P}(X \in A) = \mu_X(A) = \int I_A(x) d\mu_X(x)$$

enligt definitionen av $\mu_X$.

*Steg 1:* Resultatet är också sant för ickenegativa enkla funktioner $g = \sum_{i=1}^{n} a_i I_{A_i}$ där $a_i \geq 0$ och $A_i \in \mathcal{B}$, ty

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^{n} a_i I_{A_i}(X)\right] &= \sum_{i=1}^{n} a_i \mathbb{E}[I_{A_i}(X)] \\ &= \sum_{i=1}^{n} a_i \int I_{A_i}(x) d\mu_X(x) \\ &= \int \left[\sum_{i=1}^{n} a_i I_{A_i}(x)\right] d\mu_X(x). \end{aligned}$$

(Första och tredje likheten följer av Lemma 3.3, medan den andra följer av Steg 1.)

*Steg 2:* Låt nu $g$ vara en godtycklig mätbar ickenegativ funktion på $(\mathbb{R}, \mathcal{B})$. Vi vet att det finns en sekvens av enkla ickenegativa funktioner $g_n$ sådan att $g_n \uparrow g$ då $n \to \infty$. Vi får

$$\mathbb{E}[g(X)] = \lim_{n\to\infty} \mathbb{E}[g_n(X)] = \lim_{n\to\infty} \int g_n(x) d\mu_X(x) = \int g(x) d\mu_X(x)\,, \tag{11}$$

där första och tredje likheten följer av Monotona konvergenssatsen, medan den andra följer av Steg 2.

*Steg 3:* Satsens resultat för en godtycklig funktion $g$ fås genom följande tillämpning av (11) på $g^+$ och $g^-$ där $g = g^+ - g^-$, den vanliga uppdelningen av $g$ i sin positiva och negativa del:

$$\begin{aligned} \mathbb{E}[g(X)] &= \mathbb{E}[g^+(X) - g^-(X)] = \mathbb{E}[g^+(X)] - \mathbb{E}[g^-(X)] \\ &= \int g^-(x) d\mu_X(x) - \int g^+(x) d\mu_X(x) = \int g(x) d\mu_X(x)\,. \end{aligned}$$

□

Speciellt får vi, med $g(x) = x$, att

$$\mathbb{E}[X] = \int x d\mu_X(x)$$

för ett $X$ med ändligt väntevärde[4].

Antag nu att $X$ är en kontinuerlig stokastisk variabel, dvs det finns en mätbar funktion $f_X : \mathbb{R} \to \mathbb{R}$ sådan att $f_X \geq 0$ och

$$\mathbb{P}(X \in A) = \int f_X(x) I_A(x) dx$$

för alla $A \in \mathcal{B}$, där $\int$ betecknar Lebesgueintegralen.

**Lemma 3.8** *Om $X$ är en kontinuerlig stokastisk variabel med täthetsfunktion $f_X$, så gäller för varje mätbar funktion $g : \mathbb{R} \to \mathbb{R}$ sådan att $g(X)$ har ändligt väntevärde, att*

$$\int g(x) d\mu_X(x) = \int g(x) f_X(x) dx\,.$$

**Bevis.** Följer genom att tillämpa standardmaskinen som i beviset av Sats 3.7. Påpekandet i början av Avsnitt 3.2 att resultaten i Avsnitt 3.1 har motsvarigheter för $\sigma$-ändliga mått (och därmed speciellt för Lebesguemåttet) är här väsentligt: i Steg 1 och 3 behöver vi använda additivitet hos Lebesgueintegralen, och i Steg 2 används Lebesgueintegralsmotsvarigheten till Monotona konvergenssatsen. $\square$

Sammanfattar vi det vi lärt oss i Sats 3.7 och Lemma 3.8, får vi

$$\mathbb{E}[g(X)] = \int g(x) d\mu_X(x) = \int g(x) f_X(x) dx$$

om $X$ har täthet $f_X$ och $\mathbb{E}[|g(X)|] < \infty$. Specialfallet $g(x) = x$ ger

$$\mathbb{E}[X] = \int x f_X(x) dx$$

om $\mathbb{E}[|X|] < \infty$, och därmed är vi framme vid formeln (2) på sidan 3, som du förmodligen lärde dig ett par veckor in på din första sannolikhetskurs.

Här kan du pusta ut en stund, och begrunda din lärdomsgång vad integration beträffar. Vilka funktioner är Riemannintegrabla? Över vilka mängder kan man Riemannintegrera? Hur definierades egentligen $\mathbb{P}(X \in A)$ för "godtyckliga mängder" $A$ i din första kurs i sannolikhetsteori?

[4]Som alternativt skrivsätt för $\int g(x) d\mu_X(x)$ används ofta $\int g(x) dF_X(x)$ där $F_X$ är fördelningsfunktionen för $X$. Speciellt är då $\mathbb{E}[X] = \int x dF_X(x)$.

# 4 PRODUKTRUM

Mått och integraler på produktrum är huvudverktyg i sannolikhetsteorin som vi i detta kapitel skall försöka få åtminstone lite grepp om. Vi skall också se hur produktrum hänger ihop med det i sannolikhetsteorin så centrala begreppet oberoende. I Avsnitt 4.1 behandlar vi endast sannolikhetsrum, medan vi i Avsnitt 4.2 tittar på mer allmänna måttrum.

## 4.1 Produkter av sannolikhetsrum

Vi önskar definiera produkten $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ av två sannolikhetsrum $(\Omega_1, \mathcal{F}_1, \mu_1)$ och $(\Omega_2, \mathcal{F}_2, \mu_2)$. Vi tar en sak i taget:

För två mängder $\Omega_1$ och $\Omega_2$ definieras **produktrummet** $\Omega_1 \times \Omega_2$ som

$$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \, \omega_2 \in \Omega_2\}.$$

Exempelvis är $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, dvs planet. För $\sigma$-algebror $\mathcal{F}_1, \mathcal{F}_2$ på $\Omega_1, \Omega_2$ är **produkt-$\sigma$-algebran**

$$\mathcal{F}_1 \times \mathcal{F}_2 = \sigma\{B_1 \times B_2 : B_1 \in \mathcal{F}_1, B_2 \in \mathcal{F}_2\}$$

För att fullborda sannolikhetsrummet $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ återstår det att definiera $\mu_1 \times \mu_2$. Det sker i följande sats, som vi avstår från att bevisa.

**Sats 4.1** *Låt $(\Omega_1, \mathcal{F}_1, \mu_1)$ och $(\Omega_2, \mathcal{F}_2, \mu_2)$ vara sannolikhetsrum. Då existerar det ett unikt sannolikhetsmått på $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$, kallat* **produktmåttet** *av $\mu_1$ och $\mu_2$ och betecknat $\mu_1 \times \mu_2$, sådant att*

$$(\mu_1 \times \mu_2)(A \times B) = \mu_1(A)\mu_2(B)$$

*för alla $A \in \mathcal{F}_1, B \in \mathcal{F}_2$.*

Ett viktigt verktyg i studiet av produkter av sannolikhetsrum är följande sats, vars bevis vi åter är tvungna att hoppa över.

**Sats 4.2 (Fubinis sats)** *Låt $X$ vara en stokastisk variabel definierad på produktrummet $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$, och låt $\mu$ beteckna $\mu_1 \times \mu_2$. Då gäller*

$$\begin{aligned}\mathbb{E}[|X|] = \int_{\omega \in \Omega_1 \times \Omega_2} |X(\omega)| d\mu(\omega) &= \int_{\omega_1 \in \Omega_1} \left[ \int_{\omega_2 \in \Omega_2} |X(\omega_1, \omega_2)| d\mu_2(\omega_2) \right] d\mu_1(\omega_1) \\ &= \int_{\omega_2 \in \Omega_2} \left[ \int_{\omega_1 \in \Omega_1} |X(\omega_1, \omega_2)| d\mu_1(\omega_1) \right] d\mu_2(\omega_2) .\end{aligned}$$

*Vidare, om $\mathbb{E}[|X|] < \infty$, så är*

$$\begin{aligned}\mathbb{E}[X] = \int_{\omega \in \Omega_1 \times \Omega_2} X(\omega) d\mu(\omega) &= \int_{\omega_1 \in \Omega_1} \left[ \int_{\omega_2 \in \Omega_2} X(\omega_1, \omega_2) d\mu_2(\omega_2) \right] d\mu_1(\omega_1) \\ &= \int_{\omega_2 \in \Omega_2} \left[ \int_{\omega_1 \in \Omega_1} X(\omega_1, \omega_2) d\mu_1(\omega_1) \right] d\mu_2(\omega_2) .\end{aligned}$$

Vi önskar nu ge en definition av begreppet **oberoende** stokastiska variabler. För den sakens skull är det lämpligt att utvidga begreppet fördelning till fallet med två stokastiska variabler. Följande sats är en generalisering av Sats 2.3.

**Sats 4.3** *Låt $X$ och $Y$ vara stokastiska variabler på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Låt*

$$\mu(A) = \mathbb{P}((X, Y) \in A)$$

*för varje $A \in \mathcal{B} \times \mathcal{B}$. Då är $\mu$ ett sannolikhetsmått på $(\mathbb{R}^2, \mathcal{B} \times \mathcal{B})$.*

Satsen bevisas med en enkel modifiering av beviset av Sats 2.3, och vi lämnar detta som Övning 4.1. Måttet $\mu$ kallar vi (den bivariata) fördelningen för $(X, Y)$.

**Definition 4.1** *Låt $X$ och $Y$ vara stokastiska variabler definierade på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$ och med respektive fördelningar $\mu_X$ och $\mu_Y$. $X$ och $Y$ sägs vara* **oberoende** *om fördelningen för $(X, Y)$ på $(\mathbb{R}^2, \mathcal{B} \times \mathcal{B})$ är produktmåttet $\mu_X \times \mu_Y$.*

Två stokastiska variabler $X$ och $Y$ är alltså oberoende om och endast om

$$\mathbb{P}((X, Y) \in A \times B) = \mu_X(A)\mu_Y(B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

för alla $A, B \in \mathcal{B}$. Följande resultat är av central betydelse för hanteringen av oberoende stokastiska variabler.

**Sats 4.4** *Om $X$ och $Y$ är oberoende stokastiska variabler med ändliga väntevärden, så är*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] .$$

För att bevisa satsen krävs följande tvådimensionella motsvarighet till Sats 3.7.

**Lemma 4.1** *Låt $X$ och $Y$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, och låt $g$ vara en mätbar funktion från $(\mathbb{R}^2, \mathcal{B} \times \mathcal{B})$ till $(\mathbb{R}, \mathcal{B})$, sådan att $g(X, Y)$ har ändligt väntevärde. Låt också $\mu$ beteckna den bivariata fördelningen för $(X, Y)$. Då gäller att*

$$\mathbb{E}[g(X,Y)] = \int_{\mathbb{R}^2} g(x,y) d\mu(x,y) \,.$$

**Bevis.** Som i beviset av Sats 3.7; se Övning 4.5. □

**Bevis av Sats 4.4.** Satsen följer av följande beräkning, där första likheten utnyttjar Lemma 4.1, och den andra utnyttjar Fubinis sats:

$$\begin{aligned}
\mathbb{E}[XY] &= \int_{\mathbb{R}^2} xy \, d(\mu_X \times \mu_Y)(x,y) \\
&= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} xy \, d\mu_Y(y) \right] d\mu_X(x) \\
&= \int_{-\infty}^{\infty} x \left[ \int_{-\infty}^{\infty} y \, d\mu_Y(y) \right] d\mu_X(x) \\
&= \int_{-\infty}^{\infty} x \mathbb{E}[Y] \, d\mu_X(x) \\
&= \mathbb{E}[Y] \int_{-\infty}^{\infty} x \, d\mu_X(x) \\
&= \mathbb{E}[X]\mathbb{E}[Y] \,.
\end{aligned}$$

□

Vi bör också ha en definition av vad som menas med oberoende för mer än två stokastiska variabler:

**Definition 4.2** *Vi säger att $k \geq 2$ stokastiska variabler $X_1, \ldots, X_k$ på $(\Omega, \mathcal{F}, \mathbb{P})$, med fördelningar $\mu_1, \ldots, \mu_k$, är* **oberoende** *om fördelningen för $X_1, \ldots, X_k$ på $\mathbb{R}^k$ är lika med produktmåttet $\mu_1 \times \cdots \times \mu_k$, dvs om*

$$\mathbb{P}(X_1 \in A_1, \ldots, X_k \in A_k) = \prod_{i=1}^{k} \mathbb{P}(X_i \in A_i)$$

*för varje val av $A_1, \ldots, A_k \in \mathcal{B}$. Vidare säger vi att en oändlig följd $X_1, X_2, \ldots$ av stokastiska variabler är oberoende om $X_1, \ldots, X_k$ är oberoende för varje $k \geq 2$.*

Även begreppet ”oberoende händelser” bör definieras. Naturligtvis vet du sedan tidagare att två händelser $A$ och $B$ kallas oberoende om $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Utvidgningen till mer än två händelser kan lämpligen göras i termer av oberoende stokastiska variabler:

**Definition 4.3** *En ändlig eller oändlig följd $A_1, A_2, \ldots$ av händelser på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$ sägs vara oberoende, om deras indikatorvariabler $I_{A_1}, I_{A_2}, \ldots$ är oberoende stokastiska variabler.*

**Övningar.**

4.1 Bevisa Sats 4.3.

4.2 **Parvis oberoende.** De stokastiska variablerna $X_1, \ldots, X_k$ sägs vara parvis oberoende om $X_i$ och $X_j$ är oberoende för varje $i \neq j$.

(a) Visa att om $X_1, \ldots, X_k$ är oberoende så är de också parvis oberoende.

(b) Ge ett exempel på tre stokastiska variabler som är parvis oberoende men inte oberoende.

4.3 Visa att om $X$ och $Y$ är oberoende stokastiska variabler, och $f$ och $g$ är mätbara funktioner från $(\mathbb{R}, \mathcal{B})$ till $(\mathbb{R}, \mathcal{B})$, så är $f(X)$ och $g(Y)$ oberoende stokastiska variabler.

4.4 Visa att om $A_1, A_2, \ldots$ är oberoende händelser på $(\Omega, \mathcal{F}, \mathbb{P})$, så är även $A_1^c, A_2^c, \ldots$ oberoende händelser.

4.5 Visa Lemma 4.1. (Ledning: Använd standardmaskinen på samma sätt som i beviset av Sats 3.7.)

## 4.2 Generalisering till $\sigma$-ändliga mått.

Liksom det mesta av integrationsteorin i Kapitel 3, fungerar teorin för produktrum lika bra för $\sigma$-ändliga mått. Exempelvis gäller Sats 4.1 även i fallet då $\mu_1$ och $\mu_2$ är godtyckliga $\sigma$-ändliga mått (produktmåttet blir då naturligtvis $\sigma$-ändligt). Även Fubinis sats kan utvidgas till det $\sigma$-ändliga fallet, och den utvidgningen är så viktig att vi kostar på oss att formulera den separat:

**Sats 4.5 (Fubinis sats, $\sigma$-ändliga fallet)** *Låt $g : \Omega_1 \times \Omega_2 \to \mathbb{R}$ vara en mätbar funktion på måttrummet $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$, och låt $\mu$ beteckna $\mu_1 \times \mu_2$. Antag att $\mu_1$ och $\mu_2$ är $\sigma$-ändliga. Då gäller*

$$\begin{aligned}
\int_{\omega \in \Omega_1 \times \Omega_2} |g(\omega)| d\mu(\omega) &= \int_{\omega_1 \in \Omega_1} \left[ \int_{\omega_2 \in \Omega_2} |g(\omega_1, \omega_2)| d\mu_2(\omega_2) \right] d\mu_1(\omega_1) && (12) \\
&= \int_{\omega_2 \in \Omega_2} \left[ \int_{\omega_1 \in \Omega_1} |g(\omega_1, \omega_2)| d\mu_1(\omega_1) \right] d\mu_2(\omega_2)\,. && (13)
\end{aligned}$$

*Om ovanstående integral är ändlig, så är*

$$\begin{aligned}
\int_{\omega \in \Omega_1 \times \Omega_2} g(\omega) d\mu(\omega) &= \int_{\omega_1 \in \Omega_1} \left[ \int_{\omega_2 \in \Omega_2} g(\omega_1, \omega_2) d\mu_2(\omega_2) \right] d\mu_1(\omega_1) \\
&= \int_{\omega_2 \in \Omega_2} \left[ \int_{\omega_1 \in \Omega_1} g(\omega_1, \omega_2) d\mu_1(\omega_1) \right] d\mu_2(\omega_2)\,.
\end{aligned}$$

Som en tillämpning av denna utvidgade Fubini-sats skall vi visa följande resultat.

**Sats 4.6** *För en ickenegativ stokastisk variabel $X$ med fördelningsfunktion $F_X$, gäller*

$$\mathbb{E}[X] = \int_0^\infty (1 - F_X(y))\, dy\,.$$

**Bevis.** Låt $\mu_X$ beteckna fördelningen för $X$. Vi skall använda oss av Sats 4.5 med $(\Omega_1, \mathcal{F}_1, \mu_1) = ([0, \infty), \mathcal{B}_{[0,\infty)}, \mu_X)$ och $(\Omega_2, \mathcal{F}_2, \mu_2) = ([0, \infty), \mathcal{B}_{[0,\infty)}, m)$, där $m$ som vanligt betecknar Lebesguemåttet. Definiera funktionen $g : [0, \infty) \times [0, \infty) \to \mathbb{R}$ genom att låta $g(x, y) = I_{\{x>y\}}$, och låt $\mu$ beteckna produktmåttet $\mu_X \times m$. Observera att $g(x, y)$ är ickenegativ, så att $g(x, y) = |g(x, y)|$. Ekvation (12) i Sats 4.5 ger

$$\begin{aligned}
\int_{[0,\infty)\times[0,\infty)} g(\omega)\, d\omega &= \int_{[0,\infty)} \left[ \int_0^\infty g(x, y) dy \right] d\mu_X(x) \\
&= \int_{[0,\infty)} \left[ \int_0^x dy \right] d\mu_X(x) \\
&= \int_0^\infty x\, d\mu_X(x) \;=\; \mathbb{E}[X]\,.
\end{aligned}$$

Å andra sidan medför den ”duala” ekvationen (13) att

$$\begin{aligned}
\int_{[0,\infty)\times[0,\infty)} g(\omega)\, d\omega &= \int_0^\infty \left[ \int_{[0,\infty)} g(x, y) d\mu_X(x) \right] dy \\
&= \int_0^\infty \left[ \int_{(y,\infty)} d\mu_X(x) \right] dy \\
&= \int_0^\infty (1 - F_X(y))\, dy\,.
\end{aligned}$$

Kombinerar vi dessa räkningar får vi

$$\mathbb{E}[X] = \int_{[0,\infty)\times[0,\infty)} g(\omega)\, d\omega = \int_0^\infty (1 - F_X(y))\, dy\,.$$

$\square$

Det kanske viktigaste exemplet på $\sigma$-ändliga produktmått är följande generalisering av Lebesguemåttet till högre dimensioner.

**Definition 4.4** *Med* **det tvådimensionella Lebesguemåttet** $m^2$ *på* $(\mathbb{R}^2, \mathcal{B}\times\mathcal{B})$ *menas produktmåttet* $m^2 = m\times m$. *Mer allmännt definierar vi* **det $k$-dimensionella Lebesguemåttet** *på* $(\mathbb{R}^k, \mathcal{B}\times\cdots\times\mathcal{B})$ *som*

$$m^k = m\times\cdots\times m\,,$$

*dvs* $m^k$ *är det (unika) mått på* $(\mathbb{R}^k, \mathcal{B}\times\cdots\times\mathcal{B})$ *som uppfyller*

$$m^k(A_1\times\cdots\times A_k) = \prod_{i=1}^k m(A_i)$$

*för varje val av* $A_1,\ldots,A_k\in\mathcal{B}$.

På samma sätt som det endimensionella Lebesguemåttet $m$ kan tolkas som ett längdmått, kan vi lämpligen tolka de två- och tredimensionella Lebesguemåtten $m^2$ och $m^3$ som area- respektive volymsmått (se Övning 4.7).

En teori för Lebesgueintegration på $\mathbb{R}^k$ kan nu byggas upp på exakt samma sätt som vi gjorde för Lebesgueintegration på $\mathbb{R}$ i Avsnitt 3.2. Eftersom detta strängt taget ligger utanför sannolikhetsteorin, och vi har mycket kvar att göra i kommande kapitel, så ordar vi inte mer om den saken här.

$\square\ \square\ \square\ \square$

Låt oss avsluta detta kapitel med ytterligare några ord om oberoende stokastiska variabler. Vi påminner först om definitionen av begreppet kontinuerlig stokastisk variabel i Kapitel 2 (se (5)), och det förfinade resonemanget i Avsnitt 3.3 som visade att

$$\mathbb{P}(X\in A) = \int_A f_X(x)\, dx$$

för varje $A \in \mathcal{B}$, där $f_X$ är täthetsfunktionen för $X$, och integralen är en Lebesgueintegral. För två stokastiska variabler $X$ och $Y$, definierade på ett sannolikhetsrum, säger vi att paret $(X, Y)$ är ett **kontinuerligt** par av stokastiska variabler, om det existerar en ickenegativ funktion $f : \mathbb{R}^2 \to \mathbb{R}$ (kallad täthetsfunktionen för $(X, Y)$) sådan att

$$\mathbb{P}((X,Y) \in A) = \int_A f(x,y)\, dm^2(x,y)$$

för varje $A \in \mathcal{B} \times \mathcal{B}$ (åter är integralen en Lebesgueintegral). Man kan visa att om $X$ och $Y$ är oberoende stokastiska variabler som var för sig är kontinuerliga med täthetsfunktioner $f_X$ och $f_Y$, så är paret $(X, Y)$ kontinuerligt, med täthetsfunktion $f(x, y) = f_X(x)f_Y(y)$. Om vi nu definierar en ny stokastisk variabel $Z = X + Y$, så kan vi beräkna fördelningsfunktionen för $Z$ enligt följande. Låt $A = \{(x, y) : \ x + y \leq z\}$. Vi får

$$\begin{aligned} F_Z(z) \;=\; \mathbb{P}(Z \leq z) \;&=\; \int_A d(\mu_X \times \mu_Y) \\ &=\; \int_A f_X(x) f_Y(y)\, dm^2(x,y) \\ &=\; \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{z-x} f_X(x) f_Y(y) dy\, dx \\ &=\; \int_{-\infty}^{\infty} f_X(x) F_Y(z-x)\, dx \end{aligned}$$

och om vi deriverar med avseende på $z$ erhåller vi ett uttryck för täthetsfunktionen för $Z$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx\,. \tag{14}$$

Detta är den så kallade **faltningsformeln** för summor av oberoende kontinuerliga stokastiska variabler.

## Övningar.

4.6 Kan du generalisera Sats 4.6 till $\mathbb{E}[X^n]$ för godtyckligt $n \in \mathbb{N}$? Det vill säga: vi vill åstadkomma en identitet av typen

$$\mathbb{E}[X^n] = \int_0^{\infty} (\text{ngt}) \cdot (1 - F_X(x)) dx$$

för ickenegativa stokastiska variabler $X$.

4.7 Låt $A \subset \mathbb{R}^2$ vara en rektangel $[a_1, b_1] \times [a_2, b_2]$. Visa att $m^2(A)$ är lika med arean av $A$. Generalisera sedan resultatet till högre dimensioner.

4.8 Låt $\mathcal{B}^2$ beteckna $\sigma$(de öppna mängderna i $\mathbb{R}^2$). Visa att $\mathcal{B}^2 \subseteq \mathcal{B} \times \mathcal{B}$. (I själva verket gäller $\mathcal{B}^2 = \mathcal{B} \times \mathcal{B}$, men det är svårare att visa. Motsvarande resultat i högre dimensioner gäller också.)

4.9 Definiera Borelmängderna $\mathcal{B}^2$ i $\mathbb{R}^2$ som i föregående övning. Ange klasser av delmängder i $\mathbb{R}^2$, motsvarande dem i Övningarna 1.8 och 1.9, som generar $\mathcal{B}^2$.

4.10 Låt $X$ och $Y$ vara diskreta stokastiska variabler. Visa att $X$ och $Y$ är oberoende om och endast om $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ för alla $x, y \in \mathbb{R}^2$.

4.11 Ge ett exempel på två stokastiska variabler $X$ och $Y$ på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, sådana att $X$ och $Y$ var för sig är kontinuerliga, medan paret $(X, Y)$ inte är kontinuerligt.

# 5 NÅGRA VIKTIGA OLIKHETER

Liksom analysen arbetar sannolikhetsteorin i stor utsträckning med olikheter. I detta kapitel har vi samlat några av de olikheter man oftast stöter på i sannolikhetsteorin. Några känner du förmodligen igen från din första sannolikhetskurs, medan andra troligen är nya för dig. Flera av olikheterna inbegriper de viktiga storheterna **varians** och **kovarians**, som vi således måste börja med att diskutera.

För $p \geq 1$, låt $L^p$ beteckna klassen av stokastiska variabler $X$ sådana att $\mathbb{E}[|X|^p] < \infty$. Om vi vill markera rummet på vilket de stokastiska variablerna är definierade, kan vi skriva $L^p(\Omega, \mathcal{F}, \mathbb{P})$.

**Definition 5.1** *För* $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ *definierar vi* **kovariansen** $\mathrm{Cov}[X, Y]$ *som*

$$\mathrm{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]\,. \tag{15}$$

**Variansen** $\mathrm{Var}[X]$ *definierar vi*

$$\mathrm{Var}[X] = \mathrm{Cov}[X, X] = \mathbb{E}[(X - \mathbb{E}[X])^2]\,.$$

Om vi utvecklar högerledet i (15) får vi den användbara relationen

$$\mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\,.$$

Om $X$ och $Y$ är oberoende så har vi alltså, pga Sats 4.4, $\mathrm{Cov}[X, Y] = 0$. Om vi istället tar $X = Y$ får vi uttrycket

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \tag{16}$$

vilket i vissa läroböcker kallas Steiners sats. Observera också att en varians alltid måste vara ickenegativ, eftersom den är ett väntevärde av en ickenegativ stokastisk variabel (en kvadrat kan ju aldrig bli negativ).

Ett annat mycket användbart samband är

$$\mathrm{Cov}\left[\sum_{i=1}^{k} a_i X_i, \sum_{j=1}^{l} b_j Y_j\right] = \sum_{i=1}^{k}\sum_{j=1}^{l} a_i b_j \mathrm{Cov}[X_i, Y_j] \tag{17}$$

vars bevis vi lämnar som en övning (Övning 5.1). Speciellt blir

$$\mathrm{Var}\left[\sum_{i=1}^{k} a_i X_i\right] = \sum_{i=1}^{k}\sum_{j=1}^{k} a_i a_j \mathrm{Cov}[X_i, X_j]\,. \tag{18}$$

Om vi specialiserar ytterliggare till fallet då $X_1, \ldots, X_k$ är oberoende, så försvinner alla termer i högerledet till (18) utom de på diagonalen ($i = j$), och vi erhåller

$$\mathrm{Var}\left[\sum_{i=1}^k a_i X_i\right] = \sum_{i=1}^k a_i^2 \mathrm{Var}[X_i]\,.$$

Den andra olikheten i följande sats begränsar $\mathrm{Cov}[X, Y]$ för stokastiska variabler $X$ och $Y$ med givna varianser, och är sannolikhetsteorins kanske mest använda olikhet.

**Sats 5.1 (Cauchy–Schwarz' olikhet)** *För $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ gäller*

$$-\sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]} \le \mathbb{E}[XY] \le \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}\,. \tag{19}$$

*Vidare gäller*

$$-\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]} \le \mathrm{Cov}[X, Y] \le \sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}\,. \tag{20}$$

**Bevis.** Sätt $s_X = \sqrt{\mathbb{E}[X^2]}$ och $s_Y = \sqrt{\mathbb{E}[Y^2]}$, och bilda en ny stokastisk variabel $Z = \left(\frac{X}{s_X} - \frac{Y}{s_Y}\right)^2$. Eftersom $Z$ är ickenegativ måste även $\mathbb{E}[Z]$ vara ickenegativ. Direkt räkning ger

$$\begin{aligned}
\mathbb{E}[Z] &= \mathbb{E}\left[\left(\frac{X}{s_X} - \frac{Y}{s_Y}\right)^2\right] \\
&= \frac{\mathbb{E}[X^2]}{s_X^2} + \frac{\mathbb{E}[Y^2]}{s_Y^2} - \frac{2\mathbb{E}[XY]}{s_X s_Y} \\
&= 2\left(1 - \frac{\mathbb{E}[XY]}{s_X s_Y}\right),
\end{aligned}$$

och för att detta uttryck skall bli ickenegativt krävs att $\mathbb{E}[XY] \le s_X s_Y$. Den högra olikheten i (19) är därmed bevisad. Den vänstra olikheten följer av ett liknande resonemang med $Z = \left(\frac{X}{s_X} + \frac{Y}{s_Y}\right)^2$ istället för $\left(\frac{X}{s_X} - \frac{Y}{s_Y}\right)^2$. Slutligen följer (20) genom att tillämpa (19) på $(X - \mathbb{E}[X])$ och $(Y - \mathbb{E}[Y])$. $\square$

**Förstamomentmetoden** och **andramomentmetoden** är två kompletterande olikheter som (trots deras enkelhet) är mycket användbara för att uppskatta sannolikheten att en ickenegativ stokastisk variabel antar stora eller åtminstone strikt positiva värden; exempel på hur användbara de kan vara tillsammans kommer att ges i Avsnitt 8.3. Med $i$:te momentet hos en stokastisk variabel menas helt enkelt $\mathbb{E}[X^i]$. Förstamomentmetoden är också känd som **Markovs olikhet**.

**Sats 5.2 (förstamomentmetoden)** *Låt $X$ vara en ickenegativ stokastisk variabel med väntevärde (förstamoment) $\mathbb{E}[X]$. För varje $a > 0$ gäller att*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}. \tag{21}$$

**Bevis.** Definiera en ny stokastisk variabel $Y = aI_{\{X \geq a\}}$, dvs

$$Y = \begin{cases} a & \text{om } X \geq a \\ 0 & \text{annars.} \end{cases}$$

Vi ser att $Y \leq X$, så pga Sats 3.5 har vi alltså att $\mathbb{E}[Y] \leq \mathbb{E}[X]$. Å andra sidan ser vi att $\mathbb{E}[Y] = a\mathbb{P}(X \geq a)$. Alltså är $a\mathbb{P}(X \geq a) \leq \mathbb{E}[X]$, och (21) följer. □

**Sats 5.3 (andramomentmetoden)** *Låt $X$ vara en ickenegativ stokastisk variabel med väntevärde $\mathbb{E}[X]$ och andramoment $\mathbb{E}[X^2]$. Då gäller*

$$\mathbb{P}(X > 0) \geq \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}. \tag{22}$$

**Bevis.** Tag $Y = I_{\{X>0\}}$. Observera att $X = XY$, och att $Y = Y^2$, så att $\mathbb{E}[Y^2] = \mathbb{E}[Y] = \mathbb{P}(X > 0)$. Cauchy–Schwarz' olikhet ger därför

$$\mathbb{E}[X]^2 = \mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2] = \mathbb{E}[X^2]\mathbb{P}(X > 0),$$

och division med $\mathbb{E}[X^2]$ ger (22). □

För $p \geq 1$ definierar vi den så kallade $L^p$-normen $\|X\|_p$ av en stokastisk variabel $X$, genom

$$\|X\|_p = \mathbb{E}[|X|^p]^{1/p}.$$

I termer av $\|X\|_p$ kan vi alltså uttrycka klassen $L^p$ som

$$L^p = \{X : \|X\|_p < \infty\}.$$

Ännu en olikhet:

**Sats 5.4 (Hölders olikhet)** *För $p > 1$ och $q > 1$ sådana att $p^{-1} + q^{-1} = 1$, och stokastiska variabler $X$ och $Y$, gäller*

$$\|XY\|_1 \leq \|X\|_p \|Y\|_q \,.$$

Observera att för $p = q = 2$ säger Hölders olikhet att

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[|X|^2]\mathbb{E}[|Y|^2]}$$

vilket ju inte är något annat än Cauchy–Schwarz' olikhet för $|X|$ och $|Y|$. För beviset av Hölders olikhet är det lämpligt att börja med följande lemma.

**Lemma 5.1 (aritmetisk-geometriska medelvärdesolikheten)** *För godtyckliga positiva tal $x, y$, och $\lambda \in (0, 1)$, gäller*

$$x^\lambda y^{1-\lambda} \leq \lambda x + (1 - \lambda) y \,.$$

**Bevis.** För $y = 0$ är saken klar. Antag alltså att $y > 0$ och sätt $z = x/y$. Eftersom $r^\lambda \leq 1 + \lambda(r - 1)$ för $r \geq 0$ (eller hur?) får vi $x^\lambda y^{1-\lambda} = yz^\lambda \leq y + \lambda y(z - 1) = \lambda x + (1 - \lambda)y$. □

**Bevis av Sats 5.4.** Om vi tillämpar Lemma 5.1 med $\lambda = \frac{1}{p}$, $x = \left(\frac{|X|}{\|X\|_p}\right)^p$, och $y = \left(\frac{|Y|}{\|Y\|_q}\right)^q$, får vi

$$\frac{|X|}{\|X\|_p} \frac{|Y|}{\|Y\|_q} \leq \lambda \left(\frac{|X|}{\|X\|_p}\right)^p + (1 - \lambda) \left(\frac{|Y|}{\|Y\|_q}\right)^q$$

eftersom $1/q = 1 - 1/p = 1 - \lambda$. Tar vi väntevärden får vi

$$\begin{aligned} \frac{\mathbb{E}[|XY|]}{\|X\|_p \|Y\|_q} &\leq \lambda \frac{\mathbb{E}[|X|^p]}{\|X\|_p^p} + (1 - \lambda) \frac{\mathbb{E}[|Y|^q]}{\|Y\|_q^q} \\ &= \lambda + (1 - \lambda) \;=\; 1 \,, \end{aligned}$$

och multiplikation med $\|X\|_p \|Y\|_q$ ger

$$\mathbb{E}[|XY|] = \|XY\|_1 \leq \|X\|_p \|Y\|_q \,.$$

□

En av huvudpoängerna med att betrakta klassen $L^p = L^p(\Omega, \mathcal{F}, \mathbb{P})$ är att den utgör ett **normerat lineärt rum** med norm $\|\cdot\|_p$. Med detta menas att om $X, Y \in L^p(\Omega, \mathcal{F}, \mathbb{P})$, så gäller

(i) $aX + bY \in L^p$ för konstanter $a, b$,

(ii) $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$,

(iii) $\|aX\|_p = |a| \cdot \|X\|_p$, och

(iv) $\|X\|_p = 0 \;\Leftrightarrow\; X = 0$ (i den meningen att $\mathbb{P}(X \neq 0) = 0$).

Egenskaperna (i), (iii) och (iv) är lätta att visa, och sparas till Övning 5.6. Egenskapen (ii) har ett eget namn:

**Sats 5.5 (Minkowskis olikhet)** *För stokastiska variabler $X, Y \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ gäller att*

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p \,. \tag{23}$$

**Bevis.** Fallet $p = 1$ är självklart. Antag därför att $p > 1$, och tag $q = \frac{p}{p-1}$ så att $p^{-1} + q^{-1} = 1$. Genom att tillämpa Hölders olikhet på variablerna $|X|$ och $|X + Y|^{p-1}$, får vi

$$\begin{aligned}
\mathbb{E}[|X| \cdot |X + Y|^{p-1}] &\leq \mathbb{E}[|X|^p]^{1/p} \mathbb{E}[|X + Y|^{(p-1)(\frac{p}{p-1})}]^{\frac{p-1}{p}} \\
&= \|X\|_p \mathbb{E}[|X + Y|^p]^{\frac{p-1}{p}} \\
&= \|X\|_p \|X + Y\|_p^{p-1} \,,
\end{aligned}$$

och på samma sätt får vi att $\mathbb{E}[|Y| \cdot |X + Y|^{p-1}] \leq \|Y\|_p \|X + Y\|_p^{p-1}$. Alltså har vi

$$\begin{aligned}
\|X + Y\|_p^p &= \mathbb{E}[|X + Y|^p] \\
&= \mathbb{E}[|X + Y| \cdot |X + Y|^{p-1}] \\
&\leq \mathbb{E}[|X| \cdot |X + Y|^{p-1}] + \mathbb{E}[|Y| \cdot |X + Y|^{p-1}] \\
&\leq \|X\|_p \cdot \|X + Y\|_p^{p-1} + \|Y\|_p \cdot \|X + Y\|_p^{p-1}
\end{aligned}$$

och division med $\|X + Y\|_p^{p-1}$ ger (23). □

## Övningar.

5.1 Visa att (17) gäller för stokastiska variabler $X_1, \ldots, X_k, Y_1, \ldots, Y_l$ och reella tal $a_1, \ldots, a_k, b_1, \ldots, b_l$. Detta egenskap hos kovariansen kallas **bilinearitet**: den är lineär i båda argumenten.

5.2 **Cauchy–Schwarz' olikhet är skarp.** Visa att det, för varje val av $\sigma_X > 0$, $\sigma_Y > 0$ och $c$ sådant att $|c| \leq \sigma_X \sigma_Y$, existerar stokastiska variabler $X$ och $Y$ sådana att $\mathrm{Var}[X] = \sigma_X^2$, $\mathrm{Var}[Y] = \sigma_Y^2$ och $\mathrm{Cov}[X, Y] = c$. (Ledning: Låt $X'$ och $Y'$ vara oberoende stokastiska variabler med varians 1, och låt sedan $X = a_1 X'$, och $Y = b_1 X' + b_2 Y'$, där konstanterna $a_1$, $b_1$ och $b_2$ väljs på lämpligt sätt.)

5.3 **Tjebysjevs olikhet.** Låt $X$ vara en stokastisk variabel med väntevärde $\mathbb{E}[X]$ och varians $\mathrm{Var}[X]$. Visa att

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}[X]}{a^2}$$

för varje $a > 0$.

5.4 **Jensens olikhet.** En funktion $g : \mathbb{R} \to \mathbb{R}$ sägs vara **konvex** om det för varje $a \in \mathbb{R}$ existerar ett $\lambda \in \mathbb{R}$ sådant att $g(x) \geq g(a) + \lambda(x - a)$ (det kan vara bra att rita en figur som illustrerar denna definition). Visa att om $X$ är en stokastisk variabel med ändligt väntevärde, och $g$ är en konvex funktion, så är

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]) \, .$$

5.5 Visa på två sätt, ett med hjälp av Hölders olikhet och ett med Jensens, att $\|X\|_p$ är växande i $p$ för en stokastisk variabel $X$.

5.6 Visa att $L^p(\Omega, \mathcal{F}, \mathbb{P})$ uppfyller kraven (i), (iii) och (iv) på ett normerat lineärt rum.

# 6 KONVERGENS

## 6.1 Konvergens för följder av stokastiska variabler

Att tänka sig en lärobok i matematisk analys som inte behandlar gränsvärdesbegreppet ter sig tämligen bisarrt. På liknande sätt förhåller det sig med sannolikhetsteorin: En text på vår nivå vore ofullständig utan en ordentlig genomgång av konvergens för stokastiska variabler. Om $X_1, X_2, \ldots$ och $X$ är definierade på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, vad menar vi då med att $X_n \to X$ då $n \to \infty$?

Denna enkla fråga har ett mer komplicerat svar än vad man först skulle kunna tro. I Avsnitt 3.1 arbetade vi med **punktvis konvergens**, med vilket vi menade att $X_n(\omega) \to X(\omega)$ för varje $\omega \in \Omega$. Det finns emellertid (minst) fyra andra viktiga konvergensbegrepp, som alla används flitigt i modern sannolikhetsteori, nämligen:

- nästan säker konvergens,
- $L^p$-konvergens,
- konvergens i sannolikhet, och
- konvergens i fördelning.

Vi börjar med definitionerna:

**Definition 6.1** *Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. Om*

$$\mathbb{P}(\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)) = 1$$

*så säger vi att $X_n$ konvergerar mot $X$* **nästan säkert**, *och skriver $X_n \stackrel{n.s.}{\to} X$.*

**Definition 6.2** *Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, och låt även $p \geq 1$. Om*

$$\lim_{n \to \infty} \|X_n - X\|_p = 0$$

*(dvs om $\mathbb{E}[|X_n - X|^p] \to 0$ då $n \to \infty$), så säger vi att $X_n$ konvergerar mot $X$* **i $L^p$-mening** *(eller i $L^p$).*

**Definition 6.3** *Låt* $X_1, X_2, \ldots$ *och* $X$ *vara stokastiska variabler på* $(\Omega, \mathcal{F}, \mathbb{P})$. *Om det för varje* $\varepsilon > 0$ *gäller att*

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0\,,$$

*så säger vi att* $X_n$ *konvergerar mot* $X$ **i sannolikhet**, *och skriver* $X_n \xrightarrow{\mathcal{P}} X$.

**Definition 6.4** *Låt* $X_1, X_2, \ldots$ *och* $X$ *vara stokastiska variabler med fördelningsfunktioner* $F_{X_1}, F_{X_2}, \ldots$ *och* $F_X$. *Om*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

*för alla* $x \in \mathbb{R}$ *sådana att* $F_X$ *är kontinuerlig i* $x$, *så säger vi att* $X_n$ *konvergerar mot* $X$ **i fördelning**, *och skriver* $X_n \xrightarrow{\mathcal{D}} X$. *Vi tillåter oss också att skriva* $X_n \xrightarrow{\mathcal{D}} F_X$ *och* $F_{X_n} \xrightarrow{\mathcal{D}} F_X$.

Dessa fyra konvergensbegrepp är *inte* ekvivalenta. Vissa samband föreligger dock, som vi skall se i Sats 6.1. Innan dess skall vi först göra några allmänna kommentarer.

1. Nästan säker konvergens tillåter att det finns $\omega \in \Omega$ sådana att
$$X_n(\omega) \not\to X(\omega)\,, \tag{24}$$
under förutsättning att mängden $A = \{\omega : (24) \text{ gäller}\}$ har sannolikhet $\mathbb{P}(A) = 0$. Nästan säker konvergens är alltså svagare än punktvis konvergens, som ju kräver att $X_n(\omega) \to X(\omega)$ för *alla* $\omega \in \Omega$. För att markera kontrasten mot nästan säker konvergens brukar punktvis konvergens också kallas **säker konvergens**. För en sannolikhetsteoretiker är det naturligare att arbeta med nästan säker konvergens än med säker konvergens, eftersom det senare ju inte alls beror på sannolikhetsmåttet $\mathbb{P}$. Om du skärskådar argumenten i Avsnitt 3.1, kommer du att finna att exempelvis Monotona konvergenssatsen och Dominerade konvergenssatsen gäller även då säker konvergens ersatts med det svagare villkoret nästan säker konvergens.

2. Vad beträffar $L^p$-konvergens så är fallen $p = 1$ och $p = 2$ de överlägset mest studerade. $L^1$-konvergens kallas ofta **konvergens i medel**, medan $L^2$-konvergens brukar kallas **konvergens i kvadratiskt medel**.

3. Notera att konvergens i fördelning skiljer sig från de tre övriga konvergensbegreppen på så sätt att det inte säger något om sannolikhetsrummet som de stokastiska variablerna är definierade på, utan bara handlar om deras fördelningar. Det kan till och med vara så att variablerna är definierade på olika sannolikhetsrum. Föga förvånande skall vi strax se att konvergens i fördelning är det svagaste av de fyra konvergensbegreppen.

4. Men kan visa att konvergens i fördelning är ekvivalent med att

$$\lim_{n\to\infty} \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)]$$

för alla begränsade kontinuerliga funktioner $g : \mathbb{R} \to \mathbb{R}$. I mer avancerade texter brukar detta rentav vara definitionen av konvergens i fördelning; vår Definition 6.4 blir då en sats.

Följande samband mellan de fyra typerna av konvergens gäller.

**Sats 6.1** *Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Då gäller:*

(i) *Om $X_n \xrightarrow{\mathcal{P}} X$, så $X_n \xrightarrow{\mathcal{D}} X$.*

(ii) *Om $X_n \xrightarrow{n.s.} X$, så $X_n \xrightarrow{\mathcal{P}} X$.*

(iii) *Om $p_1, p_2 \in \mathbb{R}$ är sådana att $1 \leq p_1 \leq p_2$, och $X_n \to X$ i $L^{p_2}$, så gäller även $X_n \to X$ i $L^{p_1}$.*

(iv) *Om $X_n \to X$ i $L^p$ för något $p \geq 1$, så gäller även $X_n \xrightarrow{\mathcal{P}} X$.*

Efter att vi har bevisat satsen skall vi se ett antal exempel som visar att inga andra implikationer föreligger. Kanske kan satsen bäst sammanfattas med följande diagram, där $1 \leq p_1 \leq p_2$:

$$\begin{array}{ccccccc}
 & & & & (X_n \xrightarrow{n.s.} X) & & \\
 & & & & \Downarrow & & \\
(X_n \to X \text{ i } L^{p_2}) & \Rightarrow & (X_n \to X \text{ i } L^{p_1}) & \Rightarrow & (X_n \xrightarrow{\mathcal{P}} X) & \Rightarrow & (X_n \xrightarrow{\mathcal{D}} X)
\end{array}$$

**Bevis av Sats 6.1 (i).** Antag att $X_n \xrightarrow{\mathcal{P}} X$, och tag $\varepsilon > 0$ och $x \in \mathbb{R}$ sådant att $F_X(x)$ är kontinuerlig i punkten $x$. För givet $n$ har vi

$$\begin{aligned}
F_{X_n}(x) &= \mathbb{P}(X_n \leq x) \\
&= \mathbb{P}(X_n \leq x, X \leq x+\varepsilon) + \mathbb{P}(X_n \leq x, X > x+\varepsilon) \\
&\leq F_X(x+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon)\,,
\end{aligned}$$

och på samma sätt får vi att

$$\begin{aligned}
F_X(x-\varepsilon) &= \mathbb{P}(X \leq x-\varepsilon) \\
&= \mathbb{P}(X \leq x-\varepsilon, X_n \leq x) + \mathbb{P}(X \leq x-\varepsilon, X_n > x) \\
&\leq F_{X_n}(x) + \mathbb{P}(|X_n - X| > \varepsilon)\,.
\end{aligned}$$

Vi har alltså att

$$F_X(x-\varepsilon) - \mathbb{P}(|X_n - X| > \varepsilon) \le F_{X_n}(x) \le F_X(x+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon)\,.$$

Om vi låter $n \to \infty$ så går $\mathbb{P}(|X_n - X| > \varepsilon)$ mot 0, och vi får

$$F_X(x-\varepsilon) \le \liminf_{n\to\infty} F_{X_n}(x) \le \limsup_{n\to\infty} F_{X_n}(x) \le F_X(x+\varepsilon)\,.$$

Nu kan vi låta $\varepsilon \to 0$. Tack vare att $F_X$ är kontinuerlig i $x$ så har vi att $F_X(x-\varepsilon)$ och $F_X(x+\varepsilon)$ båda går mot $F_X(x)$. Alltså har vi

$$F_X(x) \le \liminf_{n\to\infty} F_{X_n}(x) \le \limsup_{n\to\infty} F_{X_n}(x) \le F_X(x)\,.$$

Detta är samma sak som att

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

vilket ju var definitionen av konvergens i fördelning. $\square$

**Bevis av Sats 6.1 (ii).** Antag att $X_n \overset{n.s.}{\to} X$, och fixera $\delta, \varepsilon > 0$; vi behöver visa att det finns ett $N < \infty$ sådant att

$$\mathbb{P}(|X_n - X| > \varepsilon) \le \delta \quad \text{för alla } n \ge N\,. \tag{25}$$

Definiera den stokastiska variableln $Y$ som

$$Y = \inf\{m \in \mathbb{N} : \, |X_n - X| \le \varepsilon \text{ för alla } n \ge m\}\,.$$

Tack vare att $X_n \overset{n.s.}{\to} X$ så är $Y < \infty$ med sannolikhet 1. Vi kan därför finna ett $N$ sådant att $\mathbb{P}(Y > N) \le \delta$. Vi har då

$$\mathbb{P}(|X_n - X| \le \varepsilon \text{ för alla } n \ge N) \ge 1 - \delta$$

vilket är samma sak som att

$$\mathbb{P}(|X_n - X| > \varepsilon \text{ för något } n \ge N) \le \delta\,,$$

vilket i sin tur implicerar (25). $\square$

**Bevis av Sats 6.1 (iii).** Antag att $X_n \to X$ i $L^{p_2}$, dvs att

$$\lim_{n\to\infty} \|X_n - X\|_{p_2} = 0\,. \tag{26}$$

Tack vare Övning 5.5 vet vi att $\|X_n - X\|_{p_1} \leq \|X_n - X\|_{p_2}$, vilket i kombination med (26) implicerar att

$$\lim_{n\to\infty} \|X_n - X\|_{p_1} = 0\,,$$

så att $X_n \to X$ i $L^{p_2}$. $\square$

**Bevis av Sats 6.1 (iv).** Antag att $X_n \to X$ i $L^p$, för något $p \geq 1$. Med stöd av Sats 6.1 (iii) har vi då att $X_n \to X$ i $L^1$, dvs att

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X|] = 0\,.$$

Fixera $\varepsilon > 0$. Om vi tillämpar Sats 5.2 på den stokastiska variabeln $|X_n - X|$ får vi att

$$\mathbb{P}(|X_n - X| \geq \varepsilon) \leq \frac{\mathbb{E}[|X_n - X|]}{\varepsilon}\,.$$

Alltså har vi

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0$$

och eftersom $\varepsilon > 0$ var godtyckligt har vi att $X_n \xrightarrow{\mathcal{P}} X$. $\square$

Vi skall nu ge motexempel som visar att de implikationer mellan de olika konvergensbegreppen som anges i Sats 6.1 är de enda som gäller generellt. I samtliga fall antas $X_1, X_2, \ldots$ och $X$ vara definierade på ett och samma sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$.

**Exempel 6.1: Konvergens i fördelning medför inte konvergens i sannolikhet.** Låt $X_1, X_2, \ldots$ vara oberoende likafördelade stokastiska variabler med $\mathbb{P}(X_n = 0) = \mathbb{P}(X_n = 1) = \frac{1}{2}$. Låt $X$ vara oberoende av dessa och ha samma fördelning. Eftersom $X_1, X_2, \ldots$ och $X$ har samma fördelning, så gäller självklart att $X_n \xrightarrow{\mathcal{D}} X$. Å andra sidan är $\mathbb{P}(|X_n - X| \geq 1) = \frac{1}{2}$ för varje $n$, så det gäller alltså *inte* att $X_n \xrightarrow{\mathcal{P}} X$.

**Exempel 6.2: Varken konvergens i sannolikhet eller $L^p$-konvergens medför nästan säker konvergens.** Låt $X = 0$, dvs låt $X(\omega) = 0$ för varje $\omega \in \Omega$. Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler med

$$X_n = \begin{cases} 1 & \text{med sannolikhet } \frac{1}{n} \\ 0 & \text{med sannolikhet } 1 - \frac{1}{n}\,. \end{cases}$$

För $p \geq 1$ får vi

$$\|X_n - X\|_p = \mathbb{E}[|X_n - X|^p]^{1/p} = \left(\frac{1}{n}\right)^{1/p}$$

vilket går mot 0 då $n \to \infty$, så att $X_n$ alltså konvergerar mot $X$ i $L^p$. Tack vare Sats 6.1 (iv) har vi även att $X_n$ går mot $X$ i sannolikhet.

Vi skall nu visa att $X_n$ *inte* går mot $X$ nästan säkert. Antag, i syfte att erhålla en motsägelse, att $X_n$ går mot $X$ nästan säkert, dvs att $\mathbb{P}(\lim_{n\to\infty} X_n = 0) = 1$. Eftersom variablerna $X_1, X_2, \ldots$ är 0-1-värda innebär det att det måste finnas något (stokastiskt) $Y$ sådant att $\mathbb{P}(X_n = 0 \text{ för alla } n \geq Y) = 1$. Vi kan då hitta ett tal $N$ sådant att $\mathbb{P}(Y \leq N) > 0$, så att

$$\mathbb{P}(X_n = 0 \text{ för alla } n \geq N) > 0 \,. \tag{27}$$

Men

$$\begin{aligned}
\mathbb{P}(X_n = 0 \text{ för alla } n \geq N) &= \lim_{k\to\infty} \mathbb{P}(X_N = 0, X_{N+1} = 0, \ldots, X_{N+k} = 0) \\
&= \lim_{k\to\infty} \prod_{i=N}^{N+k} \mathbb{P}(X_i = 0) \;=\; \lim_{k\to\infty} \prod_{i=N}^{N+k} \frac{i-1}{i} \\
&= \lim_{k\to\infty} \frac{N-1}{N+k} \;=\; 0 \,,
\end{aligned}$$

vilket motsäger (27). Alltså måste vårt antagande, att $X_n$ går mot $X$ nästan säkert, vara felaktigt.

**Exempel 6.3: Det är skillnad mellan $L^{p_1}$- och $L^{p_2}$-konvergens.** Fixera $q > 1$. Låt $X = 0$ och låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler med

$$X_n = \begin{cases} n^{1/q} & \text{med sannolikhet } \frac{1}{n} \\ 0 & \text{med sannolikhet } 1 - \frac{1}{n} \,. \end{cases}$$

För $p \geq 1$ får vi

$$\|X_n - X\|_p = \mathbb{E}[|X_n - X|^p]^{1/p} = (n^{\frac{p}{q}-1})^{1/p}$$

vilket går mot 0 om och endast om $p < q$. Alltså konvergerar $X_n$ mot $X$ i $L^p$-mening om och endast om $p < q$.

**Exempel 6.4: Nästan säker konvergens medför inte $L^1$-konvergens.** Låt $X = 0$ och låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler med

$$X_n = \begin{cases} n^2 & \text{med sannolikhet } \frac{1}{n^2} \\ 0 & \text{med sannolikhet } 1 - \frac{1}{n^2} \,. \end{cases}$$

Låt

$$Y = I_{\{X_1 \neq 0\}} + I_{\{X_2 \neq 0\}} + \cdots$$

och notera att

$$\mathbb{E}[Y] = \sum_{n=1}^{\infty} \mathbb{P}(X_n \neq 0) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

vilket medför att $Y < \infty$ med sannolikhet 1. Detta betyder att

$$\mathbb{P}(\exists N \text{ sådant att } X_n = 0 \text{ för alla } n \geq N) = 1 \,.$$

Alltså konvergerar $X_n$ mot 0 (dvs mot $X$) nästan säkert. Å andra sidan är

$$\|X_n - X\|_1 = \mathbb{E}[|X_n - X|] = \frac{n^2}{n^2} = 1$$

för varje $n$, så att $\|X_n - X\|_1$ *inte* går mot 0 då $n \to \infty$, dvs $X_n$ går *inte* mot $X$ i $L^1$-mening.

Dessa motexempel visar att en mängd påståenden av typen ”om $X_n$ konvergerar mot $X$ i meningen A, så konvergerar $X_n$ mot $X$ även i meningen $B$” inte äger någon generell giltighet. Å andra sidan kan det tänkas att några av dessa påståenden är sanna om man gör ytterligare antaganden om $X_n$ och $X$. Vi avslutar detta avsnitt med två sådana resultat.

**Sats 6.2** *Låt $X_1, X_2, \ldots$ vara stokastiska variabler på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och antag att $X_n \xrightarrow{\mathcal{D}} K$, där $K$ är en (deterministisk) konstant. Då gäller även att $X_n \xrightarrow{\mathcal{P}} K$.*

**Bevis.** Den stokastiska variabel som är konstant lika med $K$, har fördelningsfunktionen

$$F_K(x) = \begin{cases} 0 & \text{för } x < K \\ 1 & \text{för } x \geq K \end{cases}$$

och antagandet att $X_n \xrightarrow{\mathcal{D}} K$ innebär att $\lim_{n\to\infty} F_{X_n}(x) = F_K(x)$ för alla $x \in \mathbb{R} \setminus \{K\}$. För $\varepsilon > 0$ har vi

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(|X_n - K| > \varepsilon) &= \lim_{n\to\infty} (\mathbb{P}(X_n < K - \varepsilon) + \mathbb{P}(X_n > K + \varepsilon)) \\
&\leq \lim_{n\to\infty} (\mathbb{P}(X_n \leq K - \varepsilon) + \mathbb{P}(X_n > K + \varepsilon)) \\
&= \lim_{n\to\infty} (F_{X_n}(K - \varepsilon) + 1 - F_{X_n}(K + \varepsilon)) \\
&= 0 + 1 - 1 \;=\; 0.
\end{aligned}$$

Eftersom $\varepsilon$ var godtyckligt har vi därmed viast att $X_n \xrightarrow{\mathcal{P}} K$. □

Nästa sats är besläktad med Begränsade konvergenssatsen (Korollarium 3.1).

**Sats 6.3** *Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och antag att det finns en konstant $k$ sådan att $\mathbb{P}(|X_n| \leq k) = 1$ för alla $n$. Om $X_n \xrightarrow{\mathcal{P}} X$, så gäller även att $X_n \to X$ i $L^p$ för varje $p \geq 1$.*

**Bevis.** Observera först, att om $\mathbb{P}(|X_n| \leq k) = 1$ och $X_n \xrightarrow{\mathcal{P}} X$, så gäller även att $\mathbb{P}(|X| \leq k) = 1$ (se Övning 6.1 (a)). Detta betyder speciellt att $\mathbb{P}(|X_n - X| \leq 2k) = 1$ för alla $n$. Fixera $p \geq 1$ och $\varepsilon > 0$. Vi uppskattar $L^p$-normen för $|X_n - X|$ på följande sätt:

$$\begin{aligned}
\|X_n - X\|_p &= \mathbb{E}[|X_n - X|^p]^{1/p} \\
&= (\mathbb{E}[|X_n - X|^p I_{\{|X_n - X| \leq \varepsilon\}}] + \mathbb{E}[|X_n - X|^p I_{\{|X_n - X| > \varepsilon\}}])^{1/p} \\
&\leq (\varepsilon^p + (2k)^p \mathbb{P}(|X_n - X| > \varepsilon))^{1/p}.
\end{aligned}$$

Om vi nu låter $n \to \infty$, så går $\mathbb{P}(|X_n - X| > \varepsilon)$ mot 0, och vi erhåller

$$\lim_{n\to\infty} \|X_n - X\|_p \leq (\varepsilon^p)^{1/p} = \varepsilon$$

och eftersom $\varepsilon > 0$ var godtyckligt har vi

$$\lim_{n\to\infty} \|X_n - X\|_p = 0$$

vilket är samma sak som att $X_n \to X$ i $L^p$. $\square$

**Övningar.**

6.1 Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler sådana att $X_n \xrightarrow{\mathcal{P}} X$.

(a) Antag att $\mathbb{P}(X_n \in [a,b]) = 1$ för alla $n$ och något fixt slutet intervall $[a,b]$, och visa att $\mathbb{P}(X \in [a,b]) = 1$.

(b) Antag istället att $\mathbb{P}(X_n \in (a,b)) = 1$ för alla $n$ och något fixt öppet intervall $(a,b)$. Ge ett motexempel som visar att detta *inte* implicerar att $\mathbb{P}(X \in (a,b)) = 1$.

6.2 Låt $X_1, X_2, \ldots$ och $Y_1, Y_2, \ldots$, samt $X$ och $Y$, vara stokastiska variabler på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och låt $p \geq 1$. Visa att

(a) om $X_n \xrightarrow{n.s.} X$ och $Y_n \xrightarrow{n.s.} Y$, så gäller även $X_n + Y_n \xrightarrow{n.s.} X + Y$,

(b) om $X_n \to X$ i $L^p$, och $Y_n \to Y$ i $L^p$, så gäller även att $X_n + Y_n \to X + Y$ i $L^p$,

(c) om $X_n \xrightarrow{\mathcal{P}} X$ och $Y_n \xrightarrow{\mathcal{P}} Y$, så gäller även $X_n + Y_n \xrightarrow{\mathcal{P}} X + Y$.

Visa också medelst ett motexempel att $X_n \xrightarrow{\mathcal{D}} X$ och $Y_n \xrightarrow{\mathcal{D}} Y$ *inte* är tillräckliga villkor för att $X_n + Y_n \xrightarrow{\mathcal{D}} X + Y$ skall gälla.

6.3 **Dominerad konvergens.** Bevisa följande variant av Sats 6.3. Låt $X_1, X_2, \ldots$ samt $X$ och $Y$, vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$ och låt $p \geq 1$. Antag att

(i)[5] $X_n \xrightarrow{n.s.} X$,

(ii) $|X_n(\omega)| \leq Y(\omega)$ för alla $\omega \in \Omega$ och alla $n$, och

(iii) $Y \in L^p$.

Då gäller också att $X_n \to X$ i $L^p$.

## 6.2 Borel–Cantelli och Kolmogorovs 0-1-lag

I detta avsnitt tar vi upp ett antal ytterligare resultat som är användbara för att hantera följder av stokastiska variabler. Till de mer berömda sådana resultaten hör Borel–Cantellis båda lemmata, som vi börjar med.

[5]Detta villkor kan även ersättas med det svagare villkoret att $X_n \xrightarrow{\mathcal{P}} X$, men då blir beviset mer krävande.

Låt $A_1, A_2, \ldots \in \mathcal{F}$ vara en oändlig sekvens av händelser på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Vi påminner oss, från Övning 1.10, definitionen

$$\begin{aligned}\limsup_{n\to\infty} A_n &= \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m \\ &= \{\omega \in \Omega : \omega \in A_n \text{ för oändligt många } n\}\,.\end{aligned}$$

**Lemma 6.1 (Borel–Cantellis första lemma)** *Låt $A_1, A_2, \ldots \in \mathcal{F}$ vara händelser på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och antag att $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$. Då är*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 0\,,$$

*dvs $\mathbb{P}$(högst ändligt många $A_n$ inträffar) $= 1$.*

**Bevis.** Vi använder samma argument som i Exempel 6.4. Låt $Y$ vara antalet händelser $A_n$ som inträffar, dvs $Y = I_{A_1} + I_{A_2} + \cdots$. Vi har

$$\mathbb{E}[Y] = \sum_{n=1}^{\infty} \mathbb{E}[I_{A_n}] = \sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$$

så att[6] $Y < \infty$ med sannolikhet 1. □

Nästa exempel visar att omvändningen till Borel–Cantellis första lemma inte är sann, dvs $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$ medför inte att $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$.

**Exempel 6.5.** Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara som i Exempel 1.5, dvs $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ och $\mathbb{P}$ är Lebesguemåttet begränsat till $[0,1]$. Om vi nu låter $A_n = (0, 1/n)$, så gäller för varje $\omega \in \Omega$ att $\omega \notin A_n$ för alla tillräckligt stora $n$, så att

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = \mathbb{P}(\emptyset) = 0\,.$$

Å andra sidan har vi

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty\,.$$

Om vi däremot lägger på villkoret att $A_1, A_2, \ldots$ är oberoende händelser, så gäller den efterfrågade omvändningen:

[6] Mot detta bevis kan man (med rätta) invända att vi inte har infört någon teori för stokastiska variabler som kan anta värdet $\infty$. Emellertid kan teorin för väntevärden av positiva stokastiska variabler i Avsnitt 3.1 utan några ytterligare komplikationer utvidgas till stokastiska variabler som antar värden i $[0,\infty) \cup \{\infty\}$. En sådan stokastisk variabel $X$ med $\mathbb{P}(X = \infty) > 0$ kommer (vilket är lätt att verifiera) då alltid att ha väntevärde $\infty$.

**Lemma 6.2 (Borel–Cantellis andra lemma)** *Antag att $A_1, A_2, \ldots \in \mathcal{F}$ är oberoende händelser på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, och att $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$. Då är*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 1\,, \tag{28}$$

*dvs $\mathbb{P}($oändligt många $A_n$ inträffar$) = 1$.*

**Bevis.** Antag (i syfte att få en motsägelse) att (28) inte gäller, dvs att

$$\mathbb{P}(\text{högst ändligt många } A_n \text{ inträffar}) > 0\,. \tag{29}$$

I så fall finns det något $N < \infty$ sådant att

$$\mathbb{P}(\bigcap_{i=N}^{\infty} A_i^c) > 0\,.$$

Notera vidare att $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ implicerar $\sum_{n=N}^{\infty} \mathbb{P}(A_n) = \infty$, dvs

$$\lim_{k\to\infty} \sum_{n=N}^{N+k} \mathbb{P}(A_n) = \infty\,.$$

Eftersom $A_1, A_2, \ldots$ är oberoende, är även $A_1^c, A_2^c, \ldots$ oberoende (se Övning 4.4), och vi erhåller

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{n=N}^{\infty} A_n^c\right) &= \lim_{k\to\infty} \mathbb{P}\left(\bigcap_{n=N}^{N+k} A_n^c\right) \\
&= \lim_{k\to\infty} \prod_{n=N}^{N+k} \mathbb{P}(A_n^c) = \lim_{k\to\infty} \prod_{n=N}^{N+k} (1 - \mathbb{P}(A_n)) \\
&\leq \lim_{k\to\infty} \prod_{n=N}^{N+k} \exp(-\mathbb{P}(A_n)) \\
&= \lim_{k\to\infty} \exp\left(-\sum_{n=N}^{N+k} \mathbb{P}(A_n)\right) = 0
\end{aligned}$$

(olikheten följer av att $1 + x \leq e^x$). Detta motsäger antagandet (29), som alltså måste vara felaktigt. Således gäller (28). □

□ □ □ □

Vi skall nu presentera Kolmogorovs 0-1-lag, som är den mest kända av de så kallade 0-1-lagarna i sannolikhetsteorin. 0-1-lagar är satser som säger att sannolikheten för en viss typ av händelser måste vara (eller konvergera mot) 0 eller 1. För att kunna formulera Kolmogorovs 0-1-lag måste vi först utvidga vår terminologi beträffande $\sigma$-algebror.

**Definition 6.5** *Låt $X_1, \ldots, X_n$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. Vi definierar $\sigma(X_1, \ldots, X_n)$ genom*

$$\sigma(X_1, \ldots, X_n) = \sigma(\{\{X_i \in A\} : i \in \{1, \ldots, n\}, A \in \mathcal{B}\})$$

*och kallar den för* **den av de stokastiska variablerna $X_1, \ldots, X_n$ genererade $\sigma$-algebran**. *Om istället $X_1, X_2, \ldots$ är en oändlig följd av stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, så definierar vi $\sigma(X_1, X_2, \ldots)$ som*

$$\sigma(X_1, X_2, \ldots) = \sigma(\{\{X_i \in A\} : i \in \mathbb{N}, A \in \mathcal{B}\}) .$$

*Vi använder även beteckningarna*

$$\mathcal{F}_i^j = \sigma(X_i, X_{i+1}, \ldots, X_j)$$

*och*

$$\mathcal{F}_i^\infty = \sigma(X_i, X_{i+1}, \ldots) .$$

Vi har att $\mathcal{F}_1^\infty \subseteq \mathcal{F}$ (se Övning 6.7). Intuitivt bör man tänka på $\mathcal{F}_1^\infty = \sigma(X_1, X_2, \ldots)$ som klassen av händelser med egenskapen att vi kan avgöra om de inträffat om vi observerat variablerna $X_1, X_2, \ldots$. Exempel på händelser i $\mathcal{F}_1^\infty$ är $A_1 = \{X_5 = 2\}$, $A_2 = \{X_n > 0 \text{ för något } n \in \mathbb{N}\}$ och $A_3 = \{\liminf_{n \to \infty} X_n < 1\}$.

**Definition 6.6** *Låt $X_1, X_2, \ldots$ vara en oändlig följd av stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. Vi definierar $\sigma$-algebran $\mathcal{T} = \mathcal{T}(X_1, X_2, \ldots)$ som*

$$\mathcal{T}(X_1, X_2, \ldots) = \bigcap_{n=1}^{\infty} \mathcal{F}_n^\infty = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots)$$

*och kallar $\mathcal{T}$ för* **den av $(X_1, X_2, \ldots)$ genererade svans-$\sigma$-algebran**.

(Man måste naturligtvis verifiera att $\mathcal{T}$ verkligen är en $\sigma$-algebra. Detta lämnar vi som Övning 6.8.)

Svans-$\sigma$-algebran $\mathcal{T} = \mathcal{T}(X_1, X_2, \ldots)$ består alltså av de händelser för vilka vi kan avgöra om de inträffat genom att endast betrakta $X_i$-variabler "för godtyckligt stora $i$". Vi första påseende kan det tyckas att $\mathcal{T}$ därmed inte kan innehålla några andra händelser än de triviala händelserna $\Omega$ och $\emptyset$. Så är emellertid inte fallet – $\mathcal{T}$ kan innehålla en mängd intressanta händelser. Bland händelserna $A_1$, $A_2$ och $A_3$ ovan tillhör $A_3$ (men däremot inte $A_1$ och $A_2$) $\mathcal{T}$. Andra exempel på händelser i $\mathcal{T}$ är $A_4 = \{X_n > 0 \text{ för oändligt många } n\}$, $A_5 = \{\lim_{n \to \infty} X_n \text{ existerar}\}$, och $A_6 = \{\lim_{n \to \infty}(X_1 + \cdots + X_n)/n = 1\}$.

Vi är nu redo att formulera Kolmogorovs 0-1-lag som säger att om $X_1, X_2, \ldots$ är *oberoende* händelser, så är $\mathbb{P}(A)$ lika med 0 eller 1 för varje $A \in \mathcal{T}$. Detta föranleder många författare att kalla svans-$\sigma$-algebran $\mathcal{T}$ för "trivial", men vi vill betona att den är långt mindre trivial än den verkligt triviala $\sigma$-algebran $\{\emptyset, \Omega\}$.

**Sats 6.4 (Kolmogorovs 0-1-lag)** *Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler på ett sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Då gäller, för varje händelse $A$ i svans-$\sigma$-algebran $\mathcal{T} = \mathcal{T}(X_1, X_2, \ldots)$, att*

$$\mathbb{P}(A) \in \{0, 1\}\,.$$

För beviset av Kolmogorovs 0-1-lag skall vi använda oss av följande, här obevisade, lemma.

**Lemma 6.3** *Låt $X_1, X_2, \ldots$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. För varje $A \in \mathcal{F}_1^\infty$ och varje $\varepsilon > 0$ existerar det ett $k < \infty$ och en händelse $A_k \in \mathcal{F}_1^k$ sådan att*

$$\mathbb{P}(A \setminus A_k) + \mathbb{P}(A_k \setminus A) < \varepsilon\,. \tag{30}$$

Ett sätt att förstå vad lemmat säger är följande: För varje händelse $A$ definierad i termer av $X_1, X_2, \ldots$, kan vi med godtyckligt liten felsannolikhet avgöra om $A$ inträffat genom att observera ändligt många av $X_i$-variablerna.

**Bevis av Sats 6.4.** Fixera $A \in \mathcal{T}$ och $\varepsilon > 0$, och välj sedan $k < \infty$ och en händelse $A_k \in \mathcal{F}_1^k$ som i Lemma 6.3, dvs så att (30) är uppfylld. Då är $P(A \setminus A_k) < \varepsilon$, vilket medför att

$$\mathbb{P}(A \cap A_k) > \mathbb{P}(A) - \varepsilon\,. \tag{31}$$

Av (30) följer även att $\mathbb{P}(A_k \setminus A) < \varepsilon$, vilket medför att

$$\mathbb{P}(A_k) < \mathbb{P}(A) + \varepsilon\,. \tag{32}$$

Antagandet att $X_1, X_2, \ldots$ är oberoende gör att $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$ för varje par av händelser $B \in \mathcal{F}_1^k, C \in \mathcal{F}_{k+1}^\infty$. Men $A_k \in \mathcal{F}_k$ och $A \in \mathcal{T} \subseteq \mathcal{F}_{k+1}^\infty$, så vi har alltså att

$$\mathbb{P}(A \cap A_k) = \mathbb{P}(A)\mathbb{P}(A_k)\,.$$

Insättning i (31) ger

$$\mathbb{P}(A)\mathbb{P}(A_k) > \mathbb{P}(A) - \varepsilon$$

och med användande av (32) får vi

$$\mathbb{P}(A)(\mathbb{P}(A)+\varepsilon) > \mathbb{P}(A)-\varepsilon\,.$$

Eftersom $\varepsilon$ kunde väljas godtyckligt litet har vi alltså

$$\mathbb{P}(A)^2 \geq \mathbb{P}(A)\,. \tag{33}$$

Men $x^2 < x$ för varje $x \in (0,1)$, så enda möjligheten för (33) att vara uppfylld är om $\mathbb{P}(A)$ är lika med 0 eller 1. □

**Övningar.**

6.4 Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$ sådana att $X_n \xrightarrow{\mathcal{P}} X$. I Exempel 6.2 såg vi att detta inte medför $X_n \xrightarrow{n.s.} X$. Visa att det däremot alltid går att hitta en delsekvens som konvergerar mot $X$ nästan säkert, dvs att det finns heltal $a_1, a_2, \ldots$ sådana att $X_{a_i} \xrightarrow{n.s.} X$ då $i \to \infty$.

6.5 Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, sådana att $X_1, X_2, \ldots$ är oberoende.

(a) Visa att om $X_n \xrightarrow{n.s.} X$ så är $\mathbb{P}(X = K) = 1$ för någon konstant $K$.

(b) Utnyttja resultaten i (a) och i Övning 6.4 till att visa att om $X_n \xrightarrow{\mathcal{P}} X$ så är $\mathbb{P}(X = K) = 1$ för någon konstant $K$.

6.6 Låt $X_1, X_2, \ldots$ vara oberoende stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. Visa att

$$\sup_n X_n < \infty$$

om och endast om det finns något $A < \infty$ sådant att

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n > A) < \infty\,.$$

6.7 Låt $X_1, X_2, \ldots$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, och visa att $\sigma(X_1, X_2, \ldots) \subseteq \mathcal{F}$.

6.8 Låt $X_1, X_2, \ldots$ vara stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$. Visa att $\mathcal{T}(X_1, X_2, \ldots)$ är en $\sigma$-algebra, och att $\mathcal{T}(X_1, X_2, \ldots) \subseteq \mathcal{F}$.

6.9 Visa att händelserna $A_3$, $A_4$, $A_5$ och $A_6$ definierade i anslutning till Definition 6.6, verkligen tillhör svans-$\sigma$-algebran $\mathcal{T}$, dvs att de för varje ändligt $k$ är opåverkade av om vi ändrar värdena på $(X_1, \ldots, X_k)$.

6.10 Visa, med hjälp av ett motexempel, att Kolmogorovs 0-1-lag inte gäller om man stryker oberoendeantagandet.

## 6.3 Stora talens lag

I detta avsnitt, och senare i Avsnitt 7.2, skall vi bekanta oss med det ena av sannolikhetsteorins båda flaggskepp: Stora talens lag (det andra är naturligtvis Centrala gränsvärdessatsen). Stora talens lag säger som bekant att om $X_1, X_2, \ldots$ är oberoende och likafördelade med väntevärde $\mu$, så konvergerar medelvärdet $\frac{X_1+\cdots+X_n}{n}$ mot $\mu$. Satsen förekommer emellertid i ett antal olika varianter, beroende på vilken typ av konvergens som avses, och hur generell klass av stokastiska variabler som tillåts. Vi skall i Satserna 6.5, 6.6, 7.7 och 7.8 formulera fyra olika varianter, där typ av konvergens och klass av stokastiska variabler ges av följande tabell.

| | $X_i \in L^2$ | $X_i \in L^1$ |
|---|---|---|
| konvergens i sannolikhet | Sats 6.5 | Sats 7.7 |
| nästan säker konvergens | Sats 6.6 | Sats 7.8 |

Eftersom nästan säker konvergens medför konvergens i sannolikhet, och $L^2 \subseteq L^1$, är Sats 7.8 den starkaste och Sats 6.5 den svagaste varianten. Traditionen bjuder att ”nästan säker”-varianterna kallas Starka stora talens lag, medan medan ”i sannolikhet”-varianterna kallas Svaga stora talens lag.

Ett språkligt påpekande: ”Stora talens lag” är en direktöversättning av ”Law of large numbers”. Engelskans ”number” svarar mot svenskans ”tal”, men också mot ”antal”. I detta sammanhang syftar ”Law of large numbers” på det stora *antalet* $X_i$-variabler. Det borde alltså egentligen heta ”Stora antalens lag” snarare än ”Stora talens lag”! Vi bryr oss emellertid inte om att försöka korrigera denna felöversättning, då den numera får anses vara en inarbetad del av det svenska språket.

Här följer den enklaste, och svagaste, varianten av Stora talens lag; du känner säkert igen såväl formulering som bevis.

**Sats 6.5 (Svaga stora talens lag; ändlig varians)** *Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, med ändligt väntevärde $\mu$ och ändlig varians $\sigma^2$. Då gäller, för de successiva medelvärdena*

$$M_n = \frac{X_1 + \cdots + X_n}{n}\,, \tag{34}$$

*att*

$$M_n \xrightarrow{\mathcal{P}} \mu\,.$$

**Bevis.** Fixera $\varepsilon > 0$. Vi behöver visa att

$$\lim_{n\to\infty} \mathbb{P}(|M_n - \mu| > \varepsilon) = 0\,. \tag{35}$$

Väntevärde och varians för $M_n$ ges av

$$\mathbb{E}[M_n] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \frac{n\mu}{n} = \mu$$

och

$$\mathrm{Var}[M_n] = \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[X_i] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}\,.$$

Tjebysjevs olikhet (se Övning 5.3) ger att

$$\mathbb{P}(|M_n - \mu| > \varepsilon) \leq \frac{\mathrm{Var}[M_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \tag{36}$$

vilket går mot 0 då $n \to \infty$, så (35) är bevisad. □

**Exempel 6.6: Utan oberoendeantagandet gäller inte Stora talens lag.** Låt

$$X_1 = \begin{cases} 0 & \text{med sannolikhet } \frac{1}{2} \\ 1 & \text{med sannolikhet } \frac{1}{2}\,, \end{cases}$$

och låt sedan $X_i = X_1$ för $i = 2, 3, \ldots$. Då är $X_1, X_2, \ldots$ likafördelade med väntevärde $\frac{1}{2}$ och varians $\frac{1}{4}$, och uppfyller således alla villkor i Sats 6.5 utom oberoende. För medelvädet $M_n$ definierat som i (34) får vi $M_n = X_1$ för alla $n$, så $M_n$ konvergerar alltså *inte* i sannolikhet (eller i någon annan rimlig mening) mot väntevärdet $\frac{1}{2}$ då $n \to \infty$.

Detta exempel visar att oberoendeantagandet i Sats 6.5 är väsentligt. Å andra sidan finns det situationer där oberoendet kan ersättas med diverse svagare villkor; se exempelvis Övning 6.12. Att avgöra (eller finna lämpliga villkor på) när en följd av beroende stokastiska variabler beter sig som i Stora talens lag, är ett problem som under större delen av 1900-talet (och än idag) ägnats stor kraft och många avhandlingar.

Vår nästa uppgift blir att skärpa konvergensen i Sats 6.5 till nästan säker konvergens.

**Sats 6.6 (Starka stora talens lag; ändlig varians)** *Antag att $X_1, X_2, \ldots$ är oberoende och likafördelade stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, med ändligt väntevärde $\mu$ och ändlig varians $\sigma^2$. Då gäller, för de successiva medelvärdena $M_n$ definierade i (34), att*

$$M_n \overset{n.s.}{\to} \mu\,.$$

För att bevisa denna variant av Stora talens lag är det praktiskt att först visa följande elementära lemma (som i sig inte har något med sannolikhetsteori att göra, utan är ren aritmetik). Vi utvidgar definitionen av medelvärde genom att, för $i < j$, definiera

$$M_{(i,j)} = \frac{X_{i+1} + X_{i+2} + \cdots + X_j}{j - i}$$

dvs $M_{(i,j)}$ är medelvärdet av variablerna $X_{i+1}, X_{i+2}, \ldots, X_j$.

**Lemma 6.4** *Låt $\varepsilon > 0$ och $k \in \mathbb{N}$ vara godtyckliga, och låt $i \in \{2^k, 2^k+1, \ldots, 2^{k+1}-1\}$. Om $|M_i - \mu| > \varepsilon$ och $|M_{(i,2^{k+2})} - \mu| \leq \frac{\varepsilon}{5}$, så gäller även att $|M_{2^{k+2}} - \mu| > \frac{\varepsilon}{20}$.*

**Bevis.** Antag att $M_i - \mu > \varepsilon$; fallet $M_i - \mu < -\varepsilon$ visas analogt. Vi har då

$$\begin{aligned} M_{2^{k+2}} &= \frac{X_1 + \cdots + X_i}{2^{k+2}} + \frac{X_{i+1} + \cdots + X_{2^{k+2}}}{2^{k+2}} \\ &= \frac{i}{2^{k+2}} M_i + \frac{2^{k+2} - i}{2^{k+2}} M_{(i,2^{k+2})} \\ &= \mu + \frac{i}{2^{k+2}}(M_i - \mu) + \frac{2^{k+2} - i}{2^{k+2}}(M_{(i,2^{k+2})} - \mu) \\ &> \mu + \frac{\varepsilon}{4} - \frac{\varepsilon}{5} = \frac{\varepsilon}{20}. \end{aligned}$$

$\square$

**Bevis av Sats 6.6.** Fixera $\varepsilon > 0$. Vi behöver visa att

$$\mathbb{P}(|M_n - \mu| > \varepsilon \text{ för högst ändligt många } n) = 1\,. \tag{37}$$

Beviset går ut på att dela upp följden $(1, 2, 3, \ldots)$ i block, enligt

$$(1),\ (2,3),\ (4,5,6,7), \ldots, (2^k, 2^k+1, \ldots, 2^{k+1}-1), \ldots \tag{38}$$

och visa att $|M_n - \mu|$ överskrider $\varepsilon$ i högst ändligt många av dessa block. För $k = 1, 2, \ldots$, definiera indikatorvariabeln

$$Z_k = \begin{cases} 1 & \text{om } |M_i - \mu| > \varepsilon \text{ för något } i \in \{2^k, \ldots, 2^{k+1}-1\} \\ 0 & \text{annars.} \end{cases}$$

Om $Z_k = 1$ så är alltså $|M_i - \mu|$ större än $\varepsilon$ för något $i$ i blocket $\{2^k, \ldots, 2^{k+1}-1\}$, och vi kan då låta $T_k$ vara den första "tidpunkt" vid vilken detta inträffar, dvs

$$T_k = \min\{i \in \{2^k, \ldots, 2^{k+1}-1\} : |M_i - \mu| > \varepsilon\}\,.$$

Välj nu $K$ så stort att

$$\mathbb{P}(|M_n - \mu| \leq \frac{\varepsilon}{5}) \geq \frac{1}{2}$$

för varje $n \geq 2^{K+1}$; att ett sådant $K$ går att hitta följer av att $\lim_{n\to\infty} \mathbb{P}(|M_n - \mu| \leq \frac{\varepsilon}{5}) = 0$ pga Sats 6.5. Då gäller även att

$$\mathbb{P}(|M_{(i,i+n)} - \mu| \leq \frac{\varepsilon}{5}) \geq \frac{1}{2}$$

för alla $i$ och alla $n \geq 2^{K+1}$, ty $M_{(i,i+n)}$ och $M_n$ har samma fördelning.

Välj ett $i \in \{2^k, \dots, 2^{k+1} - 1\}$. Om händelserna $\{Z_k = 1, T_k = i\}$ och $\{|M_{(i,2^{k+2})} - \mu| \leq \frac{\varepsilon}{5}\}$ inträffar, så inträffar även händelsen $\{|M_{2^{k+2}} - \mu| > \frac{\varepsilon}{20}\}$ pga Lemma 6.4. Notera också att $\{Z_k = 1, T_k = i\}$ och $\{|M_{(i,i+n)} - \mu| \leq \frac{\varepsilon}{5}\}$ är oberoende händelser, dvs

$$\mathbb{P}(Z_k = 1, T_k = i, |M_{(i,i+n)} - \mu| \leq \frac{\varepsilon}{5}) = \mathbb{P}(Z_k = 1, T_k = i)\mathbb{P}(|M_{(i,i+n)} - \mu| \leq \frac{\varepsilon}{5})\,.$$

Detta följer av att den förstnämnda händelsen bara beror på $X_1, \dots, X_i$, medan den andra bara beror på $X_{i+1}, \dots, X_{2^{k+2}}$. För $k \geq K$ får vi därför

$$\begin{aligned}
\mathbb{P}(|M_{2^{k+2}} - \mu| > \frac{\varepsilon}{20}) &\geq \mathbb{P}(Z_k = 1, |M_{2^{k+2}} - \mu| > \frac{\varepsilon}{20}) \\
&= \sum_{i=2^k}^{2^{k+1}-1} \mathbb{P}(Z_k = 1, T_k = i, |M_{2^{k+2}} - \mu| > \frac{\varepsilon}{20}) \\
&\geq \sum_{i=2^k}^{2^{k+1}-1} \mathbb{P}(Z_k = 1, T_k = i, |M_{(i,2^{k+2})} - \mu| \leq \frac{\varepsilon}{5}) \\
&= \sum_{i=2^k}^{2^{k+1}-1} \mathbb{P}(Z_k = 1, T_k = i)\mathbb{P}(|M_{(i,2^{k+2})} - \mu| \leq \frac{\varepsilon}{5}) \\
&\geq \frac{1}{2}\sum_{i=2^k}^{2^{k+1}-1} \mathbb{P}(Z_k = 1, T_k = i) \\
&= \frac{1}{2}\mathbb{P}(Z_k = 1)\,.
\end{aligned} \tag{39}$$

Å andra sidan har vi, pga (36),

$$\mathbb{P}(|M_{2^{k+2}} - \mu| > \frac{\varepsilon}{20}) \leq \frac{\sigma^2}{2^{k+2}(\frac{\varepsilon}{20})^2}$$

och om vi kombinerar denna observation med (39) får vi

$$\mathbb{P}(Z_k = 1) \leq \frac{2\sigma^2}{2^{k+2}(\frac{\varepsilon}{20})^2}$$

för $k \geq K$. Det följer att

$$\begin{aligned}\sum_{k=K}^{\infty} \mathbb{P}(Z_k = 1) &\leq \sum_{k=K}^{\infty} \frac{2\sigma^2}{2^{k+2}(\frac{\varepsilon}{20})^2} \\ &= \frac{200\sigma^2}{\varepsilon^2} \sum_{k=K}^{\infty} \frac{1}{2^k} < \infty\,.\end{aligned}$$

Självklart gäller då även att

$$\sum_{k=1}^{\infty} \mathbb{P}(Z_k = 1) < \infty$$

och Borel–Cantellis första lemma säger oss att

$$\mathbb{P}(Z_k = 1 \text{ inträffar för högst ändligt många } k) = 1\,.$$

Men detta betyder att $|M_n - \mu|$ överskrider $\varepsilon$ i högst ändligt många av blocken i (38), vilket medför (37). $\square$

## Övningar.

6.11 Antag att vi försöker använda beviset av Sats 6.5 för att visa (det felaktiga) påståendet att $M_n \xrightarrow{\mathcal{P}} \frac{1}{2}$, där $X_1, X_2, \ldots$ och $M_1, M_2, \ldots$ är som i Exempel 6.6. Var någonstans bryter beviset samman, och varför?

6.12 **Stora talens lag för svagt stationära stokastiska processer.** Låt $X_1, X_2, \ldots$ vara stokastiska variabler i $L^2(\Omega, \mathcal{F}, \mathbb{P})$, och antag att $\mathbb{E}[X_n]$ är detsamma för alla $n$ och att $\mathrm{Cov}[X_n, X_{n+k}]$ bara beror på $k$. Under dessa villkor brukar $(X_1, X_2, \ldots)$ kallas för en **svagt stationär stokastisk process**. $E[X_n]$ brukar betecknas med $m_X$ medan $\mathrm{Cov}[X_n, X_{n+k}]$ betecknas $r_X(k)$, vilket kallas **kovariansfunktionen**. Definiera också de successiva medelvärdena $M_n$ som i (34). Visa (med samma metod som i beviset av Sats 6.5) att om kovariansfunktionen är summerbar, dvs om $\sum_{k=-\infty}^{\infty} r_X(k) < \infty$, så gäller $M_n \xrightarrow{\mathcal{P}} m_X$.

# 7 TRANSFORMMETODER

**Transformmetoder** är viktiga och ofta använda verktyg i stora delar av den högre matematiken, och sannolikhetsteorin utgör inget undantag. Vitsen med transformmetoder kan illustreras med följande exempel, som före datorernas och räknedosornas tid var välbekant även för gymnasister.

Antag att vi, utan tillgång till modern elektronik, vill multiplicera två positiva tal $x$ och $y$. Med hjälp av formeln

$$xy = \exp(\log(x) + \log(y))$$

inser vi att vi, istället för att utföra multiplikationen direkt, kan bestämma logaritmerna för $x$ och $y$, addera dessa, och sedan exponentiera svaret. På så vis **transformerar** vi problemet att multiplicera två tal till dett lättare problemet att addera dem. Priset vi får betala för detta är själva transformeringen: vi behöver beräkna två logaritmer och en exponential. Detta gjorde man förr med hjälp av färdiga tabeller.

När det är fråga om mer avancerade transformmetoder är det ofta funktioner, snarare än tal, som transformeras. De funktioner vi oftast önskar transformera i sannolikhetsteorin är de stokastiska variablerna. Två transformer av stokastiska variabler som är användbara i olika sammanhang är den sannolikhetsgenererande funktionen, och Laplacetransformen:

**Definition 7.1 Den sannolikhetsgenererande funktionen** $G_X : \mathbb{R} \to \mathbb{R}$ *av en ickenegativ heltalsvärd stokastisk variabel* $X$*, ges av*

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k)\,.$$

**Definition 7.2 Laplacetransformen** $L_X : [0,\infty) \to [0,\infty)$ *av en ickenegativ stokastisk variabel* $X$*, ges av*

$$L_X(\lambda) = \mathbb{E}[e^{-\lambda X}]\,.$$

I stället för att orda mer om dessa transformer, skall vi i nästa avsnitt stifta bekantskap med den mer allmänt användbara **karaktäristiska funktionen** $\phi_X(t)$ för en stokastisk variabel $X$. Vi skall se att karaktäristiska funktionen inte är något annat än den välkända **Fouriertransformen**.

**Övningar.**

7.1 Låt $X$ och $Y$ vara ickenegativa heltalsvärda stokastiska variabler med sannolikhetsgenererande funktioner $G_X(s)$ och $G_Y(s)$.

(a) Visa att $\mathbb{E}[X] = G'_X(1)$ och att $\mathrm{Var}[X] = G''_X(1) + G'_X(1) - (G'_X(1))^2$.

(b) Visa att om $X$ och $Y$ är oberoende, så ges den sannolikhetsgenererande funktionen för summan $X + Y$ av $G_{X+Y}(s) = G_X(s)G_Y(s)$. (Ledning: Använd resultatet i Övning 4.3.)

7.2 Låt $X$ och $Y$ vara ickenegativa stokastiska variabler med Laplacetransformer $L_X(\lambda)$ och $L_Y(\lambda)$.

(a) Visa, för $n = 1, 2, \ldots$, att $\mathbb{E}[X^n] = (-1)^n L_X^{(n)}(0)$.

(b) Visa att om $X$ och $Y$ är oberoende, så ges Laplacetransformen för summan $X + Y$ av $L_{X+Y}(\lambda) = L_X(\lambda)L_Y(\lambda)$.

## 7.1 Karaktäristiska funktioner

Eftersom karaktäristiska funktioner inbegriper komplexvärda stokastiska variabler, måste vi först definiera vad som menas med en sådan. Låt $i$ beteckna den imaginära enheten, dvs (något vanvördigt uttryckt) "$i = \sqrt{-1}$". För $z \in \mathbb{C}$ ($\mathbb{C}$ är det komplexa planet) betecknar $\Re(z)$ och $\Im(z)$ realdel respektive imaginärdel för $z$, så att $z = \Re(z) + i\Im(z)$. Vi låter också $\bar{z}$ beteckna komplexkonjugatet till $z$, dvs $\bar{z} = \Re(z) - i\Im(z)$.

**Definition 7.3** *Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum, och låt $X$ vara en avbildning från $\Omega$ till $\mathbb{C}$. $X$ sägs vara en* **komplexvärd stokastisk variabel** *om $\Re(X)$ och $\Im(X)$ är vanliga (reellvärda) stokastiska variabler. Om $X$ är en komplexvärd stokastisk variabel, så säger vi att $X$ har ändligt väntevärde om $\mathbb{E}[|\Re(X)|] < \infty$ och $\mathbb{E}[|\Im[X]|] < \infty$, och i så fall definierar vi väntevärdet av $X$ som*

$$\mathbb{E}[X] = \mathbb{E}[\Re(X)] + i\mathbb{E}[\Im(X)]\,.$$

Två komplexvärda stokastiska variabler $X$ och $Y$ sägs vara oberoende om

$$\begin{aligned} & \mathbb{P}(\Re(X) \in A_1, \Im(X) \in A_2, \Re(Y) \in B_1, \Im(Y) \in B_2) \\ = \; & \mathbb{P}(\Re(X) \in A_1, \Im(X) \in A_2)\mathbb{P}(\Re(Y) \in B_1, \Im(Y) \in B_2) \end{aligned}$$

för alla $A_1, A_2, B_1, B_2 \in \mathcal{B}$.

Komplexvärda stokastiska variabler har egenskaper som liknar de reellvärda. Exempelvis har vi följande motsvarighet till Satserna 3.4 och 4.4.

**Sats 7.1** *Låt $X$ och $Y$ vara komplexvärda stokastiska variabler med ändligt väntevärde på $(\Omega, \mathcal{F}, \mathbb{P})$, och låt $a, b \in \mathbb{R}$. Då gäller*

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]\,.$$

*Om dessutom $X$ och $Y$ är oberoende så gäller*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]\,.$$

**Bevis.** Se Övning 7.3. □

**Definition 7.4** *För en (reellvärd) stokastisk variabel $X$ på $(\Omega, \mathcal{F}, \mathbb{P})$ definierar vi den* **karaktäristiska funktionen** *$\phi_X : \mathbb{R} \to \mathbb{C}$ för $X$ genom*

$$\phi_X(t) = \mathbb{E}[e^{itX}] \tag{40}$$

*för alla $t \in \mathbb{R}$.*

Högerledet i definitionen (40) kan alternativt skrivas som $\mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)]$. Eftersom $\cos(tX)$ och $\sin(tX)$ är begränsade funktioner, så har alltså $e^{itX}$ ändligt väntevärde (vare sig $X$ själv har ändligt väntevärde eller ej). Karaktäristiska funktionen är därmed väldefinierad för godtyckliga stokastiska variabler.

Om $X$ är en kontinuerlig stokastisk variabel med täthetsfunktion $f_X$, så kan karaktäristiska funktionen $\phi_X(t)$ uttryckas som

$$\int_{-\infty}^{\infty} f_X(x)e^{itx}\,dx \tag{41}$$

vilket du (om du läst Fourieranalys) känner igen som Fouriertransformen av $f_X$. Om istället $X$ är diskret och heltalsvärd, så är $\phi_X(t)$ lika med

$$\sum_{k=-\infty}^{\infty} \mathbb{P}(X = k)e^{itk}$$

vilket är detsamma som den diskreta Fouriertransformen av funktionen $\mathbb{P}(X = \cdot) : \mathbb{Z} \to \mathbb{R}$. Härav vårt påstående i början av kapitlet att karaktäristiska funktionen är en Fouriertransform. Mycket av den underliggande teorin för karaktäristiska funktioner hör därför till Fourieranalysen snarare än till sannolikhetsteorin, och kommer därför här att behandlas en smula summariskt.

**Exempel 7.1: Normalfördelningen.** En kontinuerlig stokastisk variabel $X$ sägs (som bekant) vara **normalfördelad** med väntevärde $\mu$ och varians $\sigma^2$, om den har frekvensfunktion

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

Vi skriver då $X \stackrel{\mathcal{D}}{=} N(\mu, \sigma^2)$. Karaktäristiska funktionen fås med hjälp av (41):

$$\begin{aligned} \phi_X(t) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{itx}\, dx \qquad (42) \\ &= e^{i\mu t - \frac{1}{2}\sigma^2 t^2} . \end{aligned}$$

(Vi föreslår att du, i en sannolikhetskurs, inte lägger ner kraft på att för hand beräkna integraler som den i (42), som istället kan lösas med hjälp av något formelmanipulerande programpaket, eller slås upp i lämplig tabell över Fouriertransformer.)

Användbara egenskaper hos karaktäristiska funktioner ges i följande satser.

**Sats 7.2** *Låt $X$ vara en stokastisk variabel med karaktäristisk funktion $\phi_X(t)$, låt $a, b \in \mathbb{R}$, och låt $Y = aX + b$. Då har $Y$ karaktäristisk funktion*

$$\phi_Y(t) = e^{itb}\phi_X(at) .$$

**Bevis.** Direkt räkning ger

$$\begin{aligned} \phi_Y(t) &= \mathbb{E}[e^{it(aX+b)}] = \mathbb{E}[e^{itb}e^{itaX}] \\ &= e^{itb}\mathbb{E}[e^{itaX}] = e^{itb}\phi_X(at) . \end{aligned}$$

□

**Sats 7.3** *Låt $X_1, \ldots, X_n$ vara oberoende stokastiska variabler med karaktäristiska funktioner $\phi_{X_1}(t), \ldots, \phi_{X_n}(t)$, och låt $Y_n = X_1 + \cdots + X_n$. Då har $Y_n$ karaktäristisk funktion*

$$\phi_{Y_n}(t) = \prod_{j=1}^{n} \phi_{X_i}(t) .$$

**Bevis.** För $n = 2$ får vi

$$\begin{aligned} \phi_{Y_2}(t) &= \mathbb{E}[e^{it(X_1+X_2)}] = \mathbb{E}[e^{itX_1}e^{itX_2}] \\ &= \mathbb{E}[e^{itX_1}]\mathbb{E}[e^{itX_2}] = \phi_{X_1}(t)\phi_{X_2}(t) , \end{aligned}$$

där den tredje likheten följer av Sats 7.1. Fallen $n = 3, 4, \ldots$ följer med hjälp av induktion: Antag att satsen gäller för $n = k - 1$. Eftersom $Y_{k-1}$ och $X_k$ är oberoende får vi

$$\begin{aligned}\phi_{Y_k}(t) &= \phi_{Y_{k-1}}(t)\phi_{X_k}(t) \\ &= \left(\prod_{j=1}^{k-1} \phi_{X_j}(t)\right) \phi_{X_k}(t) = \prod_{j=1}^{k} \phi_{X_j}(t)\end{aligned}$$

så att satsen även gäller för $n = k$, och induktionssteget är avklarat. □

Särskilt Sats 7.3 är mycket användbar. Om man önskar beräkna fördelningen för summan av två (eller flera) oberoende stokastiska variabler, råkar man ofta ut för besvärliga faltningsintegraler (se (14) för fallet med kontinuerliga stokastiska variabler). Då är det ofta lättare att (i) ta fram karaktäristiska funktionen för de i summan ingående stokastiska variablerna, (ii) multiplicera dessa, och (iii) avgöra vilken fördelning som har produkten som karaktäristisk funktion (exempelvis med hjälp av Sats 7.5 nedan). I Avsnitt 7.2 och 7.3 skall vi använda denna typ av argument för att visa Stora talens lag och Centrala gränsvärdessatsen.

Vi fortsätter med fler trevliga egenskaper hos karaktäristiska funktioner:

**Sats 7.4** *Antag att $\phi_X$ är karaktäristiska funktionen för någon stokastisk variabel $X$. Då gäller:*

(i) $\phi_X(0) = 1$,

(ii) $|\phi_X(t)| \leq 1$ *för alla* $t$,

(iii) $\phi_X$ *är kontinuerlig, och*

(iv) $\phi_X$ *är positivt semi-definit, dvs*

$$\sum_{j=1}^{n} \sum_{k=1}^{n} z_j \bar{z}_k \phi(t_j - t_k) \geq 0$$

*för varje* $n$ *och varje val av* $t_1, \ldots, t_n \in \mathbb{R}$ *och* $z_1, \ldots, z_n \in \mathbb{C}$.

Beviset lämnas som Övning 7.5. Även omvändningen gäller: om $\phi_X$ är en funktion som uppfyller egenskaperna (i)–(iv), så går det att konstruera en stokastisk variabel vars karaktäristiska funktion är $\phi_X$. Detta berömda resultat bär namnet Bochners sats.

Nästa sats talar om hur man, givet en karaktäristisk funktion, kommer åt motsvarande fördelningsfunktion. För att inte ge oss alltför långt in i komplex- och Fourieranalysen avstår vi åter från bevis. För en fördelningsfunktion $F$ definierar vi $F^* : \mathbb{R} \to [0,1]$ genom att låta

$$F^*(x) = \frac{1}{2}\left(F(x) + \lim_{y \uparrow x} F(y)\right) \tag{43}$$

för varje $x \in \mathbb{R}$. ($F^*$ skiljer sig alltså från $F$ endast i diskontinuitetspunkterna, där den ligger "halvvägs upp".)

**Sats 7.5 (Fouriers inversionsformel)** *Antag att $X$ är en stokastisk variabel med karaktäristisk funktion $\phi_X(t)$ och fördelningsfunktion $F_X$, och definiera $F_X^*$ som i (43). Då gäller, för $x, y \in \mathbb{R}$, att*

$$F_X^*(x) - F_X^*(y) = \lim_{N\to\infty} \int_{-N}^{N} \frac{e^{-iyt} - e^{-ixt}}{i2\pi t} \phi_X(t)\, dt\,. \tag{44}$$

Om $X$ är en kontinuerlig stokastisk variabel, så kan vi derivera med avseende på $x$ i (44), och får (den kanske mer välbekanta) formeln

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t)\, dt\,.$$

I det allmänna fallet gäller (se Övning 7.7) att om $F_X^*(x) = F_Y^*(x)$ för alla $x$, så är även $F_X(x) = F_Y(x)$ för alla $x$. Som en konsekvens av Sats 7.5 har vi alltså att varje karaktäristisk funktion svarar mot endast en fördelningsfunktion Med andra ord: om vi vet $\phi_X$ kan vi återskapa $F_X$. Följande sats säger ännu mer.

**Sats 7.6** *Låt $X_1, X_2, \ldots$ och $X$ vara stokastiska variabler med karaktäristiska funktioner $\phi_{X_1}, \phi_{X_2}, \ldots$ och $\phi_X$. Om*

$$\lim_{n\to\infty} \phi_{X_n}(t) = \phi_X(t)$$

*för varje $t \in \mathbb{R}$, så har vi att $X_n \xrightarrow{\mathcal{D}} X$.*

**Övningar.**

7.3 Bevisa Sats 7.1. (Ledning: Skriv $X$ och $Y$ som $\Re(X) + i\Im(X)$ och $\Re(Y) + i\Im(Y)$, och utnyttja sedan Satserna 3.4 och 4.4.)

7.4 Låt $X_1, X_2, \ldots$ vara oberoende likafördelade stokastiska variabler med fördelning $N(\mu, \sigma^2)$, och låt $M_n = (X_1 + \ldots + X_n)/n$.

(a) Vilken fördelning har $M_n$?

(b) Fixera $\varepsilon > 0$. Hur beter sig $\mathbb{P}(|M_n - \mu| > \varepsilon)$ då $n \to \infty$? (Ledning: Använd approximationen $\int_x^\infty e^{-y^2/2} dy \sim x^{-1} e^{-x^2/2}$ då $x \to \infty$, där $f(x) \sim g(x)$ betyder att $f(x)/g(x) \to 1$.)

(c) Använd resultatet i (b) och Borel–Cantellis första lemma för att ge ett kort bevis av Starka Stora talens lag i fallet med normalfördelade stokastiska variabler.

7.5 Bevisa Sats 7.4. (Ledning: Sats 3.6 är till hjälp för att visa (iii).)

7.6 (a) Antag att $X \in L^1$, och visa att $\phi'_X(0) = i\mathbb{E}[X]$.

(b) Antag att $X \in L^2$ och visa att $\phi''_X(0) = -\mathbb{E}[X^2]$.

(c) Vad blir motsvarande formel för $\phi_X^{(n)}(0)$ då $X \in L^n$, för $n = 3, 4, \ldots$?

7.7 Låt $F_X$ och $F_Y$ vara fördelningsfunktioner, och definiera $F_X^*$ och $F_Y^*$ som i (43). Visa att om $F_X^*(x) = F_Y^*(x)$ för alla $x$, så är $F_X = F_Y$.

## 7.2 Stora talens lag på nytt

Vi skall nu använda karaktäristiska funktioner till att visa hur Svaga stora talens lag (Sats 6.5) kan utvidgas till fallet med stokastiska variabler i $L^1$ (dvs kravet på ändlig varians slopas; allt som krävs är ändligt väntevärde). I Övning 7.8 skall vi se att detta är en meningsfull utvidgning, dvs att det finns situationer där Sats 7.7 är tillämplig men inte Sats 6.5.

**Sats 7.7 (Svaga stora talens lag; allmänna fallet)** *Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, med ändligt väntevärde $\mu$. Då gäller, för de successiva medelvärdena $M_n = (X_1 + \cdots + X_n)/n$, att*

$$M_n \xrightarrow{\mathcal{P}} \mu \,.$$

**Bevis.** Låt $\phi_{X_1}(t)$ vara karaktäristiska funktionen för $X_1$. Vi har att $\phi_{X_1}(0) = \mathbb{E}[e^0] = 1$. Med hjälp av Övning 7.6 ser vi också att $\phi'_{X_1}(0) = i\mu$. Vi kan därför göra följande första ordningens McLaurinutveckling:

$$\phi_{X_1}(t) = 1 + i\mu t + \theta_1(t)$$

där $\theta_1(t)$ är en restterm sådan att $\frac{\theta_1(t)}{t}$ går mot 0 då $t \to 0$. Med hjälp av Satserna 7.2 och 7.3 får vi att $M_n$ har karaktäristisk funktion

$$\begin{aligned}\phi_{M_n}(t) &= \prod_{j=1}^{n} \phi_{X_j}\left(\frac{t}{n}\right) = \left(\phi_{X_1}\left(\frac{t}{n}\right)\right)^n \\ &= \left(1 + \frac{i\mu t}{n} + \theta_1\left(\frac{t}{n}\right)\right)^n .\end{aligned}$$

Om vi nu låter $n \to \infty$, så kan vi uttnyttja att $\lim_{n\to\infty}(1 + \frac{a}{n})^n = e^a$ och att resttermen kan försummas. Vi får därmed

$$\lim_{n\to\infty} \phi_{M_n}(t) = e^{i\mu t}$$

vilket inte är något annat än den karaktäristiska funktionen för konstanten $\mu$. Detta medför, pga Sats 7.6, att $M_n \xrightarrow{\mathcal{D}} \mu$. Med stöd av Sats 6.2 har vi då även att $M_n \xrightarrow{\mathcal{P}} \mu$. □

Vi har nu fyllt i tre av de fyra hörnen i tabellen på sidan 58. Den fjärde, och starkaste, satsen ger vi utan bevis. Förhoppningsvis har vi härmed retat din nyfikenhet såpass att du konsulterar någon av de mer avancerade texter som omnämns i Kapitel 9 för att se hur den bevisas.

**Sats 7.8 (Starka stora talens lag; allmänna fallet)** *Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, med ändligt väntevärde $\mu$. Då gäller, för de successiva medelvärdena $M_n = (X_1 + \cdots + X_n)/n$, att*

$$M_n \xrightarrow{n.s.} \mu \,.$$

**Övningar.**

7.8 Ge exempel på en stokastisk variabel $X$ som har ändligt väntevärde men inte ändlig varians.

7.9 **Cauchyfördelningen.** En kontinuerlig stokastisk variabel $X$ sägs vara **Cauchyfördelad** om den har täthetsfunktion $f_X(x) = \frac{1}{\pi(1+x^2)}$. Låt $X_1, X_2, \ldots$ vara oberoende Cauchyfördelade stokastiska variabler, och låt som vanligt $M_n = (X_1 + \cdots + X_n)/n$.

(a) Bestäm täthetsfunktionen för $M_n$. (Ledning: Utnyttja att $\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{itx}}{1+x^2} dx = e^{-|t|}$.)

(b) Varför motsäger inte resultatet i (a) Stora talens lag?

## 7.3 Centrala gränsvärdessatsen

När vi nu behandlat Stora talens lag i några olika former, så inställer sig den naturliga frågan: Hur mycket avviker $M_n = (X_1 + \cdots + X_n)/n$ från $\mu$? I fallet med ändlig varians ges svaret av **Centrala gränsvärdessatsen** som säger att, om $X_i$-variablerna är oberoende och likafördelade med väntevärde $\mu$ och varians $\sigma^2$, så gäller för stora $n$ att

(i) avvikelsen $M_n - \mu$ är av storleksordningen $\frac{1}{\sqrt{n}}$, och

(ii) $\sqrt{\frac{n}{\sigma^2}}(M_n - \mu)$ är approximativt $N(0,1)$, dvs normalfördelad med väntevärde 0 och varians 1.

En exakt formulering är följande:

**Sats 7.9 (Centrala gränsvärdessatsen)** *Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, med ändligt väntevärde $\mu$ och ändlig varians $\sigma^2$. Då gäller, för de successiva medelvärdena $M_n = (X_1 + \cdots + X_n)/n$, att*

$$\sqrt{\frac{n}{\sigma^2}}(M_n - \mu) \xrightarrow{\mathcal{D}} N(0,1)\,. \tag{45}$$

**Bevis.** Beviset går till på liknande sätt som det av Sats 7.7; den väsentliga skillnaden är att vi tar med en term till i McLaurinutvecklingen. Låt

$$Z_j = \frac{X_j - \mu}{\sqrt{\sigma^2}}$$

för $j = 1, 2, \ldots$. Notera att $Z_j$ då har väntevärde 0 och varians 1, vilket pga (16) medför att $\mathbb{E}[Z_j^2] = \mathbb{E}[Z_j]^2 + \mathrm{Var}[Z_j] = 1$. Av Övning 7.6 vet vi att karaktäristiska funktionen $\phi_{Z_j}$ för $Z_j$ därmed uppfyller $\phi'_{Z_j}(0) = i\mathbb{E}[Z_j] = 0$ och $\phi''_{Z_j}(0) = -\mathbb{E}[Z_j^2] = -1$. Vi får därför följande andra ordningens McLaurinutveckling av $\phi_{Z_j}$:

$$\phi_{Z_j}(t) = 1 - \frac{t^2}{2} + \theta_2(t)$$

där resttermen $\theta_2(t)$ uppfyller $\lim_{t\to 0} \frac{\theta_2(t)}{t^2} = 0$. Låt $U_n$ beteckna vänsterledet i (45), och notera att

$$U_n = \sqrt{\frac{n}{\sigma^2}}(M_n - \mu) = \sum_{j=1}^{n} \frac{Z_j}{\sqrt{n}}\,.$$

Karaktäristiska funktionen för $U_n$ ges därför av

$$\begin{aligned}
\phi_{U_n}(t) &= \prod_{j=1}^{n} \phi_{Z_j}\left(\frac{t}{\sqrt{n}}\right) \\
&= \left(1 - \frac{t^2}{2n} + \theta_2\left(\frac{t}{\sqrt{n}}\right)\right)^n,
\end{aligned}$$

och om vi låter $n$ gå mot $\infty$ får vi, på samma sätt som i beviset av Sats 7.7,

$$\lim_{n\to\infty} \phi_{U_n}(t) = e^{-t^2/2}\,.$$

Här är högerledet lika med karaktäristiska funktionen för en stokastisk variabel som är $N(0,1)$ (se Exempel 7.1), och pga Sats 7.6 är beviset klart. $\square$

Det är värt att påpeka att medan Stora talens lag gäller även för stokastiska variabler med oändlig varians (men ändligt väntevärde), så gäller Centrala gränsvärdessatsen endast i fallet med ändlig varians. Summor av oberoende likafördelade stokastiska variabler med oändlig varians har i allmänhet ett helt annorlunda (och viss mening rikare) ”centrala gränsvärdesbeteende” än de med ändlig varians.

# 8 NÅGRA ILLUSTRATIVA EXEMPEL

Vi är nu så gott som färdiga med den allmänna teori vi hade tänkt gå igenom. De konkreta exemplen har inte precis stått som spön i backen, och för att i någon mån råda bot på detta ämnar vi i detta kapitel ge några ytterst konkreta exempel, där en del av våra lärdomar kommer till hjälp. Dessa exempel är valda så att de i någon mån ansluter till två av den moderna sannolikhetsteorins mest dynamiska forskningsområden: sannolikhetsteori för fraktaler, och perkolationsteori.

## 8.1 Cantorfördelningen

I Exempel 0.1 bjöd vi på en stokastisk variabel $X$ som varken är diskret eller kontinuerlig. Emellertid kan dess fördelning $\mu_X$ delas upp i en diskret och en kontinuerlig komponent: hälften av sannolikhetsmassan (den diskreta komponenten) ligger i punkten $\{0\}$, och resten (den kontinuerliga komponenten) ligger jämnt utsmetad på intervallet $[0, 30]$, där täthetsfunktionen $f_X$ existerar och är lika med $\frac{1}{60}$. Som utlovats i Exempel 1.7, skall vi i detta avsnitt presentera ett mer ”genuint” exempel på en stokastisk variabel som varken är diskret eller kontinuerlig: dess fördelning kan ej heller delas upp i en diskret och en kontinuerlig komponent.

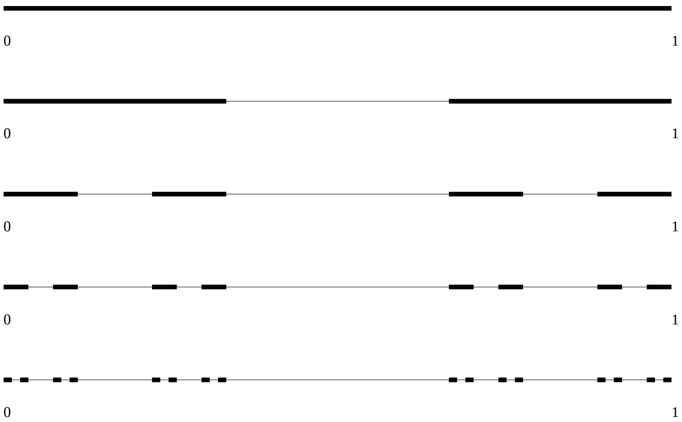
**Fraktaler** är viktiga objekt i den moderna matematiken, och för tiotalet år sedan nådde de stor spridning även i populärvetenskapen, främst i samband med så kallad kaosteori. Även i vissa delar av sannolikhetsteorin spelar fraktaler en framträdande roll, och det exempel vi valt att presentera inbegriper en av de enklaste och mest kända fraktalerna: den så kallade **Cantormängden**.

För att definiera Cantormängden börjar vi med att definiera mängder $\mathcal{C}_0, \mathcal{C}_1, \ldots \subset \mathbb{R}$ på följande vis. Låt $\mathcal{C}_0 = [0, 1]$. Bilda därefter $\mathcal{C}_1$ genom att ta bort den mittersta tredjedelen $(\frac{1}{3}, \frac{2}{3})$ av $\mathcal{C}_0$ (men lämna kvar randpunkterna), så att $\mathcal{C}_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Vi fortsätter sedan induktivt: varje $\mathcal{C}_k$ är en disjunkt union av ändligt många slutna delintervall, och $\mathcal{C}_{k+1}$ fås från $\mathcal{C}_k$ genom att ta bort den mittersta öppna tredjedelen i varje sådant delintervall. Exempelvis blir $\mathcal{C}_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. Se Figur 1.

Observera att $\mathcal{C}_0 \supset \mathcal{C}_1 \supset \mathcal{C}_2 \cdots$. Följden $\mathcal{C}_0, \mathcal{C}_1, \ldots$ har därför ett ”gränsvärde” $\mathcal{C}$ definierat av

$$\mathcal{C} = \mathcal{C}_\infty = \bigcap_{k=1}^{\infty} \mathcal{C}_k \,, \tag{46}$$

och det är $\mathcal{C}$ som vi kallar Cantormängden. En första fråga som man ställer sig när man ser Figur 1 är huruvida $\mathcal{C}$ överhuvudtaget innehåller några punkter. Svaret är ja, eftersom




Figur 1: Mängderna $\mathcal{C}_0, \ldots, \mathcal{C}_4$.

punkten 1 tillhör $\mathcal{C}_k$ för varje $k$, och detsamma gäller exempelvis punkterna $\frac{1}{3}$ och $\frac{20}{27}$. Men vi skall strax se att $\mathcal{C}$ innehåller långt fler punkter utöver dessa "randpunkter".

Varför sägs Cantormängden vara en fraktal? En utmärkande egenskap hos fraktaler är att de i viss mening kan anses vara $r$-dimensionella, där $r$ inte är ett heltal. För att förklara vad Cantormängden har för dimension börjar vi med ett enklare exempel: kvadraten $[0,1)^2 \subset \mathbb{R}^2$. Om vi förstorar upp $[0,1)^2$ med en (längd-)faktor 2 får vi $[0,2)^2$, som ju kan delas upp i 4 translaterade kopior av $[0,1)^2$, och vi anser därför att $[0,1)^2$ har dimension

$$\dim([0,1)^2) = \frac{\log(4)}{\log(2)} = 2\,.$$

Vi kunde välja en annan skalfaktor än 2 och få samma resultat: om vi exempelvis skalar upp $[0,1)^2$ med en faktor 7 får vi mängden $[0,7)^2$ som ju består av 49 translaterade kopior av $[0,1)^2$, och vi får $\dim([0,1)^2) = \frac{\log(49)}{\log(7)} = 2$. Ett liknande resonemang leder till att kuben $[0,1)^3$ har dimension 3, osv. Antag nu att vi förstorar upp Cantormängden $\mathcal{C}$ med en faktor 3, dvs vi betraktar mängden

$$3\mathcal{C} = \{x \in \mathbb{R} : \frac{x}{3} \in \mathcal{C}\}\,.$$

Det är lätt att konstatera att vi då får 2 kopior av $\mathcal{C}$: en i intervallet $[0,1]$ och en i intervallet $[2,3]$. Vi säger därför att Cantormängden har dimension[7]

$$\dim(\mathcal{C}) = \frac{\log(2)}{\log(3)} \approx 0.6309\,.$$

Vi skall nu ge en alternativ beskrivning av Cantormängden. Du är van vid att varje tal

[7]Vi medger att det dimensionsbegrepp vi använder här är en smula vagt, och endast fungerar för så kallat "självsimilära" mängder. Ett mer allmänt användbart begrepp är den så kallade Hausdorffdimensionen $\dim_H$. För samtliga mängder vilkas dimension vi här diskuterat, sammanfaller $\dim_H$ med $\dim$.

$a \in [0,1]$ kan representeras med sin decimalutveckling

$$a = 0.\mathrm{d}_1\mathrm{d}_2\mathrm{d}_3\ldots = \sum_{k=1}^{\infty} \mathrm{d}_k 10^{-k}\,,$$

där $d_1, d_2, \ldots$ antar värden i $\{0, 1, \ldots, 9\}$. Troligen har du också stött på motsvarande binärutveckling

$$a = (0.\mathrm{b}_1\mathrm{b}_2\mathrm{b}_3\ldots)_{\text{två}} = \sum_{k=1}^{\infty} \mathrm{b}_k 2^{-k}\,,$$

där $b_1, b_2, \ldots$ är $\{0,1\}$-värda. På samma sätt kan man skriva $a$ i termer av dess **trinärutveckling**

$$a = (0.\mathrm{t}_1\mathrm{t}_2\mathrm{t}_3\ldots)_{\text{tre}} = \sum_{k=1}^{\infty} \mathrm{t}_k 3^{-k}\,,$$

där $t_1, t_2, \ldots$ är $\{0,1,2\}$-värda. Det är nu lätt att verifiera att $\mathcal{C}_n$ består av exakt de tal $a \in [0,1]$ vilkas trinärutveckling $(0.\mathrm{t}_1\mathrm{t}_2\mathrm{t}_3\ldots)_{\text{tre}}$ uppfyller[8]

$$t_k \in \{0,2\} \quad \text{för } k = 1, \ldots, n\,. \tag{47}$$

Sedan följer det direkt av definitionen (46) att $\mathcal{C}$ består av de $a \in [0,1]$ vars trinärutveckling uppfyller

$$t_k \in \{0,2\} \quad \text{för alla } k \geq 1\,. \tag{48}$$

Eftersom det finns överuppräkneligt många sådana trinärutvecklingar, har vi alltså att Cantormängden är överuppräknelig.

Så här långt har vi inte diskuterat någon sannolikhetsteori i samband med Cantormängden. Nu är tiden kommen för detta, och vi vill definiera en stokastisk variabel $X$ som är ”likformigt fördelad” på $\mathcal{C}$. Vad är då rimligt att mena med likformig fördelning på $\mathcal{C}$?

Notera först att $\mathcal{C}$ kan delas upp i två likadana bitar (så när som på translation), där den ena ligger i $[0, \frac{1}{3}]$ och den andra i $[\frac{2}{3}, 1]$. Om $X$ skall kunna kallas likformigt fördelad på $\mathcal{C}$, bör vi därför rimligen kräva att

$$\mathbb{P}(X \in [0, \tfrac{1}{3}]) = \mathbb{P}(X \in [\tfrac{2}{3}, 1]) = \frac{1}{2}\,,$$

vilket betyder att första koefficienten i trinärutvecklingen för $X$ skall anta värdet 0 eller 2 med sannolikhet $\frac{1}{2}$ vardera. På samma sätt inser vi att vi bör ha $\mathbb{P}(X \in I) = \frac{1}{4}$ för

[8] Vissa rationella tal har inte bara en, utan *två olika* trinärutvecklingar. Exempelvis kan $\frac{1}{3}$ skrivas både som $(0.10000\cdots)_{\text{tre}}$ och som $(0.02222\cdots)_{\text{tre}}$. Konventionen att randpunkter behålles då $\mathcal{C}_k$ bildas från $\mathcal{C}_{k-1}$ innebär att sådana tal $a$ behandlas ”maximalt generöst”: det räcker att den ena av dess två trinärutvecklingar uppfyller (47) för att $a$ skall tillhöra $\mathcal{C}_k$. Motsvarande konvention gäller då naturligtvis även (48).

vart och ett av de fyra intervallen $I = [0, \frac{1}{9}], [\frac{2}{9}, \frac{1}{3}], [\frac{2}{3}, \frac{7}{9}], [\frac{8}{9}, 1]$. Detta är samma sak som att säga att de två första koefficienterna i $X$'s trinärutveckling skall vara oberoende och likafördelade enligt likformig fördelning på $\{0, 2\}$. Genom att fortsätta resonemanget på samma sätt inser vi att samtliga koefficienter i $X$'s trinärutveckling bör vara oberoende och anta värdena 0 eller 2 med sannolikhet $\frac{1}{2}$ vardera. Detta motiverar följande definition.

**Definition 8.1** *Låt $Y_1, Y_2, \ldots$ vara oberoende och likafördelade stokastiska variabler på $(\Omega, \mathcal{F}, \mathbb{P})$, och låt deras fördelning vara given av*

$$\mathbb{P}(Y_k = x) = \begin{cases} \frac{1}{2} & \textit{för } x = 0\,, \\ \frac{1}{2} & \textit{för } x = 2\,. \end{cases}$$

*Låt sedan*

$$X = \sum_{k=1}^{\infty} Y_k 3^{-k}\,.$$

*Fördelningen $\mu_X$ för $X$ kallas då* **Cantorfördelningen**.

För att formulera huvudresultatet om Cantorfördelningen är det lämpligt att göra ytterligare en definition.

**Definition 8.2** *Fördelningen $\mu_X$ för en stokastiska varibel $X$ sägs* **sakna diskret komponent** *om den inte har några atomer, dvs om*

$$\mu_X(\{x\}) = 0 \textit{ för varje } x \in \mathbb{R}\,.$$

*Fördelningen sägs* **sakna kontinuerlig komponent** *om det finns en mängd $A \in \mathcal{B}$ sådan att*

$$m(A) = 0 \quad \textit{och} \quad \mu_X(A) = 1\,. \tag{49}$$

*Om $\mu_X$ saknar såväl diskret som kontinuerlig komponent, så säger vi att $\mu_X$ är en* **singulär** *fördelning, och att $X$ är en* **singulär** *stokastisk variabel.*

Observera att en singulär stokastisk variabel $X$ inte har någon täthet $f_X$, ty om den hade det så skulle vi, med $A$ valt som i (49), få

$$\mu_X(A) = \mathbb{P}(X \in A) = \int_A f_X(x)\, dm(x) = 0$$

vilket ju skulle motsäga (49).

En diskret eller kontinuerlig stokastisk variabel är helt klart inte singulär. Ej heller är den stokastiska variabeln i Exempel 0.1 singulär, eftersom ju dess fördelning kan delas upp i en diskret och en kontinuerlig komponent. För Cantorfördelningen har vi däremot följande.

**Sats 8.1** *Cantorfördelningen är singulär.*

**Bevis.** Låt $\mu_X$ beteckna Cantorfördelningen. Vi beföver visa (i) att $\mu_X$ saknar diskret komponent, och (ii) att $\mu_X$ saknar kontinuerlig komponent. För att visa (i), låt $x \in \mathbb{R}$ vara godtycklig. Om $x \notin \mathcal{C}$ är saken klar, ty $\mu_X(\mathcal{C}) = 1$. Om å andra sidan $x \in \mathcal{C}$, så har $x$ en trinärutveckling $0.t_1t_2t_3\ldots$ sådan att $t_k \in \{0, 2\}$ för alla $k$. Om vi nu låter $Y_1, Y_2, \ldots$ och $X$ vara som i Definition 8.1, så har vi

$$\begin{aligned}\mu_X(\{x\}) &= \mathbb{P}(X = x) = \mathbb{P}(Y_1 = t_1, Y_2 = t_2, \ldots) \\ &= \prod_{k=1}^{\infty} \mathbb{P}(Y_k = t_k) = \prod_{k=1}^{\infty} \frac{1}{2} = 0\,,\end{aligned}$$

och (i) är bevisad.

För att visa (ii) räcker det att visa att $m(\mathcal{C}) = 0$. För varje $k$ gäller att $\mathcal{C}_k$ består av $2^k$ disjunkta intervall av längd $3^{-k}$ vardera, så att $m(\mathcal{C}_k) = (\frac{2}{3})^k$. Eftersom $\mathcal{C} \subset \mathcal{C}_k$ har vi $m(\mathcal{C}) \leq m(\mathcal{C}_k) = (\frac{2}{3})^k$, och eftersom $(\frac{2}{3})^k \to 0$ då $k \to \infty$ måste vi ha $m(\mathcal{C}) = 0$. $\square$

Del (ii) av ovanstående bevis kan ersättas med följande mer probabilistiska bevis för att $m(\mathcal{C}) = 0$, som å ena sidan är något mera omständligt, men som å andra sidan är mer användbart för generaliseringar (se Övning 8.4).

**Alternativt bevis av att $m(\mathcal{C}) = 0$.** Eftersom $\mathcal{C} \subset [0, 1]$, så är $m(\mathcal{C}) = m_{[0,1]}(\mathcal{C})$. Låt nu $Z$ vara en stokastisk variabel på $(\Omega, \mathcal{F}, \mathbb{P})$ som är likformigt fördelad på $[0, 1]$, dvs $Z$ har fördelning $m_{[0,1]}$. Låt vidare $W_1, W_2, \ldots$ vara koefficienterna i $Z$'s trinärutveckling. Då är $W_1, W_2, \ldots$ oberoende och likafördelade med $\mathbb{P}(W_i = 0) = \mathbb{P}(W_i = 1) = \mathbb{P}(W_i = 2) = \frac{1}{3}$ (se Övning 8.3). Då har vi $\sum_{k=1}^{\infty} \mathbb{P}(W_k = 1) = \infty$, så med stöd av Borel–Cantellis andra lemma har vi att $\mathbb{P}(W_k = 1 \text{ för något } k) = 1$, så att

$$\begin{aligned}m_{[0,1]}(\mathcal{C}) = \mathbb{P}(Z \in \mathcal{C}) &= \mathbb{P}(W_k \in \{0, 2\} \text{ för alla } k) \\ &= 1 - \mathbb{P}(W_k = 1 \text{ för något } k) = 0\,.\end{aligned}$$

$\square$

## Övningar.

8.1 Visa att Cantormängden är en Borelmängd.

8.2 Bestäm väntevärdet av en Cantorfördelad stokastisk variabel $X$. (Ledning: Skriv $X$ som $X = \sum_{k=1}^{\infty} Y_k 3^{-k}$, där $Y_1, Y_2, \ldots$ är som i Definition 8.1, och använd lämplig konvergenssats för att visa att $\mathbb{E}[X] = \sum_{k=1}^{\infty} 3^{-k}\mathbb{E}[Y_k]$.)

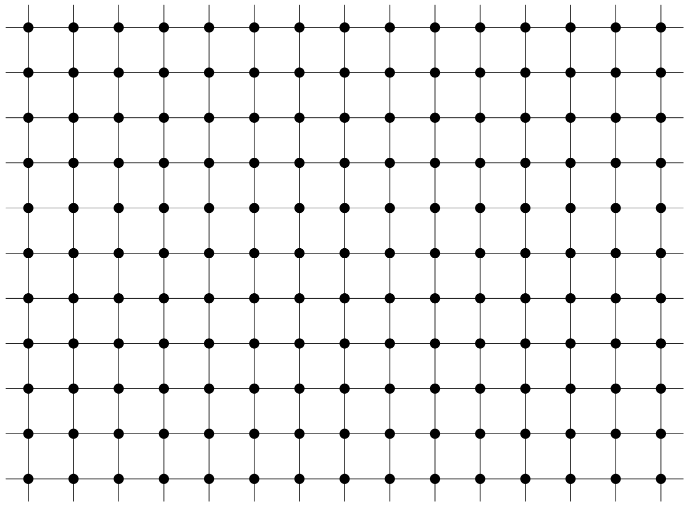
8.3 Låt den stokastiska variablen $X$ vara likformigt fördelad på $[0,1]$, och låt $Y_1, Y_2, \ldots$ vara decimalerna i decimalutvecklingen för $X$. Visa att $Y_1, Y_2, \ldots$ är oberoende och likafördelade, med fördelning $(\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10})$ på $\{0, 1, \ldots, 9\}$. Generalisera sedan resultatet till andra baser än 10.

8.4 Definiera en $[0,1]$-värd stokastisk variabel $X$ genom att låta dess decimaler vara oberoende och likafördelade enligt någon fördelning $\mu$ på $\{0, 1, \ldots, 9\}$. Vi konstaterade i föregående övning att om $\mu = (\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10})$, så är $X$ likformigt fördelad på $[0,1]$. Visa att varje annat val av $\mu$ medför att $X$ saknar kontinuerlig komponent. (Ledning: Resonera som i det alternativa beviset av $m(\mathcal{C}) = 0$, men använd Starka stora talens lag istället för Borel–Cantelli.)

## 8.2 Perkolation på $\mathbb{Z}^2$

I detta avsnitt och nästa skall vi kort beröra det forskningsområde inom sannolikhetsteorin som kallas perkolationsteori. Perkolationsteorin kan kort beskrivas som studiet av konnektivitetsegenskaper hos slumpmässiga grafer.

Låt oss börja med det exempel som ägnats de mest omfattande studierna: perkolation på heltalsgittret $\mathbb{Z}^2$. Betrakta den graf vars nodmängd utgörs av $\mathbb{Z}^2$ (dvs heltalspunkterna i planet), och i vilken två noder delar en kant om och endast om deras avstånd är 1. Se Figur 2.

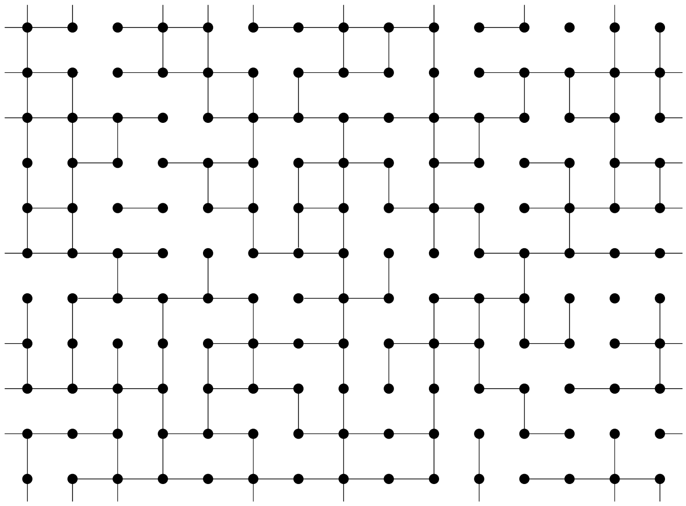


Figur 2: En bit av heltalsgittret $\mathbb{Z}^2$.

Låt $\mathcal{E}(\mathbb{Z}^2)$ beteckna mängden av kanter i denna graf. Antag nu att vi suddar ut kanter slumpmässigt. En lämplig modell för detta är följande. Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett lämpligt sannolikhetsrum. Till varje kant $e \in \mathcal{E}(\mathbb{Z}^2)$ associerar vi en stokastisk variabel $X_e$ på $(\Omega, \mathcal{F}, \mathbb{P})$. Vi fixerar $p \in [0,1]$, och låter $X_e$-variablerna vara oberoende av varandra och anta värden 0 och 1 med sannolikheter

$$\mathbb{P}(X_e = x) = \begin{cases} p & \text{för } x = 1\,, \\ 1-p & \text{för } x = 0\,. \end{cases}$$

Vi låter "$X_e = 1$" stå för att "kanten $e$ är närvarande" medan "$X_e = 0$" står för att "kanten $e$ är borta". Varje kant är alltså närvarande med sannolikhet $p$, oberoende av alla andra kanter. Hur detta kan te sig visas i Figur 3.



Figur 3: Perkolation på $\mathbb{Z}^2$ med $p = 0.6$.

Den slumpmässiga graf vi därvid erhåller har $\mathbb{Z}^2$ som nodmängd, medan kantmängden är en delmängd av $\mathcal{E}(\mathbb{Z}^2)$. Vi är intresserade av hur "sammanhängande" denna slumpgraf är. Vi kan dela upp den i maximala sammanhängande komponenter, som vi kallar **kluster**. Två noder $x, y \in \mathbb{Z}^2$ tillhör alltså samma kluster om och endast om det går att ta sig från $x$ till $y$ längs närvarande kanter. Den grundläggande frågan i perkolationsteorin är följande.

$$\text{Finns det något oändligt kluster?} \tag{50}$$

Den perkolationsmodell vi här har definierat, kan exempelvis ses som förenklad modell för genomsläppligheten i ett poröst material[9]. Tänk dig en porös sten[10], med ett oordnat system av mikroskopiska kanaler genomsläppliga för vatten. Hur tätt måste kanalerna ligga för att själva stenen skall bli genomsläpplig? Om kanalerna representeras av kanter i $\mathbb{Z}^2$, så krävs det mycket stora kluster för att det skall motsvara någon genomsläpplighet

[9] *The American Heritage Dictionary*, Houghton Mifflin 1982, anger att "percolate" betyder "to cause liquid, powder, or small particles to pass through a porous substance or small holes".

[10] Eftersom stenen är tredimensionell, bör nog $\mathbb{Z}^2$ i denna tillämpning helst ersättas med $\mathbb{Z}^3$. Perkolation på $\mathbb{Z}^3$ fungerar i princip på samma sätt som $\mathbb{Z}^2$-fallet.

på makroskopisk nivå. En lämplig matematisk idealisering är att anse materialet vara genomsläppligt om och endast om perkolationsmodellen uppvisar oändliga kluster.

Svaret på frågan (50) beror på $p$. För $p = 1$ är svaret ja, ty då är med sannolikhet 1 alla kanter närvarande, och hela $\mathbb{Z}^2$ blir ett enda oändligt kluster. För $p = 0$ är däremot svaret nej, eftersom då med sannolikhet 1 alla kanter är frånvarande, så att varje nod utgör ett eget kluster, och alla kluster är då naturligtvis ändliga. Men vad händer för värden på $p$ som ligger däremellan? Följande sats ger en del av svaret.

**Sats 8.2** *Låt $\pi(p)$ beteckna sannolikheten för existens av oändligt kluster för perkolation på $\mathbb{Z}^2$. Det existerar ett kritiskt värde $p_c = p_c(\mathbb{Z}^2) \in [0,1]$ sådant att*

$$\pi(p) = \begin{cases} 0 & om\ p < p_c \\ 1 & om\ p > p_c\,. \end{cases} \tag{51}$$

**Bevis.** Det räcker att bevisa att

(i) $\pi(p) \in \{0,1\}$ för varje $p \in [0,1]$, och

(ii) $\pi(p)$ är växande i $p$,

ty om (i) och (ii) gäller, så gäller även (51) med

$$p_c = \inf\{p : \pi(p) = 1\}\,.$$

Beviset av (i) sparas till Övning 8.5. För att visa (ii) skall vi använda att så kallat kopplingsargument[11]: vi definierar två perkolationsprocesser $\{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ och $\{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ på samma sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$. Fixera $p_1, p_2 \in [0,1]$ sådana att $p_1 < p_2$; vad som krävs är att vi visar att

$$\pi(p_1) \leq \pi(p_2)\,. \tag{52}$$

Låt $\{Z_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ vara oberoende likafördelade stokastiska variabler med fördelning given av

$$\mathbb{P}(Z_e = x) = \begin{cases} p_1 & \text{för } x = 2 \\ p_2 - p_1 & \text{för } x = 1 \\ 1 - p_2 & \text{för } x = 0\,. \end{cases}$$

[11]Kopplingsmetoden, som skördat stora framgångar i sannolikhetsteorin de senaste decenierna, går i stora drag ut på att definiera två eller flera stokastiska variabler eller stokastiska processer på ett och samma sannolikhetsrum, och, genom att på något sätt jämföra dem, dra slutsatser om deras respektive fördelningar.

För varje $e \in \mathcal{E}(\mathbb{Z}^2)$ definierar vi en stokastisk variabel $X'_e$ genom

$$X'_e = \begin{cases} 1 & \text{om } Z_e = 2 \\ 0 & \text{annars.} \end{cases}$$

Då är $\{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ oberoende likafördelade stokastiska variabler sådana att $\mathbb{P}(X'_e = 1) = p_1$ och $\mathbb{P}(X'_e = 0) = 1 - p_1$. Men detta är inget annat än fördelningen för $\{X_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ vid perkolation på $\mathbb{Z}^2$ med parameter $p_1$. Alltså har vi

$$\pi(p_1) = \mathbb{P}(\exists \text{ oändligt kluster i } \{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)})\,. \tag{53}$$

Vi definierar också, för varje $e \in \mathcal{E}(\mathbb{Z}^2)$,

$$X''_e = \begin{cases} 1 & \text{om } Z_e \geq 1 \\ 0 & \text{annars.} \end{cases}$$

Men då är $\{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ oberoende likafördelade stokastiska variabler sådana att $\mathbb{P}(X''_e = 1) = p_2$ och $\mathbb{P}(X''_e = 0) = 1 - p_2$, och vi får, med samma argument som för $\{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$, att

$$\pi(p_2) = \mathbb{P}(\exists \text{ oändligt kluster i } \{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)})\,. \tag{54}$$

Observera också att det för varje kant $e$ gäller att om $X'(e) = 1$, så är även $X''(e) = 1$, vilket är samma sak som att säga att varje kant som är närvarande i $\{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$ också är det i $\{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$. Detta betyder speciellt att om det finns ett oändligt kluster i $\{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$, så finns det ett sådant även i $\{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}$. Härav följer att

$$\mathbb{P}(\exists \text{ oändligt kluster i } \{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}) \leq \mathbb{P}(\exists \text{ oändligt kluster i } \{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)})\,. \tag{55}$$

Om vi nu kombinerar (53), (54) och (55) så får vi alltså

$$\begin{aligned} \pi(p_1) &= \mathbb{P}(\exists \text{ oändligt kluster i } \{X'_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}) \\ &\leq \mathbb{P}(\exists \text{ oändligt kluster i } \{X''_e\}_{e\in\mathcal{E}(\mathbb{Z}^2)}) \;=\; \pi(p_2) \end{aligned}$$

och (52) är bevisad. □

Observera att vi ingenstans i beviset utnyttjade några speciella egenskaper hos $\mathbb{Z}^2$. Alltså finns det, för varje oändlig graf $G$, något kritiskt värde $p_c = p_c(G)$, sådant att motsvarigheten till (51) gäller.

Nästa fråga är naturligtvis: Vilket värde har $p_c$? För $\mathbb{Z}^2$ ges svaret, $p_c = \frac{1}{2}$, i följande djupa resultat, som visades av Harry Kesten 1980.

**Sats 8.3** *För perkolation på $\mathbb{Z}^2$ med parameter $p$ gäller*

$$\mathbb{P}(\exists \textit{ oändligt kluster}) = \begin{cases} 0 & \textit{om } p \leq \frac{1}{2} \\ 1 & \textit{om } p > \frac{1}{2}\,. \end{cases}$$

Ett bevis av denna sats ligger utom räckhåll för oss, men vi kan komma en bit på vägen genom att visa följande resultat.

**Lemma 8.1** *Kritiska värdet* $p_c = p_c(\mathbb{Z}^2)$ *för perkolation på* $\mathbb{Z}^2$ *uppfyller*

$$p_c \geq \frac{1}{3}\,.$$

För beviset är det lämpligt att införa följande definition. Låt $\theta(p)$ beteckna sannolikheten att en given nod $x \in \mathbb{Z}^2$ ligger i ett oändligt kluster, när perkolationsparametern är $p$. $\theta(p)$ beror naturligtvis inte på valet av $x$, eftersom $\mathbb{Z}^2$ "ser likadant ut överallt".

Med en **väg** $\gamma$ av längd $k$ från en nod $x$ till en annan nod $y$ menar vi en sekvens $(e_1, \ldots, e_k)$ av kanter, och en sekvens $(v_1, \ldots, v_{k+1})$ av noder, sådana att (i) $v_1 = x$, (ii) $v_{k+1} = y$, (iii) för $i = 1, \ldots, k$ är $v_i$ och $v_{i+1}$ ändpunkter till kanten $e_i$, och (iv) ingen kant eller nod förekommer mer än en gång. En väg $\gamma$ sägs vara **öppen** om $X_e = 1$ för alla kanter $e$ i $\gamma$, dvs om alla kanter längs vägen är närvarande i perkolationsprocessen.

**Bevis av Lemma 8.1.** Fixera $p < \frac{1}{3}$. Vi behöver visa att $\pi(p) = 0$, vilket är detsamma som att $\theta(p) = 0$ (se Övning 8.6 (b)). Fixera också en nod $x \in \mathbb{Z}^2$. Låt $A$ vara händelsen att $x$ ligger i ett oändligt kluster, så att $\theta(p) = \mathbb{P}(A)$. För $k = 1, 2, \ldots$ låter vi $A_k$ vara händelsen att det finns någon öppen väg av längd $k$ utgående från $x$. Vi har då att

$$A_1 \supset A_2 \supset \cdots$$

och att

$$A = \bigcap_{k=1}^{\infty} A_k$$

vilket (se Övning 1.18 (b)) medför att

$$\mathbb{P}(A) = \lim_{k \to \infty} \mathbb{P}(A_k)\,.$$

Vi skall nu uppskatta $\mathbb{P}(A_k)$ med hjälp av förstamomentmetoden (Sats 5.2). Låt $Z_k$ vara antalet olika (men eventuellt överlappande) öppna vägar från $x$ av längd $k$, så att händelsen $A_k$ är detsamma som att $Z_k \geq 1$. Totala antalet vägar från $x$ av längd $k$ är högst $4 \cdot 3^{k-1}$, ty den första kanten kan väljas på 4 olika sätt, och därefter kan varje ytterligare kant väljas på högst 3 olika sätt. Varje väg av längd $k$ har sannolikhet $p^k$ att vara öppen. Därför är

$$\begin{aligned} \mathbb{E}[Z_k] &= \sum_{\gamma} \mathbb{E}[I_{\{\gamma \text{ är öppen}\}}] = \sum_{\gamma} \mathbb{P}(\gamma \text{ är öppen}) \\ &= \sum_{\gamma} p^k \leq 4 \cdot 3^{k-1} p^k = \frac{4(3p)^k}{3}\,, \end{aligned}$$

där summorna är över alla vägar av längd $k$ utgående från $x$. Enligt Sats 5.2 är $\mathbb{P}(Z_k \geq 1) \leq \mathbb{E}[Z_k]$, så att

$$\begin{aligned} P(A) &= \lim_{k\to\infty} \mathbb{P}(A_k) = \lim_{k\to\infty} \mathbb{P}(Z_k \geq 1) \\ &\leq \lim_{k\to\infty} \mathbb{E}[Z^k] \leq \lim_{k\to\infty} \frac{4(3p)^k}{3} = 0\,, \end{aligned}$$

ty $3p < 1$. □

I tre dimensioner känner man till att det kritiska värdet $p_c(\mathbb{Z}^3)$ är (strikt) lägre än i två dimensioner, men dess exakta värde förblir än så länge okänt. Detsamma gäller perkolation i fyra eller fler dimensioner.

För $p > \frac{1}{2}$ existerar alltså (minst) ett oändligt kluster. Kan man avgöra hur många det finns? Svaret ges i nästa sats.

**Sats 8.4** *För perkolation på $\mathbb{Z}^2$ med $p > \frac{1}{2}$ gäller*

$$\mathbb{P}(\exists \textit{ exakt ett oändligt kluster}) = 1\,.$$

Motsvarande resultat gäller även i högre dimensioner: För perkolation på $\mathbb{Z}^d$ ($d \geq 3$) med $p > p_c(\mathbb{Z}^d)$ existerar (med sannolikhet 1) exakt ett oändligt kluster.

## Övningar.

8.5 Betrakta perkolation på $\mathbb{Z}^2$ med parameter $p$. Argumentera, med hjälp av Kolmogorovs 0-1-lag, för att $\mathbb{P}(\exists$ oändligt kluster) är lika med 0 eller 1 för varje val av $p$.

8.6 Låt $\theta(p)$ vara som i beviset av Lemma 8.1.

- (a) Visa att $\theta(p)$ är växande i $p$.
- (b) Visa att $\theta(p) > 0$ om och endast om $\pi(p) = 1$.

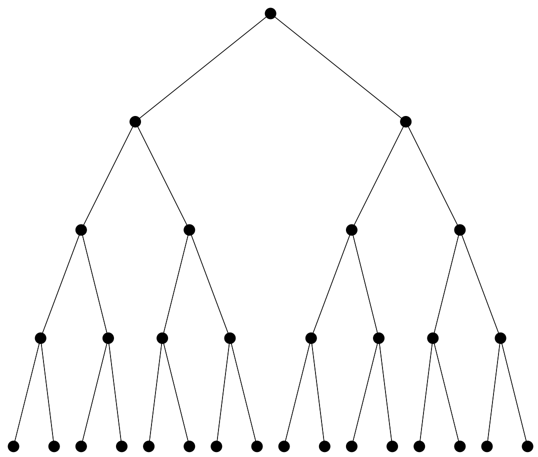
8.7 **Perkolation i en dimension är mindre intressant.** Betrakta perkolation på $\mathbb{Z}^1$, och definiera $\theta(p)$ analogt med $\mathbb{Z}^2$-fallet.

- (a) Visa att $\theta(p) = 0$ för varje $p < 1$.
- (b) Konkludera att $p_c(\mathbb{Z}^1) = 1$.

## 8.3 Perkolation på träd

Kanske ser du det som ett antiklimax att vi inte beräknade $p_c(\mathbb{Z}^2)$ i föregående avsnitt. Att visa Sats 8.3 hade emellertid krävt minst ett par dussin sidor av avancerade geometriska och probabilistiska argument. Som plåster på såren skall vi i detta avsnitt studera perkolation på en annan grafstruktur, vars kritiska värde ligger inom mer bekvämt räckhåll för oss att beräkna.

Ett **träd** är en sammanhängande graf utan cykler, så att det alltså finns exakt en väg att gå mellan varje par av noder $x$ och $y$. Vi skall studera perkolation på ett speciellt oändligt träd $\mathbb{T}_2$, känt som **det binära trädet**, som byggs upp på följande sätt. Vi startar med en nod $\rho$ som vi kallar **roten**. Från denna går två kanter till två olika noder, som sägs vara $\rho$'s barn. Var och en av dessa noder har, förutom kanten till $\rho$, en kant till vardera av sina respektive barn. Konstruktionen fortsätter sedan (i evighet) på ett sådant sätt att varje nod har exakt två barn. Mängden av noder på avstånd exakt $n$ till $\rho$ betecknas $\Gamma_n$, och kallas för den $n$:te generationen. Se Figur 4. Observera att $\Gamma_n$ innehåller $2^n$ noder.



Figur 4: De fem första generationerna $\Gamma_0, \Gamma_1, \ldots, \Gamma_4$ i det binära trädet $\mathbb{T}_2$.

Perkolation på $\mathbb{T}_2$ med parameter $p \in [0, 1]$ går till på samma sätt som perkolation på $\mathbb{Z}^2$: Varje kant $e \in \mathcal{E}(\mathbb{T}_2)$ associeras med en stokastisk variabel $X_e$, och vi låter $X_e$-variablerna vara oberoende av varandra och anta värden 0 och 1 med respektive sannolikheter $1 - p$ och $p$. Varje kant är alltså närvarande med sannolihet $p$, oberoende av alla andra kanter. För vilka $p$ får vi något oändligt kluster?

Vi har redan konstaterat att Sats 8.2 (existensen av kritiskt värde) fungerar för godtyckliga grafer. Tillämpat på $\mathbb{T}_2$ har vi alltså ett kritiskt värde $p_c = p_c(\mathbb{T}_2)$ sådant att perkolation på $\mathbb{T}_2$ ger något oändligt kluster med sannolikhet 1 om $p > p_c$, och med sannolikhet 0 om $p < p_c$. Om $\theta(p)$ betecknar sannolikheten att $\rho$ ligger i ett oändligt kluster, så är detta samma sak som att säga att

$$\theta(p)\begin{cases} = 0 & \text{för } p < p_c \\ > 0 & \text{för } p > p_c \end{cases}$$

(att så är fallet följer av en $\mathbb{T}_2$-motsvarighet till Övning 8.6 (b)).

**Sats 8.5** *Det kritiska värdet* $p_c = p_c(\mathbb{T}_2)$ *för perkolation på* $\mathbb{T}_2$ *ges av*[12]

$$p_c = \frac{1}{2}\,.$$

För beviset av denna sats behöves ytterligare en del notation och terminologi. För två noder $x, y$ i samma generation $\Gamma_n$, låt $x \wedge y$ beteckna den nod där vägarna från $x$ till $\rho$ och från $y$ till $\rho$ möts, så att med andra ord $x \wedge y$ är "den yngste gemensamme förfadern" till $x$ och $y$. Kalla $x$ och $y$ "$d$-männingar" om $x \wedge y$ återfinns $d$ generationer ovanför $x$ och $y$. 1-männingar kan alltså anses vara "syskon", 2-männingar är "kusiner", etc. Vår definition av $d$-männing innebär också att varje nod $x$ är sin egen 0-männing.

För två noder $x$ och $y$, låt $(x \leftrightarrow y)$ beteckna händelsen att $x$ och $y$ ligger i samma kluster, dvs att alla kanter på vägen mellan $x$ och $y$ är närvarande. Definiera också

$$(\rho \leftrightarrow \Gamma_n) = \bigcup_{x \in \Gamma_n} (\rho \leftrightarrow x)\,,$$

dvs $(\rho \leftrightarrow \Gamma_n)$ är händelsen att det finns en väg av öppna kanter från $\rho$ till någon nod i $\Gamma_n$. Låt även $(\rho \leftrightarrow \infty)$ vara händelsen att $\rho$ ligger i ett oändligt kluster.

**Bevis av Sats 8.5:** Observera att

$$(\rho \leftrightarrow \Gamma_1) \supset (\rho \leftrightarrow \Gamma_2) \supset \cdots$$

och att

$$(\rho \leftrightarrow \infty) = \bigcap_{n=1}^{\infty} (\rho \leftrightarrow \Gamma_n)\,.$$

Härav följer att

$$\mathbb{P}(\rho \leftrightarrow \infty) = \lim_{n\to\infty} \mathbb{P}(\rho \leftrightarrow \Gamma_n)$$

(se Övning 1.18 (b)). Vi behöver alltså visa att

(i) $\lim_{n\to\infty} \mathbb{P}(\rho \leftrightarrow \Gamma_n) = 0$ för $p < \frac{1}{2}$, och

[12] Att $\mathbb{Z}^2$ och $\mathbb{T}_2$ har samma kritiska värde bör ses som en ren tillfällighet, och kan *inte* tas som intäkt för att andra grafer också har kritiskt värde $\frac{1}{2}$. Exempelvis är ju, som vi påpekade i Avsnitt 8.2, $p_c(\mathbb{Z}^3) < \frac{1}{2}$. Se även Övning 8.8.

(ii) $\lim_{n\to\infty} \mathbb{P}(\rho \leftrightarrow \Gamma_n) > 0$ för $p > \frac{1}{2}$.

Vi börjar med att visa (i). Definiera, för $n = 1, 2, \ldots$, den stokastiska variablen $Z_n$ genom

$$Z_n = \sum_{x \in \Gamma_n} I_{(\rho \leftrightarrow x)}\,,$$

så att $Z_n$ alltså är antalet noder i $\Gamma_n$ som ligger i samma kluster som $\rho$. För $x \in \Gamma_n$ finns det bara en väg att gå från $\rho$ till $x$, och den vägen har innehåller $n$ kanter. Dessa kanter är oberoende, var och en med sannolikhet $p$ att vara närvarande, så $\mathbb{P}(\rho \leftrightarrow x) = p^n$. Eftersom antalet noder i $\Gamma_n$ är $2^n$ får vi därför

$$\begin{aligned}
\mathbb{E}[Z_n] &= \sum_{x \in \Gamma_n} \mathbb{E}[I_{(\rho \leftrightarrow x)}] = \sum_{x \in \Gamma_n} \mathbb{P}(\rho \leftrightarrow x) \\
&= \sum_{x \in \Gamma_n} p^n = 2^n p^n = (2p)^n\,.
\end{aligned}$$

Eftersom $Z_n$ är ickenegativ kan vi tillämpa Sats 5.2 (förstamomentmetoden) på $Z_n$, och vi får

$$\begin{aligned}
\mathbb{P}(\rho \leftrightarrow \Gamma_n) &= \mathbb{P}(Z_n \geq 1) \\
&\leq \mathbb{E}[Z_n] = (2p)^n\,.
\end{aligned} \tag{56}$$

Om nu $p < \frac{1}{2}$, så går (56) mot 0 då $n \to \infty$, och (i) är bevisad.

För att visa (ii) skall vi tillämpa andramomentmetoden (Sats 5.3), och då behöver vi förutom $\mathbb{E}[Z_n]$ även bestämma $\mathbb{E}[Z_n^2]$. Direkt räkning ger

$$\begin{aligned}
\mathbb{E}[Z_n^2] &= \mathbb{E}\left[\left(\sum_{x \in \Gamma_n} I_{(\rho \leftrightarrow x)}\right)^2\right] = \mathbb{E}\left[\sum_{x \in \Gamma_n} \sum_{y \in \Gamma_n} I_{(\rho \leftrightarrow x)} I_{(\rho \leftrightarrow y)}\right] \\
&= \sum_{x \in \Gamma_n} \sum_{y \in \Gamma_n} \mathbb{E}\left[I_{(\rho \leftrightarrow x)} I_{(\rho \leftrightarrow y)}\right] = \sum_{x \in \Gamma_n} \sum_{y \in \Gamma_n} \mathbb{P}(\rho \leftrightarrow x, \rho \leftrightarrow y)\,.
\end{aligned} \tag{57}$$

Vi behöver således bestämma $\mathbb{P}(\rho \leftrightarrow x, \rho \leftrightarrow y)$. För att händelsen $(\rho \leftrightarrow x, \rho \leftrightarrow y)$ skall inträffa, krävs att alla kanter i vägarna från $\rho$ till $x \wedge y$, från $x \wedge y$ till $x$, och från $x \wedge y$ till $y$, är närvarande. Om $x$ och $y$ är $d$-männingar, så har dessa vägar längder $n - d$, $d$ och $d$. De tre vägarna är också disjunkta, så de har totalt $(n - d) + d + d = n + d$ kanter, så att

$$\mathbb{P}(\rho \leftrightarrow x, \rho \leftrightarrow y) = p^{n+d}$$

Vidare har varje nod $x$ i $\Gamma_n$ exakt en 0-männing (sig själv), och, för $k = 1, \ldots, n$, exakt $2^{k-1}$ $k$-männingar. Antag att $p > \frac{1}{2}$. Om vi fortsätter beräkningen i (57), får vi

$$\mathbb{E}[Z_n^2] = \sum_{x \in \Gamma_n} \sum_{y \in \Gamma_n} \mathbb{P}(\rho \leftrightarrow x, \rho \leftrightarrow y)$$

$$
\begin{aligned}
&= \sum_{x\in\Gamma_n}\sum_{d=0}^{n}(\text{antal } d\text{-männingar till } x)\cdot p^{n+d} \\
&= \sum_{x\in\Gamma_n}\left(p^n+\sum_{d=1}^{n}2^{d-1}p^{n+d}\right) = 2^n\left(p^n+\sum_{d=1}^{n}2^{d-1}p^{n+d}\right) \\
&= (2p)^n\left(1+\frac{1}{2}\sum_{d=1}^{n}(2p)^d\right) = (2p)^n\left(1+\frac{p((2p)^n-1)}{2p-1}\right)
\end{aligned}
$$

där sista steget följer av formeln $\sum_{k=1}^{n} a^k = a\frac{a^n-1}{a-1}$ (med $a\neq 1$) för en geometrisk summa. Sats 5.3 säger oss att $\mathbb{P}(Z_n\geq 1)\geq\frac{\mathbb{E}[Z_n]^2}{\mathbb{E}[Z_n^2]}$, så vi får

$$
\begin{aligned}
\mathbb{P}(\rho\leftrightarrow\Gamma_n) &= \mathbb{P}(Z_n\geq 1)\geq\frac{\mathbb{E}[Z_n]^2}{\mathbb{E}[Z_n^2]} \\
&= \frac{(2p)^{2n}}{(2p)^n(1+\frac{p((2p)^n-1)}{2p-1})} \\
&= \frac{(2p-1)(2p)^n}{2p-1+p((2p)^n-1)},
\end{aligned}
$$

vilket (då $p>\frac{1}{2}$) går mot $\frac{2p-1}{p}$ då $n\to\infty$. Då är

$$
\lim_{n\to\infty}\mathbb{P}(\rho\leftrightarrow\Gamma_n)\geq\frac{2p-1}{p}>0
$$

och (ii) är bevisad. □

Vi har alltså visat att $\mathbb{P}(\rho\leftrightarrow\infty)=0$ då $p<\frac{1}{2}$, och att $\mathbb{P}(\rho\leftrightarrow\infty)>0$ då $p>\frac{1}{2}$. Men vad händer vid det kritiska värdet $p=\frac{1}{2}$? Sats 8.5 säger inget om detta. Med något mer arbete kan man visa att $\mathbb{P}(\rho\leftrightarrow\infty)=0$ då $p=\frac{1}{2}$.

Det kluster som innehåller $\rho$ kan ses som en så kallad **förgreningsprocess**, där $\rho$ är anfadern i en (enkönad) släkt där varje individ får 0, 1 eller 2 barn med respektive sannolikheter $(1-p)^2$, $2p(1-p)$ och $p^2$. Väntevärdet för antal barn är $2p$, och resultatet att

$$
\mathbb{P}(\rho\leftrightarrow\infty)>0 \;\text{ om och endast om }\; p>\frac{1}{2}
$$

betyder därmed att släkten har positiv sannolikhet för att leva vidare i evighet om och endast om förväntade antalet barn i en barnaskara är strikt större än 1. Detta resultat gäller även andra fördelningar för antalet barn[13].

[13]Det finns ett enda undantag, nämligen då varje individ får 1 barn med sannolikhet 1. Då är $\mathbb{E}[\text{antal barn}]=1$, samtidigt som släkten fortlever i evighet.

De båda exempel på perkolationsprocesser vi hittills koncentrerat oss på (perkolation på $\mathbb{Z}^2$ och på $\mathbb{T}_2$) delar egenskapen att oändliga kluster saknas vid det kritiska värdet. Man kan fråga sig om detta gäller för perkolation på godtyckliga grafstrukturer. De flesta perkolationsprocesser som dyker upp naturligt i olika sammanhang tycks ha denna egenskap[14], men svaret är ändå nej: det finns grafer som uppvisar oändliga kluster vid det kritiska värdet. Ett trivialt motexempel är perkolation på $\mathbb{Z}^1$ (se Övning 8.7), där $p_c = 1$ och det då givetvis blir ett oändligt kluster då $p = p_c$. Finns det då något motexempel med $p_c < 1$? Svaret är ja, och som exempel kan man ta följande modifiering av det binära trädet. Fixera $k$, och konstruera trädet $\mathbb{T}^*_{2,k}$ på samma sätt som det binära trädet $\mathbb{T}_k$, med undantag av att noder i generationerna $\Gamma_1, \Gamma_2, \Gamma_4, \Gamma_8, \Gamma_{16}, \ldots$ får $k$ barn istället för 2 (alla andra noder får som vanligt 2 barn). Detta träd visar sig vara tillräckligt "likt" $\mathbb{T}_2$ (generationerna som beter sig annorlunda blir ju allt ovanligare) för att det kritiska värdet $p_c$ skall bli detsamma, dvs $\frac{1}{2}$; detta kan visas genom att modifiera beviset av Sats 8.5. Men vid det kritiska värdet händer det mer förvånande saker: en noggrann analys visar att oändliga kluster förekommer vid det kritiska värdet $p = \frac{1}{2}$ om och endast om $k \geq 5$.

Vi avslutar (analogt med Sats 8.4) med frågan om hur många oändliga kluster man får vid perkolation på $T_2$ med $p > \frac{1}{2}$. För $p = 1$ får vi naturligtvis bara ett enda, men för $\frac{1}{2} < p < 1$ är svaret radikalt annorlunda än i $\mathbb{Z}^2$-fallet:

**Sats 8.6** *För perkolation på $\mathbb{T}_2$ med $p \in (\frac{1}{2}, 1)$, gäller att*

$$\mathbb{P}(\exists \text{ oändligt många oändliga kluster}) = 1\,.$$

Beviset av denna sats är faktiskt inte svårare än att det går att knäcka med de kunskaper du hittills inhämtat, plus en liten smula klurighet; se Övning 8.9.

## Övningar.

8.8 Det så kallade **trinära trädet** $\mathbb{T}_3$ definieras på samma sätt som $\mathbb{T}_2$, med skillnaden att varje nod har tre barn istället för två.

(a) Visa att $p_c(\mathbb{T}_3) = \frac{1}{3}$.

(b) Om $\mathbb{T}_k$, $k \geq 4$, konstrueras på motsvarande sätt, vad blir då den allmänna formeln för $p_c(\mathbb{T}_k)$?

8.9 Bevisa Sats 8.6. (Ledning: Det räcker att visa att $\mathbb{P}(\exists$ högst $M$ oändliga kluster$) \leq \varepsilon$ för godtyckliga $\varepsilon > 0$ och $M$. Detta kan göras genom att först visa att om $n$ är tillräckligt stort så är antalet frånvarande kanter i $\Gamma_n$ större än $M$ med sannoliket minst $1 - \varepsilon$, och sedan visa att det "under" varje sådan frånvarande kant finns något oändligt kluster med sannolikhet 1.)

[14] Simuleringsstudier och diverse heuristiska argument tycks indikera att detta (dvs frånvaro av oändliga kluster vid det kritiska värdet) gäller även på $\mathbb{Z}^d$ i högre dimensioner $d \geq 3$, men än så länge har man endast lyckats åstadkomma ett bevis för $d \geq 19$ (!)

# 9 LITTERATURTIPS

Som förkunskap till den kurs för vilken detta kompendium är avsett, duger i stort sett vilken bok som helst ur den breda floran av nybörjarböcker i sannolikhetsteori. Två böcker som ofta används är de av Ross [R] och Rudemo & Råde [RR]. Det kan vara en god idé att i början av kursen titta tillbaka i någon av dessa böcker, och repetera de centrala delarna.

Mycket av innehållet i kompendiet finns presenterat på ungefär samma nivå i böckerna av Brémaud [Bd] och Grimmett & Stirzaker [GS]. Dessa båda böcker kan passa bra som bredvidläsning för den som upplever vår framställning som varande bitvis alltför kortfattad.

På en aningen högre nivå finns de utomordentliga böckerna av Williams [W] och Durrett [D], som båda används flitigt i svensk doktorandutbildning i matematisk statistik. Ett äldre men fortfarande användbart alternativ är Breimans klassiska bok [B] från 1968, som kommit i ny upplaga på 90-talet. Folland [F] ger en introduktion till mått- och integrationsteorin, dock utan någon större tonvikt på sannolikheter.

Avsnitten om perkolation är kanske det mest originella inslaget i en text på vår nivå, och saknar motsvarigheter i exempelvis [Bd] och [GS]. För den som tänder till på innehållet i Avsnitt 8.2 och 8.3, och vill läsa mer, rekommenderar vi böckerna av Grimmett [G] och Lyons & Peres [LP].

Låt oss slutligen nämna den lättsamma boken av Blom, Holst och Sandell [BHS]. Från en helt elementär nivå tar den upp en mängd konkreta sannolikhetsproblem om exempelvis slumpvandringar, hasardspel, och ”kulor och urnor”, olika varianter av de klassiska födelsedags- och kupongsamlarproblemen, och det ökända ”bilen och getterna”-problemet. Denna bok rekommenderas för den som då och då känner för att ta en paus från den abstrakta teorin till förmån för mer handfasta problem.

□ □ □ □

Vi riktar ett varmt tack till Johan Jonasson och Jeff Steif, för rättelser och värdefulla synpunkter på manuskriptet.

# Referenser


[BHS] Blom, G., Holst, L. & Sandell, D. (1994) *Problems and Snapshots from the World of Probability*, Springer, New York.

[B] Breiman, L. (1992) *Probability. Corrected reprint of the 1968 original*, SIAM, Philadelphia.

[Bd] Brémaud, P. (1988) *An Introduction to Probabilistic Modeling*, Springer, New York.

[D] Durrett, R. (1991) *Probability: Theory and Examples*, Wadsworth & Brooks/Cole, Pacific Grove.

[F] Folland, G. (1984) *Real Analysis. Modern Techniques and Applications*, Wiley, New York.

[G] Grimmett, G. (1989) *Percolation*, Springer, New York.

[GS] Grimmett, G. & Stirzaker, D. (1992) *Probability and Random Processes*, Clarendon, Oxford.

[LP] Lyons, R. & Peres, Y. (1998) *Probability on Trees and Networks*, ännu ej färdigskriven bok, preliminär version tillgänglig på `http://php.indiana.edu/~rdlyons`

[R] Ross, S. (1998) *A First Course in Probability*, 5th edition, Macmillan, New York.

[RR] Rudemo M. & Råde, L. (1992) *Sannolikhetsteori och Statistik för Teknisk Högskola*, andra upplagan, Studentlitteratur, Lund.

[W] Williams, D. (1991) *Probability with Martingales*, Cambridge University Press.