Transformation af tætheder.
Ved brug af resultaterne i sidste afsnit kan man i 'pæne' situationer løse problemet vedrørende bestemmelse af fordelingen af en transformation af en given variabel med en kendt tæthed. Thi lad $\underline{X} = (X_1, \ldots, X_n)$ betegne en stokastisk vektor, som er absolut kontinuert m.h.t. λ_n med tæthed f, d.v.s. for alle $A \in \mathcal{B}(\mathbf{R}^n)$ er

$$P_{\underline{X}}(A) = P(\underline{X} \in A) = \int_A f \, d\lambda_n.$$

Man er nu ofte interesseret i en transformation af $\underline{X}$, d.v.s. en stokastisk vektor $\underline{Y}$ af formen $T(\underline{X})$, hvor $T : \mathbf{R}^n \to \mathbf{R}^n$ er en given målelig afbildning. Da vi ønsker at beskrive en situation, hvor $\underline{Y}$ igen har tæthed, vil vi antage, at T opfylder følgende betingelse:

$$\exists\, U \in \mathcal{B}(\mathbf{R}^n) : P_{\underline{X}}(U) = 1 \text{ og } \exists\, S : \mathbf{R}^n \to \mathbf{R}^n \text{ målelig: } S(T(\underline{x})) = \underline{x},\ \ \underline{x}\ \in U.$$

Bemærk at betingelsen specielt er opfyldt, hvis T er en bijektion med en målelig invers, f.eks. hvis T er en affin bijektion, thi her kan $U = \mathbf{R}^n$ og $S = T^{-1}$ bruges. Antagelsen og regneregler for billedmål sikrer for ethvert $A \in \mathcal{B}(\mathbf{R}^n)$, at

$$P_{\underline{Y}}(A) = P(\underline{Y} \in A) = P_{\underline{X}}(T^{-1}(A)) = P_{\underline{X}}(T^{-1}(A) \cap U)$$

$$\int_{T^{-1}(A)\cap U} f \, d\lambda_n = \int f(S \circ T) \cdot \mathbf{1}_A(T) \, d\lambda_{n,U} = \int_A f(S) \, d(\lambda_{n,U} \circ T^{-1}),$$

d.v.s. $P_{\underline{Y}}$ har tæthed $\underline{y} \mapsto f(S(\underline{y}))$ m.h.t. billedmålet $\lambda_{n,U} \circ T^{-1}$.
Det er derfor afgørende, om vi for et U, som opfylder ovenstående betingelse, er i stand til at bestemme $\lambda_{n,U} \circ T^{-1}$. Dette er f.eks. muligt, hvis T er en affin bijektion, d.v.s.

$$T : \underline{x} \mapsto \underline{a} + L(\underline{x}),$$

hvor $\underline{a} \in \mathbf{R}^n$ og L er en lineær bijektion i $\mathbf{R}^n$. Thi da T^{-1} her eksisterer som en kontinuert og derfor målelig funktion, kan U vælges lig $\mathbf{R}^n$, og da

$$\lambda_{n,U} \circ T^{-1} = \lambda_n \circ T^{-1} = |\det L|^{-1} \cdot \lambda_n$$

ifølge Proposition 20, viser dette, at $\underline{Y} = T(\underline{X})$ har tæthed (m.h.t. λ_n)

$$\underline{y} \mapsto |\det L|^{-1} \cdot f(T^{-1}(\underline{y})).$$

Dette 'lineære' eksempel er et specialtilfælde af flg. ofte forekommende mere generelle situation, hvor der findes en åben mængde $U \subseteq \mathbf{R}^n$, så at

$$P_{\underline{X}}(U) = 1 \text{ og } T \text{ er injektiv på } U \text{ af klasse } C^1 \text{ med } \det T'(\underline{x}) \neq 0 \text{ for alle } \underline{x} \in U.$$

I denne situation sikrer den såkaldte *open mapping theorem*, at $V := T(U)$ er en åben mængde i $\mathbf{R}^n$, samt at afbildningen $S : \mathbf{R}^n \to \mathbf{R}^n$ defineret ved

$$S(\underline{y}) := T^{-1}(\underline{y}) \;\; \underline{y} \in V \text{ og } \; S(\underline{y}) := \underline{0} \;\; \underline{y} \notin V$$

er målelig endog af type C^1 på V. Identifikationen af billedmålet $\lambda_{n,U} \circ T^{-1}$ følger nu af Sætning 13, som viser, at

$$\lambda_{n,U} \circ T^{-1} \ll \lambda_{n,V} \;\; \text{med tæthed} \;\; \underline{y} \mapsto |\det T'(S(\underline{y}))|^{-1},$$

Vi kan derfor med udgangspunkt i Sætning 13 formulere flg. vigtige resultat angående transformation af tætheder.

Lad U, V, T og S være som i Sætning 13 og lad $\underline{X}$ betegne en n-dimensional stokastisk vektor, som er absolut kontinuert m.h.t. λ_n med tæthed f og opfylder $P(\underline{X} \in U) = 1$. Da har $\underline{Y} := T(\underline{X})$ tæthed m.h.t. λ_n givet ved

$$\underline{y} \mapsto |\det T'(S(\underline{y}))|^{-1} \cdot f(S(\underline{y})) \cdot \mathbf{1}_V(\underline{y}).$$

Husk at $V = T(U)$ og at $S(\underline{y})$ for ethvert $\underline{y} \in V$ er bestemt ved identiteterne

$$S(\underline{y}) \in U \quad \text{og} \quad T(S(\underline{y})) = \underline{y}.$$

En simpel men vigtig anvendelse af den lineære transformationssætning viser, at hvis $X_1, \ldots, X_n$ er uafhængige $N(0,1)$-variable og T en lineær bijektion på $\mathbf{R}^n$, så er for ethvert $\underline{\mu} \in \mathbf{R}^n$

$$Y := \underline{\mu} + T(\underline{X}),$$

hvor $\underline{X} := (X_1, \ldots, X_n)$, absolut kontinuert med tæthed givet ved

$$\underline{y} \mapsto \frac{1}{\sqrt{(2\pi)^n \det \underline{\underline{\sigma}}}} \, exp\left(-\frac{1}{2} \sum_{1 \leq i,j \leq n} (y_i - \mu_i) \, \underline{\underline{\sigma}}^{-1}(i,j) \, (y_j - \mu_j)\right) \quad \underline{y} \in \mathbf{R}^n,$$

hvor $\underline{\underline{\sigma}}$ er $n \times n$-matricen svarende til den lineære afbildning $T \circ T^*$ beregnet m.h.t. den kanoniske basis i $\mathbf{R}^n$.

Sandsynlighedsfelter og stokastiske funktioner.
Som allerede nævnt i indledningen består et sandsynlighedsfelt $(\Omega, \mathcal{F}, P)$ af et måleligt rum $(\Omega, \mathcal{F})$ og et sandsynlighedsmål P på $(\Omega, \mathcal{F})$. Ω kaldes *udfaldsrummet* og elementerne i $\mathcal{F}$ *hændelser*. Værdien $P(A)$ af P på en hændelse A omtales som 'sandsynligheden for, at A indtræffer'.
Ofte er det ikke de enkelte udfald, der er interessante, men snarere en konsekvens heraf, d.v.s. en funktion, hvis værdimængde normalt er $\mathbf{R}$ eller $\mathbf{R}^n$, men mere generelle rum kan dog forekomme. Sådanne objekter kaldes i teorien stokastiske variable, stokastiske vektorer eller generelt stokastiske funktioner, og de betegnes med store bogstaver X, Y o.s.v. I overensstemmelse med at sandsynlighedsteoretiske modeller beskriver systemer, hvor man forud ikke nøjagtigt kan vide eller beregne, hvilket udfald der indtræffer, er det ikke de funktionsmæssige egenskaber, der har interesse. Det interessante er derimod størrelsen af sandsynlighederne for, at en given stokastisk funktion antager en værdier i en given pæn delmængde af værdimængden, d.v.s. sandsynligheder for forskellige urbilleder. Men da sandsynligheder kun kan tilordnes hædelser, har dette kun mening for $\mathcal{F}$-målelige funktioner, og det fører defor til flg. notation.
$X : \Omega \to \mathbf{R}$ kaldes en *stokastisk variabel*, hvis og kun hvis $X \in \mathrm{M}(\mathcal{F})$, og en $(\mathcal{F}, \mathcal{B}(\mathbf{R}^n))$-målelig afbildning $\underline{X} : \Omega \to \mathbf{R}^n$ kaldes en *n-dimensional stokastisk vektor*. Helt generelt kaldes en $(\mathcal{F}, \mathcal{E})$-målelig afbildning Y defineret på Ω med værdier i et måleligt rum $(E, \mathcal{E})$ en *stokastisk funktion*.
Ved brug af M 6 ses, at $\underline{X} : \Omega \to \mathbf{R}^n$ er en n-dimensional stokastisk vektor, hvis og kun hvis $\underline{X}$ er på formen

$$\underline{X} = (X_1, \ldots, X_n),$$

hvor X_i'erne er stokastiske variable. Disse kaldes ofte de til $\underline{X}$ hørende *koordinatvariable*.
Måleligheden bevirker, at man til enhver stokastisk funktion Y ved hjælp af billedmålsideen kan tilordne det såkaldte *fordelingsmål* P_Y defineret ved

$$P_Y := P \circ Y^{-1} \quad \text{d.v.s.} \quad P_Y(A) = P(Y \in A) \quad A \in \mathcal{E}.$$

Y er her en stokastisk funktion med værdier i et måleligt rum $(E, \mathcal{E})$. Specielt hvis X og $\underline{X}$ er h.h.v. en stokastisk variabel og en n-dimensional stokastisk vektor, så er P_X og $P_{\underline{X}}$ Borel sandsynlighedsmål på h.h.v. $\mathbf{R}$ og $\mathbf{R}^n$ givet ved

$$P_X(A) = P(X \in A) \quad A \in \mathcal{B}(\mathbf{R}) \quad \text{og} \quad P_{\underline{X}}(B) = P(\underline{X} \in B) \quad B \in \mathcal{B}(\mathbf{R}^n).$$

Fordelingsmålene indeholder altså præcis den information, vi søger, og man kan derfor med god ret sige, at en stokastisk funktion er fuldstændigt beskrevet ved det tilhørende fordelingsmål.
Ved at kombinere Lemma 3 og Proposition 2 ses, at *fordelingsfunktionerne*

$$\mathbf{R} \ni x \mapsto F_X(x) := P(X \le x) \quad \text{og} \quad \mathbf{R}^n \ni \underline{x} \mapsto F_{\underline{X}}(\underline{x}) := P(X_1 \le x_1, \ldots, X_n \le x_n)$$

entydigt bestemmer fordelingsmålene for en stokastisk variabel X og en n-dimensional stokastisk vektor $\underline{X}$.

Med udgangspunkt i fordelingsmålene inddeles de stokastiske funktioner i forskellige typer, hvoraf to er specielt vigtige. F.eks. siges en stokastisk variabel X at være *diskret*, hvis der findes en højst tællelig delmængde $E \subseteq \mathbf{R}$, så at

$$P(X \in E) = 1.$$

Fordelingsmålet P_X er i denne situation entydigt bestemt ved den såkaldte *diskrete tæthedsfunktion*

$$\mathbf{R} \ni x \mapsto P(X = x),$$

idet

$$P(X \in A) = \sum_{x \in A} P(X = x) := \sum_{i=1}^{\infty} \mathbf{1}_A(x_i) \cdot P(X = x_i)$$

for alle $A \in \mathcal{B}(\mathbf{R})$, hvor $(x_i)_{i \geq 1}$ er en nummerering af E. Begreberne diskrete stokastiske vektorer og funktioner defineres tilsvarende.

Simple egenskaber ved tællelige mængder viser, at hvis $\underline{X}$ er en n-dimensional stokastisk vektor med koordinatvariable $X_1, \ldots, X_n$, så er

$$\underline{X} \text{ diskret } \Leftrightarrow \ X_i \text{ diskret for } i = 1, \ldots, n,$$

og i givet fald gælder for ethvert i og ethvert $x \in \mathbf{R}$, at

$$P(X_i = x) = P(\underline{X} \in E_i^x) = \sum_{\underline{x} \in E_i^x} P(\underline{X} = \underline{x}) \ \text{ hvor } E_i^x = \{\underline{x} \in \mathbf{R}^n \,|\, x_i = x\}.$$

D.v.s. de marginale diskrete tæthedsfunktioner fremkommer ud fra den simultane diskrete tæthedsfunktion ved at 'summere de andre variable ud'.

Fordelingsmålet for en diskret stokastiske variabel er singulært m.h.t. Lebesgue målet. I modsætning hertil er den anden type, vi vil omtale, karakteriseret ved, at de tilhørende fordelingsmål er absolut kontinuerte m.h.t. Lebesgue målet. En stokastisk variabel X siges nemlig at være *absolut kontinuert*, hvis

$$P_X \ll \lambda_1$$

og den Radon-Nikodym afledede $dP_x/d\lambda_1$ kaldes i givet fald en *tæthed* for X. Tætheden, der som bekendt er entydig bestemt op til λ_1-nulmængder, bestemmer oplagt fordelingsmålet for en absolut kontinuert stokastisk variabel.

Det flerdimensionale begreb defineres analogt, men i modsætning til det diskrete tilfælde kan vi her kun slutte

$$\underline{X} \text{ absolut kontinuert } \Rightarrow \ X_i \text{ absolut kontinuert for } i = 1, \ldots, n.$$

hvor $\underline{X}$ er en n-dimensional stokastisk vektor med koordinatvariable $X_1, \ldots, X_n$. Den omvendte implikation er generelt ikke sand, hvilket f.eks. ses ved at betragte en stokastisk vektor af formen (X, X), hvor X er en absolut kontinuert stokastisk variabel. Men hvis en stokastisk vektor er absolut kontinuert, er det som ovenfor igen muligt at beregne de marginale tætheder ud fra den simultane. Resultatet er et specialtilfælde af flg. mere generelle resultat.

Hvis $\underline{X} = (X_1, \ldots, X_n)$ er absolut kontinuert med tæthed f, så er enhver marginal $(X_{n_1}, \ldots, X_{n_k})$ også absolut kontinuert, og en tæthed findes ved at integrere de ikke-relevante variable ud.

Nedenstående eksemplificerer bevismetoden. Antag at $\underline{X} = (X_1, \ldots, X_5)$ er absolut kontinuert med tæthed f, d.v.s.

$$f \in \overline{\mathrm{M}}_+(\mathcal{B}(\mathbf{R}^5)) \text{ og } P(\underline{X} \in A) = \int_A f \, d\lambda_5 \qquad A \in \mathcal{B}(\mathbf{R}^5).$$

Definer $\underline{U} := (X_1, X_3, X_5)$. D.v.s. $\underline{U}$ er en 3-dimensional stokastisk vektor, og vi vil nu vise, at den er absolut kontinuert med tæthed

$$\mathbf{R}^3 \ni \underline{u} \mapsto \int f(u_1, v_1, u_2, v_2, u_3) \, \lambda_2(d\underline{v}).$$

Betragt afbildningen

$$p : \mathbf{R}^5 \to \mathbf{R}^5 \quad p(\underline{x}) := (x_1, x_3, x_5, x_2, x_4).$$

p er kontinuert og derfor målelig med en målelig invers q givet ved

$$q : \mathbf{R}^5 \to \mathbf{R}^5 \quad q(\underline{x}) := (x_1, x_4, x_2, x_5, x_3).$$

En simpel overvejelse viser, at $\lambda_5 \circ p^{-1} = \lambda_5$ samt at

$$P(p(\underline{X}) \in A) = P(\underline{X} \in p^{-1}(A)) = \int_{p^{-1}(A)} f \, d\lambda_5$$

$$= \int f \circ q(p) \cdot \mathbf{1}_A(p) \, d\lambda_5 = \int_A f \circ q \, d\lambda_5 \circ p^{-1} = \int_A f \circ q \, d\lambda_5.$$

for alle $A \in \mathcal{B}(\mathbf{R}^5)$. D.v.s. $p(\underline{X})$ har tæthed $g := f \circ q$.

Da $(\mathbf{R}^5, \mathcal{B}(\mathbf{R}^5))$ kan opfattes som produktrummet $(\mathbf{R}^3 \times \mathbf{R}^2, \mathcal{B}(\mathbf{R}^3) \otimes \mathcal{B}(\mathbf{R}^2))$ og ligeledes $\lambda_5 = \lambda_3 \otimes \lambda_2$ følger af Lemma 9, at

$$\underline{u} \mapsto \int g(\underline{u}, \underline{v}) \, \lambda_2(d\underline{v}) \in \overline{\mathrm{M}}(\mathcal{B}(\mathbf{R}^3))_+,$$

hvor

$$g(\underline{u}, \underline{v}) := g(u_1, u_2, u_3, v_1, v_2) = f(u_1, v_1, u_2, v_2, u_3)$$

for $\underline{u} = (u_1, u_2, u_3)$ og $\underline{v} = (v_1, v_2)$. Men dette viser netop det ønskede, thi for alle $A \in \mathcal{B}(\mathbf{R}^3)$ gælder ifølge Tonelli's Sætning

$$P(\underline{U} \in A) = P(p(\underline{X}) \in A \times \mathbf{R}^2) = \int_{A \times \mathbf{R}^2} g \, d\lambda_5$$

$$= \int \Big\{ \int (g \cdot \mathbf{1}_{A \times \mathbf{R}^2})(\underline{u}, \underline{v})\, \lambda_2(d\underline{v}) \Big\}\, \lambda_3(d\underline{u}) = \int_A \Big\{ \int g(\underline{u}, \underline{v})\, \lambda_2(d\underline{v}) \Big\}\, \lambda_3(d\underline{u}). \quad \diamond$$

Ved brug af integralteorien indføres endvidere det såkaldte middelværdibegreb, idet en stokastisk variabel X siges at *have middelværdi* h.h.v. *have endelig middelværdi*, hvis

$$X \in L(P) \quad \text{h.h.v.} \quad X \in L^1(P)$$

og i givet fald defineres tallet

$$E[X] := \int X \, dP$$

og kaldes *middelværdien* af X. For stokastiske variable med endelig middelværdi er middelværdien derfor et reelt tal. Ved brug af den lille transformationssætning ses, at

$$X \in L(P) \text{ h.h.v. } L^1(P) \;\Leftrightarrow\; x \mapsto |x| \in L(P_X) \text{ h.h.v. } L^1(P_X)$$

og

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

hvis X har middelværdi. Formlen viser, at middelværdiens eksistens og værdi, som ventet, kun afhænger af fordelingsmålet. Men den muliggør også en eksplicit beregning; f.eks. hvis X er absolut kontinuert med tæthed f, er

$$E[X] = \int_{\mathbf{R}} x \cdot f(x) \, dx,$$

og hvis X er diskret, er

$$E[X] = \sum_{x \in \mathbf{R}} x \cdot P(X = x) := \sum_{i=1}^{\infty} x_i \cdot P(X = x_i),$$

hvor $(x_i)_{i \geq 1}$ er reelle tal, så at

$$\sum_{i=1}^{\infty} P(X = x_i) = 1.$$

Bemærk at for enhver ikke-negativ stokastisk variabel X er

$$E[X] = \int_0^\infty P(X > t) \, dt = \int_0^\infty (1 - F_X(t)) \, dt$$

ifølge integrationsformlen på side 50.

Uafhængighed.
Lad $(\Omega, \mathcal{F}, P)$ betegne et givent sandsynlighedsfelt. Som det er velkendt fra elementær sandsynlighedteori, siges hændelser $A_1, \dots, A_n$ at være uafhængige hvis

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$$

for alle valg af indices $1 \leq i_1 < \dots < i_k \leq n$. D.v.s. i alt $2^n - 1$ ligninger, hvoraf de n, svarende til $k = 1$, dog er trivielle. En simpel overvejelse viser, at betingelsen for uafhængighed ækvivalent kan formuleres som

$$P(\bigcap_{j=1}^{n} B_j) = \prod_{j=1}^{n} P(B_j)$$

hvor $B_j = A_j$ eller Ω for $j = 1, \dots, n$. Med baggrund heri siges hændelsessystemer $\mathcal{G}_1, \dots, \mathcal{G}_n$ indeholdt i $\mathcal{F}$ at være uafhængige, hvis

$$G_1, \dots G_n \text{ er uafhængige for alle valg af } G_j \in \mathcal{G}_j \text{ for } j = 1, \dots, n,$$

eller ækvivalent

$$P(\bigcap_{j=1}^{n} G_j) = \prod_{j=1}^{n} P(G_j) \quad \text{ for } G_j \in \mathcal{G}_j^* \; j = 1, \dots, n,$$

hvor $\mathcal{G}_j^* = \mathcal{G}_j \cup \{\Omega\}$ for alle j. Dette betyder specielt, at

$$\mathcal{G}_1, \dots, \mathcal{G}_n \text{ uafhængige} \Leftrightarrow \mathcal{G}_1^*, \dots, \mathcal{G}_n^* \text{ uafhængige.}$$

Mere generelt siges en vilkårlig familie $(\mathcal{G}_i)_{i \in I}$ af hændelssystemer i $\mathcal{F}$ at være uafhængige, hvis for ethvert $n \geq 1$

$$\mathcal{G}_{i_1}, \dots \mathcal{G}_{i_n} \text{ er uafhængige for alle valg af forskellige indices } i_1 \dots, i_n \in I.$$

Begrebet uafhængige σ-algebraer er specielt vigtigt, og i denne sammenhæng er det næste resultat af stor betydning.

Ua 1 *Lad $\mathcal{G}_1, \dots, \mathcal{G}_n$ betegne mængdesystemer i $\mathcal{F}$, som alle er $\cap f$-stabile. Da er σ-algebraerne $\sigma(\mathcal{G}_1), \dots, \sigma(\mathcal{G}_n)$ uafhængige, hvis og kun hvis*

$$P(\bigcap_{j=1}^{n} G_j) = \prod_{j=1}^{n} P(G_j) \qquad \text{for } G_j \in \mathcal{G}_j^* \quad j = 1, \dots, n,$$

d.v.s. $\sigma(\mathcal{G}_1), \dots, \sigma(\mathcal{G}_n)$ er uafhængige, hvis og kun hvis $\mathcal{G}_1, \dots, \mathcal{G}_n$ er uafhængige.
Bevis. Lad $G_k \in \mathcal{G}_k^*$ $k = 2, \dots, n$ være givet. Definer for $A \in \mathcal{F}$

$$\mu_1(A) := P(A \cap G_2 \cap \dots \cap G_n) \text{ og } \nu_1(A) := P(A) \cdot \prod_{k=2}^{n} P(G_k).$$

μ_1 og ν_1 er da endelige mål på $\mathcal{F}$, og pr. antagelse er de ens på $\mathcal{G}_1^*$ og dermed på $\sigma(\mathcal{G}_1)$ ifølge Proposistion 2. Vælg dernæst $G_1 \in \sigma(\mathcal{G}_1)$ og $G_k \in \mathcal{G}_k^*$ $k = 3, \ldots, n$ og definer som før for $A \in \mathcal{F}$

$$\mu_2(A) := P(G_1 \cap A \cap G_3 \cap \cdots \cap G_n) \text{ og } \nu_2(A) := P(G_1) \cdot P(A) \cdot \prod_{k=3}^{n} P(G_k).$$

Ifølge det netop viste er disse to endelige mål ens på $\mathcal{G}_2^*$ og derfor ifølge Proposistion 2 også på $\sigma(\mathcal{G}_2)$. Sådan fortsættes n-gange og vi ender op med, at

$$P(G_1 \cap \cdots \cap G_n) = \prod_{k=1}^{n} P(G_k)$$

for alle $G_k \in \sigma(\mathcal{G}_k)$ $k = 1, \ldots, n$, hvilket netop er det ønskede resultat. $\diamond$

For at kunne udnytte uafhængighedsbegrebet fuldt ud skal man vide, hvordan man opererer med uafhængige mængdesystemer. Resultatet, vi skal kende, siger løst sagt, at uafhængighed bevares, hvis man undgår at opbryde de enkelte systemer. For at budskabet ikke skal drukne helt i bar notation, formuleres det kun for en tællelig indeksmængde. Men det generelle tilfælde er åbenbar og overlades til læseren. Da mængdesystemet $\{\emptyset, \Omega\}$ er uafhængigt af alle mængdesystemer inklusiv sig selv, omfatter det formulerede resultat uden videre situationen, hvor udgangspunktet er endelig mange uafhængige mængdesystemer.

Ua 2 *Lad $(\mathcal{G}_i)_{i \geq 1}$ betegne en følge af uafhængige mængdesystemer i $\mathcal{F}$, som hver især er $\cap f$-stabile. Lad endvidere $(\alpha_j)_{j \geq 1}$ betegne en endelig eller tællelig familie af parvis disjunkte delmængder af $\{1, 2, \ldots, \}$. Da er $(\mathcal{H}_j)_{j \geq 1}$ uafhængige, hvor $\mathcal{H}_j := \sigma(\bigcup_{i \in \alpha_j} \mathcal{G}_i)$ for $j \geq 1$.*

Bevis. Som allerede nævnt kan vi antage, at $\Omega \in \mathcal{G}_i$ for alle i. Lad $j_1, \ldots, j_n$ være givet. Vi skal vise, at $\mathcal{H}_{j_1}, \ldots, \mathcal{H}_{j_n}$ er uafhængige. Ifølge Ua 1 er det nok at vise, at $\tilde{\mathcal{H}}_{j_1}, \ldots, \tilde{\mathcal{H}}_{j_n}$ er uafhængige, hvor

$$\tilde{\mathcal{H}}_{j_k} = \{ \bigcap_{i \in \beta_k} A_i \mid A_i \in \mathcal{G}_i \ i \in \beta_k \text{ hvor } \beta_k \subseteq \alpha_{j_k} \ k = 1, \ldots, n \text{ endelig delmængde} \},$$

thi $\tilde{\mathcal{H}}_{j_k}$'erne er hver for sig $\cap f$-stabile, og $\sigma(\tilde{\mathcal{H}}_{j_k}) = \mathcal{H}_{j_k}$ for alle k, da

$$\mathcal{G}_i \subseteq \tilde{\mathcal{H}}_{j_k} \quad i \in \alpha_{j_k}, \ k = 1, \ldots, n.$$

Specielt er $\Omega \in \tilde{\mathcal{H}}_{j_k}$ for alle k. Vi skal altså vise

$$P(\bigcap_{k=1}^{n} B_k) = \prod_{k=1}^{n} P(B_k)$$

for alle valg af $B_k \in \tilde{\mathcal{H}}_{j_k}$ $k = 1, \ldots, n$. Men for ethvert sådant valg er

$$\bigcap_{k=1}^{n} B_k = \bigcap_{k=1}^{n} \bigcap_{i \in \beta_k} G_i^k, \text{ hvor } \beta_k \subseteq \alpha_{j_k} \text{ endelig og } G_i \in \mathcal{G}_i \ i \in \beta_k \text{ for } k = 1, \ldots, n.$$

D.v.s. $\bigcap_{k=1}^n B_k$ er ligesom de enkelte B_k'er et endeligt gennemsnit af G_i'er stammende fra $\mathcal{G}_i$ for forskellige i. Antagelsen sikrer derfor, at

$$P(\bigcap_{k=1}^n B_k) = \prod_{k=1}^n \prod_{i\in\beta_k} P(G_i^k) \;\text{ og }\; P(B_k) = \prod_{i\in\beta_k} P(G_i^k) \quad k = 1,\ldots,n,$$

hvoraf resultatet umiddelbart følger. ◇

Med udgangspunkt i begrebet uafhængige mængdesystemer overføres uafhængighedsbegrebet til stokastiske variable og mere generelt stokastiske funktioner. Idet en familie $(X_i)_{i\in I}$ af målelige variable defineret på $(\Omega,\mathcal{F},P)$ med værdier i målelige rum $(E_i,\mathcal{E}_i)_{i\in I}$ siges at være uafhængige, hvis og kun hvis $(\sigma(X_i))_{i\in I}$ er uafhængige σ-algebraer.
Som en vigtig konsekvens af Ua 1 fås derfor, at hvis $\mathcal{E}_i = \sigma(\mathcal{G}_i)$ for $i \in I$, hvor $\mathcal{G}_i$ er $\cap f$-stabil for alle i, så er

$$(X_i)_{i\in I} \text{ uafhængige } \Leftrightarrow\; (X_i^{-1}(\mathcal{G}_i))_{i\in I} \text{ uafhængige},$$

hvor

$$X_i^{-1}(\mathcal{G}_i) = \{X_i^{-1}(B) \mid B \in \mathcal{G}_i\} \quad i \in I.$$

F.eks. er to stokastiske variable X og Y uafhængige, hvis og kun hvis

$$P(X \le x,\, Y \le y) = P(X \le x)\cdot P(Y \le y) \quad \text{for alle } x, y \in \mathbf{R}.$$

Der gælder et lignende resultat for stokastiske vektorer. D.v.s. for givne stokastiske variable $X_1,\ldots,X_n$ med marginale fordelingsfunktioner F_k $k = 1,\ldots,n$ og simultan fordelingsfunktion F gælder flg. biimplikation.

Ua 3

$$X_1,\ldots,X_n \text{ uafhængige } \Leftrightarrow F(\underline{x}) = \prod_{k=1}^n F_k(x_k) \;\text{ for alle } \underline{x} = (x_1,\ldots,x_n) \in \mathbf{R}^n.$$

Ligheden

$$P_X(A) = P(X\in A) \;\text{ og }\; P_{(X,Y)}(A\times B) = P(X\in A,\, Y\in B)$$

for $A \in \mathcal{E}$ og $B \in \mathcal{F}$, hvor X og Y er stokastiske funktioner med værdier i målelige rum $(E,\mathcal{E})$ og $(F,\mathcal{F})$, viser ved brug af Proposition 2 den vigtige biimplikation.

Ua 4

$$X \text{ og } Y \text{ uafhængige } \Leftrightarrow P_{(X,Y)} = P_X \otimes P_Y.$$

Der gælder et tilsvarende resultat for endelig eller endog tællelig mange variable.

2.18 Maximal Inequalities for Subadditive Schemes

Ottaviani's ulighed.
Lad $X_1, \ldots, X_n$ *betegne uafhængige stokastiske variable. Sæt* $M_n := \max_{1\leq j\leq n} |S_j|$ *hvor* $S_k := X_1 + \cdots + X_k$ *for* $k = 1, \ldots, n$. *For alle reelle tal* x *og* y *gælder da*

$$P(M_n > x + y) \cdot \min_{1\leq j\leq n} P(|S_n - S_j| \leq y) \leq P(|S_n| > x),$$

Korollar *Hvis* X_i*'erne yderligere har middelværdi* 0*, gælder for alle* $p \geq 1$

$$E[\, M_n^p] \leq 3 \cdot 2^p \cdot E[\, |S_n|^p].$$

Bevis for uligheden. Da uligheden er triviel, hvis enten x eller y er negativ, lader vi $x, y \geq 0$ være givet. Sæt

$D_1 = \{|S_1| > x+y\}$ og $D_j = \{|S_j| > x+y,\ |S_1| \leq x+y, \ldots, |S_{j-1}| \leq x+y\}$ $j \geq 2$.

Da D_j'erne er disjunkte og $\{M_n > x + y\} = \bigcup_{j=1}^n D_j$, er

$$P(M_n > x + y) = \sum_{j=1}^n P(D_j).$$

For ethvert j fås endvidere af trekantsuligheden at

$$\{|S_j| > x + y\} \subseteq \{|S_n| > x\} \cup \{|S_n - S_j| > y\}.$$

Heraf følger, da $(X_1, \ldots, X_j)$ og dermed D_j og $|S_n - S_j|$ er uafhængige, at

$$P(M_n > x + y) \leq \sum_{j=1}^n \left(P(\{|S_n| > x\} \cap D_j) + P(\{|S_n - S_j| > y\} \cap D_j)\right)$$

$$\leq \sum_{j=1}^n \left(P(\{|S_n| > x\} \cap D_j) + P(|S_n - S_j| > y) \cdot P(D_j)\right)$$
$$\leq P(|S_n| > x) + \max_{1\leq j\leq n} P(|S_n - S_j| > y) \cdot P(M_n > x + y),$$

hvoraf resultatet følger, da

$$1 - \max_{1\leq j\leq n} P(|S_n - S_j| > y) = \min_{1\leq j\leq n} P(|S_n - S_j| \leq y).$$

Bevis for Korollaret. Lad $p \geq 1$ være valgt. Da S_j og $S_n - S_j$ er uafhængige og centrerede fås af formel (4.5.6), at for $j = 1, \ldots, n$ er

$$E[\, |S_n|^p] = E[\, |S_n - S_j + S_j|^p] \geq E[\, |S_n - S_j|^p].$$

For $\tau := (2 \cdot E[\,|S_n|^p])^{1/p}$ gælder derfor ifølge Markov's Ulighed

$$P(|S_n - S_j| > \tau) \leq E[\,|S_n - S_j|^p]/\tau^p \leq E[\,|S_n|^p]/\tau^p = 1/2$$

og dermed

$$\min_{1 \leq j \leq n} P(|S_n - S_j| \leq \tau) \geq 1 - 1/2 = 1/2.$$

D.v.s.

$$P(M_n > x) = P(M_n > (x - \tau) + \tau) \leq 2 \cdot P(|S_n| > x - \tau) = 2 \cdot P(|S_n| + \tau > x)$$

for alle $x > 0$ og dermed ved integration

$$E[\,M_n^p] \leq 2 \cdot E[\,(|S_n| + \tau)^p] \leq 2^p \cdot E[\,|S_n|^p + \tau^p] \leq 3 \cdot 2^p \cdot E[\,|S_n|^p]. \quad \diamond$$

Ottaviani's Ulighed gælder for alle uafhængige stokastiske variable, men er variablene yderligere symmetriske, gælder med samme notation som ovenfor flg. mere præcise resultat.

Lévy's ulighed.
Lad $X_1, \ldots, X_n$ betegne uafhængige symmetriske stokastiske variable. Da er

$$P(M_n > t) \leq 2 \cdot P(|S_n| > t) \quad \text{for alle } t > 0,$$

og dermed $E[M_n^p] \leq 2 \cdot E[\,|S_n|^p]$ for alle $p > 0$.
Bevis. Lad $t > 0$ være givet. Sæt igen

$$D_1 = \{|S_1| > t\} \text{ og } D_j = \{|S_j| > x, |S_1| \leq t, \ldots, |S_{j-1}| \leq t\} \quad j \geq 2.$$

Da D_j'erne er disjunkte gælder som ovenfor

$$\begin{aligned} P(M_n > t) &= \sum_{j=1}^{n} P(|S_j| > t, D_j) = \sum_{j=1}^{n} P(|S_n + S_n^j| > 2t, D_j) \\ &\leq \sum_{j=1}^{n} P(|S_n| > t, D_j) + \sum_{j=1}^{n} P(|S_n^j| > t, D_j), \end{aligned}$$

hvor $S_n^j := X_1 + \cdots + X_j - (X_{j+1} + \cdots + X_n)$. Men da X_i'erne er symmetriske og uafhængige, er

$$(X_1, \ldots, X_n) \sim (X_1, \ldots, X_j, -X_{j+1}, \ldots, -X_n) \quad \text{for alle } j,$$

og dermed specielt

$$P(|S_n| > t, D_j) = P(|S_n^j| > t, D_j) \quad \text{for alle } j.$$

Indsættes dette fås

$$P(M_n > t) \leq 2 \sum_{j=1}^{n} P(|S_n| > t, D_j) \leq 2 \cdot P(|S_n| > t). \quad \diamond$$

3.40 The Uniqueness Theorem III

Vi skal kun bruge en lille del af dette meget generelle resultat, som behandler spørgsmålet, om en delmængde $\mathcal{H} \subseteq \mathrm{M}(\mathcal{E})$ på et måleligt rum $(E, \mathcal{E})$ er såkaldt *målseparerende* for en mængde m af endelige mål på $(E, \mathcal{E})$, d.v.s., hvornår der gælder

$$\int f\, d\mu = \int f\, d\nu \quad \forall f \in \mathcal{H} \;\Rightarrow\; \mu = \nu$$

for par af mål μ, $\nu \in m$. Et simpelt resultat af denne type er.

Lad (S, d) være et metrisk rum. Da er $bC(S)_+$ målseparerende for mængden af Borel sandsynlighedsmål på S.

Bevis. Lad μ og ν betegne to Borel sandsynlighedsmål på S. Ifølge Proposition 2 er det nok at vise, at

$$\mu(U) = \nu(U) \quad \text{for alle } U \subseteq S \text{ åben.}$$

Men dette fås af Monoton konvergens, thi betragt for givet U funktionerne

$$x \mapsto f_n(x) := (n \cdot d(x, U^c)) \wedge 1 \quad n \geq 1,$$

disse er klart ikke-negative, kontinuerte og begrænsede, og da $f_n \uparrow \mathbf{1}_U$ overalt har vi

$$\mu(U) = \lim_n \int f_n\, d\mu = \lim_n \int f_n\, d\nu = \nu(U). \quad \diamond$$

Som konsekvens heraf fås flg. resultat, som viser, at mængden

$$x \mapsto x^n \quad n \geq 1$$

er målseparerende for mængden af Borel sandsynlighedsmål på $\mathbf{R}$, som er koncentreret på $[0,1]$ ($[0,1]$ kan udskiftes med ethvert andet begrænset interval).

To Borel sandsynlighedsmål μ og ν på $\mathbf{R}$, som begge er koncentreret på $[0,1]$, d.v.s. $\mu([0,1]) = \nu([0,1]) = 1$, er ens, hvis

$$\int x^n\, \mu(dx) = \int x^n\, \nu(dx) \quad \text{for } n \geq 1.$$

D.v.s. et Borel sandsynlighedsmål μ på $\mathbf{R}$ koncentreret på $[0,1]$ er bestemt ved den tilhørende momentfølge, d.v.s. talfølgen $(m_n)_{n\geq 1}$, hvor for $n \geq 1$

$$m_n := \int x^n\, \mu(dx).$$

Bevis. Det er nok at vise den sidste påstand. $(m_n)_{n\geq 1}$ bestemmer ifølge integralets linearitet ethvert integral af typen $\int p\, d\mu$, hvor p er et polynomium, d.v.s. en funktion af formen

$$x \mapsto a_0 + a_1 \cdot x + \cdots + a_n \cdot x^n \;\; \text{hvor } a_0, \ldots, a_n \in \mathbf{R}.$$

Ifølge Lebegue's og Weierstrass - Bernstein's Sætning (se Exc. 4.29) bestemmer momentfølgen derfor også alle integraler $\int f\,d\mu$, hvor $f \in C(\mathbf{R})$, da enhver sådan funktion kan approksimeres uniformt på $[0,1]$ af en følge af polynomier. Resultatet er derfor en konsekvens af ovenstående generelle resultat. $\diamond$

Korollar *To Borel sandsynlighedsmål μ og ν på $\mathbf{R}$ begge koncentreret på $[0,\infty)$ er ens, hvis*

$$\int e^{-nx}\,\mu(dx) = \int e^{-nx}\,\nu(dx) \quad \text{for } n \geq 1.$$

D.v.s. et Borel sandsynlighedsmål μ koncentreret på $[0,\infty)$, er bestemt ved talfølgen $(l_n)_{n\geq 1}$, hvor

$$l_n := \int e^{-nx}\,\mu(dx) \quad \text{for } n \geq 1.$$

Bevis. Igen er det nok at vise den sidste påstand. Definer sandsynlighedsmålet $\tilde{\mu} := \mu \circ \psi^{-1}$, hvor ψ er den reelle funktion $x \mapsto e^{-x}$. $\tilde{\mu}$ er da koncentreret på $[0,1]$, da

$$\tilde{\mu}([0,1]) = \mu(\psi^{-1}([0,1])) = \mu([0,\infty)) = 1,$$

og ifølge den lille transformationssætning er

$$\int x^n\,\tilde{\mu}(dx) = \int e^{-nx}\,\mu(dx) = l_n \quad \text{for } n \geq 1$$

D.v.s. $(l_n)_{n\geq 1}$ bestemmer $\tilde{\mu}$ og dermed μ, da $\mu = \tilde{\mu} \circ \varphi^{-1}$, hvor $\varphi : \mathbf{R} \to \mathbf{R}$ er givet ved

$$\varphi(x) = -\log x \text{ for } x > 0 \text{ og } 0 \text{ ellers.}$$

Bemærk at $\varphi \circ \psi(x) \equiv x$. $\diamond$

4.18 Characteristic Functions

For ethvert Borel sandsynlighedsmål μ på $\mathbf{R}$ er $x \mapsto e^{itx} \in L^1(\mu, \mathbf{C})$ for alle $t \in \mathbf{R}$, d.v.s.

$$\hat{\mu}(t) := \int e^{itx}\, \mu(dx) \qquad t \in \mathbf{R}$$

er en vel defineret funktion på $\mathbf{R}$ med komplekse værdier. $\hat{\mu}$ kaldes den *Fourier transformerede* af μ. Tilsvarende defineres hvis ν er et Borel sandsynlighedsmål på $\mathbf{R}^n$ den Fourier transformerede af ν som

$$\underline{t} \mapsto \hat{\nu}(\underline{t}) := \int e^{i\underline{t}\cdot\underline{x}}\, \mu(d\underline{x}) \qquad \underline{t} \in \mathbf{R}^n,$$

hvor $\underline{t} \cdot \underline{x} = \sum_{j=1}^n t_j \cdot x_j$ er det sædvanlige skalarprodukt i $\mathbf{R}^n$.
Integralets linearitet og kontinuitet på $L^1(\mu, \mathbf{C})$ sikrer at for ethvert Borel sandsynlighedsmål μ på $\mathbf{R}^n$ er

Kf 1 $|\,\hat{\mu}(\underline{t})\,| \leq \hat{\mu}(\underline{0}) = 1 \quad \underline{t} \in \mathbf{R}^n$.

Kf 2 $\underline{t} \mapsto \hat{\nu}(\underline{t})$ er kontinuert.

Kf 3 $\hat{\mu}(-\underline{t}) = \overline{\hat{\mu}(\underline{t})} \quad \underline{t} \in \mathbf{R}^n$.

Endvidere viser Fubini's Sætning, at hvis μ og ν er Borel sandsynlighedsmål på h.h.v. $\mathbf{R}^n$ og $\mathbf{R}^m$, så gælder.

Kf 4

$$\int_{\mathbf{R}^m} \hat{\mu}(\underline{s})\, \nu(d\underline{s}) = \int_{\mathbf{R}^n} \hat{\nu}(\underline{t})\, \mu(d\underline{t}).$$

Kf 5 Den Fourier transformerede af $\mu \otimes \nu$ er givet ved

$$(\underline{t}, \underline{s}) \mapsto \hat{\mu}(\underline{t}) \cdot \hat{\nu}(\underline{s}) \quad (\underline{t}, \underline{s}) \in \mathbf{R}^n \times \mathbf{R}^m = \mathbf{R}^{n+m}.$$

Via fordelingsmålet overføres dette til stokastiske variable og vektorer, idet vi til enhver stokastisk variabel X h.h.v. stokastisk vektor $\underline{X}$ tilordner den såkaldte *karakteristiske funktion* φ_X h.h.v. $\varphi_{\underline{X}}$ defineret ved

$$\varphi_X := \hat{P}_X \quad \text{og} \quad \varphi_{\underline{X}} := \hat{P}_{\underline{X}}.$$

Ifølge den lille transformationssætning gælder altså

$$\varphi_X(t) = E[e^{itX}] \quad t \in \mathbf{R} \;\text{ og }\; \varphi_{\underline{X}}(\underline{t}) = E[e^{i\underline{t}\cdot\underline{X}}] \quad \underline{t} \in \mathbf{R}^n.$$

Oversættes ovenstående simple egenskaber til karakteristiske funktioner fås for enhver n-dimensional stokastisk vektor $\underline{X}$

Kf 1 $|\,\varphi_{\underline{X}}(\underline{t})\,| \leq \varphi_{\underline{X}}(\underline{0}) = 1 \quad \underline{t} \in \mathbf{R}^n$.

Kf 2 $\underline{t} \mapsto \varphi_{\underline{X}}(\underline{t})$ er kontinuert.

Kf 3 $\varphi_{\underline{X}}(-\underline{t}) = \overline{\varphi_{\underline{X}}(\underline{t})} = \varphi_{-\underline{X}}(\underline{t}) \quad \underline{t} \in \mathbf{R}^n$ og mere generelt

$$\varphi_{\underline{\underline{A}}\underline{X}+\underline{b}}(\underline{t}) = e^{i\underline{t}\cdot\underline{b}} \cdot \varphi_{\underline{X}}(\underline{\underline{A}}^t\underline{t}) \quad \underline{t} \in \mathbf{R}^m$$

hvor $\underline{b} \in \mathbf{R}^m$ og $\underline{\underline{A}}$ en $m \times n$ matrice og $\underline{\underline{A}}^t$ den tilhørende transponerede matrice.

Ligeledes bliver Kf 4 og Kf 5 for $\underline{X}$ og $\underline{Y}$ en h.h.v. n og m-dimensional stokastisk vektor til lighederne

Kf 4 $E[\,\varphi_{\underline{X}}(\underline{Y})\,] = E[\,\varphi_{\underline{Y}}(\underline{X})\,]$.

Kf 5 Hvis $\underline{X}$ og $\underline{Y}$ er uafhængige, er

$$\varphi_{(\underline{X},\underline{Y})}(\underline{t},\underline{s}) = \varphi_{\underline{X}}(\underline{t}) \cdot \varphi_{\underline{Y}}(\underline{s}) \quad (\underline{t},\underline{s}) \in \mathbf{R}^n \times \mathbf{R}^m,$$

samt hvis $n = m$

$$\varphi_{\underline{X}+\underline{Y}}(\underline{t}) = \varphi_{\underline{X}}(\underline{t}) \cdot \varphi_{\underline{Y}}(\underline{t}) \quad \underline{t} \in \mathbf{R}^n.$$

Udover disse mere eller mindre trivielle egenskaber gælder, hvad der er yderst vigtigt, den såkaldte entydighedssætning for karakteristiske funktioner d.v.s.

Entydighedssætningen for karakteristiske funktioner.
Ethvert Borel sandsynlighedsmål μ på $\mathbf{R}^n$ er bestemt ved $\hat{\mu}$, d.v.s. for enhver n-dimensional stokastisk vektor $\underline{X}$ bestemmer $\varphi_{\underline{X}}$ fordelingen for $\underline{X}$.
Dette kan ækvivalent udtrykkes som

$$\mu = \nu \;\Leftrightarrow\; \hat{\mu} = \hat{\nu} \quad \text{og} \quad \underline{X} \sim \underline{Y} \;\Leftrightarrow\; \varphi_{\underline{X}} = \varphi_{\underline{Y}},$$

hvor μ og ν er Borel sandsynlighedsmål på $\mathbf{R}^n$ og $\underline{X}$ og $\underline{Y}$ n-dimensionale stokastiske vektorer.

Bevis. Vi betragter kun det 1-dimensionale tilfælde, men beviset generaliserer ordret til højere dimensioner, hvis vi udnytter, at et endeligt Borel mål i $\mathbf{R}^n$ er bestemt ved dets værdier på Euklidiske kugler.
For et givet sandsynlighedsmål μ gælder for alle $a < b$, at hvis $c := (a+b)/2$ og $r := (b-a)/2$ så er

$$\mu(\,]\,a,b\,[\,) = \mu(\,]\,c-r,c+r\,[\,) = \mu(\phi_c^{-1}(\,[\,0,r^2\,[\,)) = \mu\circ\phi_c^{-1}(\,[\,0,r^2\,[\,),$$

hvor $\phi_c(x) := (x-c)^2 \; x \in \mathbf{R}$. D.v.s. $\hat{\mu}$ bestemmer μ, hvis $\hat{\mu}$ bestemmer sandsynlighedsmålene $\mu\circ\phi_c^{-1}$ for ethvert $c \in \mathbf{R}$.
Lad $c \in \mathbf{R}$ være givet. Da sandsynlighedsmålet $\mu\circ\phi_c^{-1}$ er koncentreret på $\mathbf{R}_+$ er det ifølge 3.40 bestemt ved tallene

$$\int_0^\infty e^{-nt}\,\mu\circ\phi_c^{-1}(dt) = \int_{-\infty}^\infty e^{-n(x-c)^2}\,\mu(dx) \quad n \geq 1.$$

Men for ethvert a afhænger

$$\int_{-\infty}^\infty \hat{\mu}(at)\,e^{-iatc}\,\frac{1}{\sqrt{2\pi}}\,e^{-t^2/2}\,dt,$$

kun af $\hat{\mu}$, og ifølge Fubini's Sætning er

$$\int_{-\infty}^{\infty} \hat{\mu}(at)\, e^{-iatc}\, \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{iatx}\, e^{-iatc}\, \mu(dx) \right\} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt$$

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{iat(x-c)}\, \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt \right\} \mu(dx) = \int_{-\infty}^{\infty} e^{-a^2(x-c)^2/2}\, \mu(dx)$$

Ved at sætte $a = \sqrt{2n}$ ses specielt, at $\hat{\mu}$ bestemmer tallene

$$\int_{-\infty}^{\infty} e^{-n(x-c)^2}\, \mu(dx) \qquad n \geq 1$$

og dermed μ. $\diamond$

Bemærk at beviset viser, at Entydighedssætningen kan skærpes til

Kf 6 *X og Y er identisk fordelte hvis og kun hvis $\varphi_X(t) = \varphi_Y(t)$ for alle $t \geq 0$.*

Advarsel. Der findes stokastiske variable X og Y, som <u>ikke</u> er identisk fordelte, men hvor $\varphi_X(t) = \varphi_Y(t)$ for alle t i et åbent interval omkring 0.

Kædes Entydighedssætningen sammen med Kf 5 fås flg. uafhængighedskriterium.

Kf 7 *Stokastiske variable $X_1, \ldots, X_n$ er uafhængige hvis og kun hvis*

$$\varphi_{(X_1,\ldots,X_n)}(\underline{t}) = \prod_{k=1}^{n} \varphi_{X_k}(t_k) \;\; \text{for alle } \underline{t} = (t_1, \ldots, t_n) \in \mathbf{R}^n.$$

Resltatet generaliserer umiddelbart til stokastiske vektorer. Se formel (4.18.11).

Til senere brug genopfriskes et par resultater om funktionen $t \mapsto e^{it}$. Vi starter med den simple men basale ulighed

$$|e^{it} - e^{is}| \leq 2 \wedge |t - s| \qquad t,\, s \in \mathbf{R},$$

hvilket for $s = 0$ giver $|e^{it} - 1| \leq 2 \wedge |t|$ for $t \in \mathbf{R}$.

For ethvert $n \geq 0$ har vi en Taylorapproksimation af n-te grad, hvor restleddet er defineret, så ligheden holder, d.v.s.

$$e^{it} = \sum_{k=0}^{n} \frac{i^k t^k}{k!} + r_n(t) \quad t \in \mathbf{R}.$$

Bemærk at $r_n(-t) = \overline{r_n(t)}$. Vi får vi brug for flg. vurderinger

$$|r_n(t)| \leq \frac{2|t|^n}{n!} \wedge \frac{|t|^{n+1}}{(n+1)!} \qquad n \geq 0,\, t \in \mathbf{R}.$$

Bevis. Da $|r_n(-t)| = |\overline{r_n(t)}| = |r_n(t)|$ behøver vi kun at betragte $t \geq 0$, Vi vil benytte induktion efter n. Da

$$|r_0(t)| = |e^{it} - 1| = |e^{it} - e^{i0}| \leq 2 \wedge |t| \;\; \text{for alle } t,$$

holder uligheden for $n = 0$. Lad $t \in \mathbf{R}_+$ være givet. Integreres Taylorapproksimationen af orden n fra 0 til t fås efter multiplikation med i

$$e^{it} - 1 = i\int_0^t e^{is}\,ds = i\int_0^t \sum_{k=0}^{n} \frac{i^k s^k}{k!}\,ds + i\int_0^t r_n(s)\,ds$$

$$= \sum_{k=0}^{n} \frac{i^{k+1}t^{k+1}}{(k+1)!} + i\int_0^t r_n(s)\,ds = \sum_{k=1}^{n+1} \frac{i^k t^k}{k!} + i\int_0^t r_n(s)\,ds.$$

Pr. definition af restleddet r_{n+1} giver dette derfor sammenhængen

$$r_{n+1}(t) = i\int_0^t r_n(s)\,ds \quad \text{og dermed } |r_{n+1}(t)| \le \int_0^t |r_n(s)|\,ds,$$

hvoraf induktionsskridtet umiddelbart følger. $\diamond$

For n lig 1 og 2 fås specielt vurderingerne

$$|\, e^{it} - 1 - it|/2 \le |t| \wedge \frac{|t|^2}{4} \quad \text{og} \quad |\, e^{it} - 1 - it + \frac{t^2}{2}| \le |t|^2 \wedge \frac{|t|^3}{6} \quad \text{for } t \in \mathbf{R},$$

og dermed for alle $\alpha \in [2,3]$ uligheden

$$|\, e^{it} - 1 - it + \frac{t^2}{2}| \le |t|^\alpha \quad \text{for } t \in \mathbf{R}.$$

Kombineres den basale ulighed $|e^{it} - e^{is}| \le |t-s|$ med reglerne for differentation af et integral afhængig af en reel parameter fås let formel (4.18.13), som viser en sammenhæng mellem eksistens af momenter og differentiabillitet af den karakteristiske funktion. Der gælder nemlig flg. udsagn for enhver stokastisk variabel X og ethvert $n \ge 1$.

Kf 8 *Hvis* $E[\,|X|^n] < \infty$ *er* $t \mapsto \varphi_X(t)$ *n-gange differentiabel, og for* $k = 1, \dots, n$ *er*

$$\varphi_X^{(k)}(a) = E[i^k X^k e^{iaX}] \quad \text{for } a \in \mathbf{R}, \text{ specielt } \varphi_X^{(k)}(0) = i^k \cdot E[X^k].$$

Det er værd at bemærke, at det omvendte ikke er rigtigt, idet der findes variable uden endelig middelværdi, hvis karakteristiske funktion er differentiabel overalt.

Ved brug af identiteten $e^{itx} - e^{iax} = e^{iax} \cdot (e^{i(t-a)x} - 1)$ har vi som ovenfor for ethvert $n \ge 0$ en n-te ordens Taylorapproksimation af typen

$$e^{itx} = \sum_{k=0}^{n} \frac{i^k x^k e^{iax}}{k!}(t-a)^k + \tilde{r}_n(t,x,a) \quad x,\, t,\, a \in \mathbf{R},$$

hvor $\tilde{r}_n(t,x,a) := e^{iax} \cdot r_n((t-a)x)$ med tilhørende restledsvurdering

$$|\tilde{r}_n(t,x,a)| \le \frac{2\,|(t-a)x|^n}{n!} \wedge \frac{|(t-a)x|^{n+1}}{(n+1)!} \quad n \ge 0,\, x,\, t,\, a \in \mathbf{R}.$$

Substitueres i denne vurdering x med X og tages middelværdi på begge sider fås under antagelsen om endeligt n'te moment formel (4.18.14), d.v.s.

$$\varphi_X(t) = \sum_{k=0}^{n} \frac{E[i^k X^k e^{iaX}]}{k!}(t-a)^k + E[\tilde{r}_n(t,X,a)] = \sum_{k=0}^{n} \frac{\varphi_X^{(k)}(a)}{k!}(t-a)^k + R_n(t,a)$$

for $n \geq 0$ og $t, a \in \mathbf{R}$, hvor $R_n(t,a) := E[\tilde{r}_n(t,X,a)]$. Specielt fås for $a = 0$ ifølge sammenhængen mellem momenter og afledede i 0 formel (4.18.17), d.v.s.

$$\varphi_X(t) = \sum_{k=0}^{n} \frac{i^k \cdot E[X^k]}{k!} t^k + R_n(t,0).$$

Vurderingerne på $\tilde{r}_n$ kan umiddelbart oversættes til R_n, og for $n \geq 0$ og $t, a \in \mathbf{R}$ har vi

$$|R_n(t,a)| \leq \frac{|t-a|^n}{n!} \cdot E[\,|X|^n \cdot \left(2 \wedge \frac{|(t-a) \cdot X|}{n+1}\right)] \leq \frac{2\,|t-a|^n}{n!} \cdot E[\,|X|^n].$$

D.v.s. hvis $E[\,|X|^n] < \infty$ fås for alle t og a af Lebesgue's sætning, at

$$\lim_{t \to a} E[\,|X|^n \cdot \left(2 \wedge \frac{|(t-a) \cdot X|}{n+1}\right)] = 0$$

og dermed implikationen

$$E[\,|X|^n] < \infty \Rightarrow \lim_{t \to a} \frac{R_n(t,a)}{(t-a)^n} = 0.$$

For $n = 2$ og $a = 0$ viser dette specielt, at hvis $E[X] = 0$ og $Var(X) = \sigma^2$, så er

$$\lim_{t \to 0} \frac{1 - \varphi_X(t)}{t^2} = \sigma^2/2,$$

og hvis X har endelig andet moment, så gælder for ethvert $\alpha \in [2,3]$ og $t \in \mathbf{R}$ ifølge det ovenstående uligheden

$$|\varphi_X(t) - 1 - it \cdot E[X] + t^2/2 \cdot E[X^2]| \leq |t|^\alpha \cdot E[\,|X|^\alpha].$$

Uligheden, som vi skal benytte i beviset for Lyapounov's Sætning, er selvfølgelig kun interessant for $X \in L^\alpha(P)$. Den kan opfattes som en generalisation af formel (4.18.17), der omhandler heltallige α.

Momentproblemet.
Lad X betegne en stokastisk variabel med momenter af enhver orden. *Momentfølgen* $(E[X^n])_{n\geq 1}$ er derfor en vel defineret reel talfølge bestemt ved fordelingen, og spørgsmålet om, den omvendt bestemmer fordelingen entydigt, er kendt under navnet *Momentproblemet*. Ifølge tillægget til Sektion 3.40 gælder det for begrænsede stokastiske variable, men da eksempler viser, at momentfølgen ikke altid bestemmer fordelingen entydigt, er man naturligvis interesseret i at vide, hvornår det er tilfældet. En simpel men dog interessant tilstrækkelig betingelse siger

$$(E[X^n])_{n\geq 1} \text{ bestemmer } P_X \text{ hvis } (*) : \sum_{n=0}^{\infty} \frac{E[\,|X|^n]}{n!} \cdot \rho^n < \infty \text{ for et } \rho > 0.$$

Vel kendt potensrækketeori viser, at $(*)$ ækvivalent kan formuleres som

$$\limsup_n \left(\frac{E[\,|X|^n]}{n!}\right)^{1/n} < \infty \quad \text{d.v.s.} \quad \exists\, c \in \mathbf{R}_+ : E[\,|X|^n] \leq c^n \cdot n! \quad n \geq 1.$$

Bevis. Antag $(*)$ holder, d.v.s. specielt, at $\lim_n \rho^n \cdot E[\,|X|^n]/n! = 0$ og dermed ifølge en af ulighederne ovenfor, at

$$\lim_n R_n(t,a) = 0 \quad \text{for alle } t, a \in \mathbf{R} \text{ forudsat at } |t-a| < \rho.$$

Ifølge ovenstående Taylor udvikling kan φ_X derfor rækkeudvikles omkring ethvert punkt a med en konvergensradius, som er større end ρ. Heraf kan resultatet nu vises, thi ved først at rækkeudvikle omkring 0 ses, at $(E[X^n])_{n\geq 1}$ bestemmer φ_X og dermed alle dens afledede i intervallet $]-\rho,\rho\,[$. Ved fornyet rækkeudvikling omkring punkter tæt ved ρ og $-\rho$ ses derfor, at dette også gælder i intervallet $]-2\rho, 2\rho\,[$. Sådan fortsættes og vi ser, at momentfølgen alt i alt bestemmer φ_X og dermed ifølge Entydighedssætningen fordelingsmålet P_X. $\diamond$

Bemærk at $(*)$ er ækvivalent med at $E[e^{\rho|X|}] < \infty$ for et $\rho > 0$, $(*)$ holder derfor specielt hvis $P(|X| < M) = 1$ for et $M \in \mathbf{R}_+$.

$(*)$ er en betingelse på de absolutte momenter, men der gælder flg. resultat.

Mp 1 *Lad X og Y være stokastiske variable med momenter af enhver orden, så at $E[X^k] = E[Y^k]$ for alle $k \geq 1$. Da holder $(*)$ for X, hvis og kun hvis $(*)$ holder for Y; og i givet fald er X og Y derfor identisk fordelte.*

Bevis. Antag at X opfylder $(*)$, d.v.s. $\exists\, c \in \mathbf{R}_+ : E[\,|X|^n] \leq c^n \cdot n! \quad n \geq 1$. Ifølge Cauchy-Schwarz's gælder derfor

$$E[\,|Y|^n] \leq \sqrt{E[Y^{2n}]} = \sqrt{E[X^{2n}]} \leq \sqrt{c^{2n} \cdot (2n)!} = c^n \cdot \sqrt{(2n)!}.$$

Men da $\sqrt{(2n)!} \leq 2^n \cdot n!$ for alle n ses, at Y også opfylder $(*)$. $\diamond$

Lad stadig X betegne en given stokastisk variabel. Da

$$e^{aX} \vee e^{-aX} \leq e^{a|X|} \leq e^{aX} + e^{-aX} \quad \text{for alle } a > 0,$$

følger det umiddelbart af sætningen om Monoton konvergens, at hvis vi med $\mathcal{R}(L_X)$ betegner mængden $\{t \in \mathbf{R} \mid E[e^{tX}] < \infty\}$, så gælder biimplikationen

$$X \text{ opfylder } (*) \;\Leftrightarrow\; \mathcal{R}(L_X) \text{ indeholder et åbent interval omkring } 0.$$

$\mathcal{R}(L_X)$ er altid et interval indeholdende 0, men det kan bestå af 0 alene eller have 0 som enten venstre eller højre endepunkt. Definer

$$M_X(t) := E[e^{tX}] \quad \text{for } t \in \mathcal{R}(L_X).$$

$M_X(\cdot)$ kaldes ofte den *momentfrembringende funktion*. Begrundelsen for dette er klar ud fra det ovenstående, thi hvis $\mathcal{R}(L_X)$ indeholder et åbent interval af formen $]-\epsilon, \epsilon[$, så har X momenter af enhver orden og

$$M_X(t) = \sum_{n=1}^{\infty} \frac{E[X^n]}{n!} \cdot t^n \quad \text{for } |t| < \epsilon.$$

Ifølge potensrækketeori er $t \mapsto M_X(t)$ derfor uendelig ofte differentiabel i 0 med

$$M_X^{(n)}(0) = E[X^n] \quad n \geq 1.$$

Heraf sluttes umiddelbart flg. resultat.

Mp 2 *Stokastiske variable X og Y er identisk fordelte, hvis $M_X(t) = M_Y(t) < \infty$ for alle t i et åbent interval omkring* 0.

Bemærkning. Da momenterne er bestemt som afledede i punktet 0, kan man forholdsvis nemt vise, at det er nok, at M_X og M_Y begge er endelige i en omegn af 0, og $M_X(t_n) = M_Y(t_n)$ for en følge $(t_n)_{n\geq 1}$, som konvergerer mod 0.

Et såkaldt 'målskifte' argument viser, at resultatet gælder uændret, uanset hvilket interval der er tale om.

'Bevis'. Antag $M_X(t) = M_Y(t) < \infty$ for alle $t \in]\lambda_1, \lambda_2[$, hvor $\lambda_1 < \lambda_2$. Lad for $\lambda_0 \in]\lambda_1, \lambda_2[$ Q_X og Q_Y betegne sandsynlighedsmålene på $(\Omega, \mathcal{F})$ givet ved

$$Q_X := e^{\lambda_0 X}/a\, dP \quad \text{og} \quad Q_Y := e^{\lambda_0 Y}/a\, dP, \text{ hvor } a = E[e^{\lambda_0 X}] = E[e^{\lambda_0 Y}].$$

Lad M_X^Q og M_Y^Q betegne de momentfrembringende funktioner for X under Q_X og Y under Q_Y. Reglerne for integration med hensyn til afledte mål viser, at M_X^Q og M_Y^Q er endelige og ens i intervallet $]\lambda_1 - \lambda_0, \lambda_2 - \lambda_0[$. Da dette interval indeholder 0, følger af det netop viste, at $Q_X \circ X^{-1} = Q_Y \circ Y^{-1}$. Specielt er

$$E[f(X) \cdot e^{\lambda_0 X}] = a \cdot E^{Q_X}[f(X)] = a \cdot E^{Q_Y}[f(Y)] = E[f(Y) \cdot e^{\lambda_0 Y}]$$

og dermed $E[f(X)] = E[f(Y)]$ for alle kontinuerte funktioner f med kompakt støtte, hvilket kun er muligt, hvis X og Y har samme fordeling. $\diamond$

Den flerdimensionale normalfordeling.
Som en simpel konsekvens af entydighedssætningen for og øvrige egenskaber ved karakteristiske funktioner genfinder vi flg. egenskab ved klassen af endimensionale normalfordelinger.

Hvis $X_1, \ldots, X_n$ er uafhængige normalfordelte stokastiske variable, er $\sum_{i=1}^n a_i X_i$ igen normalfordelt for ethvert valg af reelle konstanter $a_1, \ldots a_n$.

Med udgangspunkt heri indføres flg. definition.

Definition. En n-dimensional stokastisk vektor $\underline{X}$ siges at være *n-dimensional normalfordelt*, hvis

$$\underline{t} \cdot \underline{X} := \sum_{i=1}^{n} t_i X_i$$

er en normalfordelt stokastisk variabel for alle $\underline{t} = (t_1, \ldots, t_n) \in \mathbf{R}^n$.

Vælges $\underline{t}$ som en passende enhedsvektor ses, at alle koordinatvariable i en flerdimensional normalfordeling $\underline{X}$ er endimensionale normalfordelinger, d.v.s. de har både middelværdi og varians. Middelværdivektoren og kovariansmatricen

$$\mu_{\underline{X}} := (E[X_1], \ldots, E[X_n]) \quad \text{og} \quad \underline{\underline{\sigma}}_{\underline{X}} := \{Cov(X_i, X_j)\}_{1 \leq i, j \leq n}$$

er derfor vel definerede, og ligesom i det endimensionale tilfælde er en n-dimensional normalfordeling bestemt ved sin tilhørende middelværdivektor og kovariansmatrice. Der gælder nemlig flg. resultat.

N 1 *Hvis $\underline{X}$ og $\underline{Y}$ er n-dimensional normalfordelt med $\mu_{\underline{X}} = \mu_{\underline{Y}}$ og $\underline{\underline{\sigma}}_{\underline{X}} = \underline{\underline{\sigma}}_{\underline{Y}}$, så er $\varphi_{\underline{X}} = \varphi_{\underline{Y}}$, d.v.s. $\underline{X}$ og $\underline{Y}$ er identiske fordelte.*

Bevis. Lad $\underline{t} \in \mathbf{R}^n$ være givet. Da $\underline{t} \cdot \underline{X}$ og $\underline{t} \cdot \underline{Y}$ begge er normalfordelte, er de identisk fordelte, da de har samme middelværdi og varians, da

$$E[\underline{t} \cdot \underline{X}] = \underline{t} \cdot \mu_{\underline{X}} \quad \text{og} \quad Var(\underline{t} \cdot \underline{X}) = \sum_{1 \leq i, j \leq n} t_i \, \underline{\underline{\sigma}}_{\underline{X}}(i,j) \, t_j = \underline{t} \cdot \underline{\underline{\sigma}}_{\underline{X}} \cdot \underline{t}^*$$

og tilsvarende

$$E[\underline{t} \cdot \underline{Y}] = \underline{t} \cdot \mu_{\underline{Y}} \quad \text{og} \quad Var(\underline{t} \cdot \underline{Y}) = \underline{t} \cdot \underline{\underline{\sigma}}_{\underline{Y}} \cdot \underline{t}^*.$$

D.v.s.

$$\varphi_{\underline{X}}(\underline{t}) = E[exp(i\,(\underline{t} \cdot \underline{X}))] = E[exp(i\,(\underline{t} \cdot \underline{Y}))] = \varphi_{\underline{Y}}(\underline{t}).$$

for $\underline{t} \in \mathbf{R}^n$, og ifølge entydighedssætningen er $\underline{X}$ og $\underline{Y}$ derfor identisk fordelte. $\diamond$

Det har derfor mening, at tale om den n-dimensionale normalfordeling med middelværdi vektor $\underline{\mu}$ og kovariansmatrice $\underline{\underline{\sigma}}$, og vi vil i denne forbindelse skrive

$$\underline{X} \sim N_n(\underline{\mu}, \underline{\underline{\sigma}}),$$

hvis $\underline{X}$ er n-dimensional normalfordelt med $\underline{\mu}_{\underline{X}} = \underline{\mu}$ og $\underline{\underline{\sigma}}_{\underline{X}} = \underline{\underline{\sigma}}$. Flg. vigtige egenskaber ved flerdimensionale normalfordelinger er nu åbenbare. Som det er sædvane bruges samme notation for den lineære afbildning og den tilhørende matrice udregnet i h.h.t. den kanoniske basis.

N 2 *Klassen af flerdimensionale normalfordelinger er stabil under affine transformationer, d.v.s. hvis* $\underline{X} \sim N_n(\underline{\mu}, \underline{\underline{\sigma}})$ *og* $T : \mathbf{R}^n \to \mathbf{R}^m$ *lineær, så er*

$$\underline{Y} := \underline{y} + T(\underline{X}) \sim N_m(\underline{y} + T(\underline{\mu}),\, T \circ \underline{\underline{\sigma}} \circ T^*)$$

for ethvert $\underline{y} \in \mathbf{R}^m$*. Specielt er* $\underline{Y} \sim N_m(\underline{y}, T \circ T^*)$*, hvis* $\underline{X} \sim N_n(\underline{0}, \underline{\underline{I_n}})$*.*

Bevis. Da en linearkombination af koordinaterne i $\underline{Y}$ er en affin linearkombination af koordinaterne i $\underline{X}$, er $\underline{Y}$ m-dimensionalt normalfordelt. Resten følger nu ved beregning af den tilhørende middelværdivektor og kovariansmatrice. ◇

Regnerierne ovenfor viser sammen med entydighedssætningen, at

$$\underline{X} \sim N_n(\underline{\mu}, \underline{\underline{\sigma}}) \;\Leftrightarrow\; \varphi_{\underline{X}}(\underline{t}) = exp\Big(i\,(\underline{t} \cdot \underline{\mu}) - 1/2 \cdot \underline{t} \cdot \underline{\underline{\sigma}} \cdot \underline{t}^*\Big) \quad \underline{t} \in \mathbf{R}^n.$$

F.eks. er $\underline{X} \sim N_n(\underline{0}, \underline{\underline{I_n}})$, hvis $X_1, \ldots X_n$ er uafhængige $N(0,1)$-fordelte stokastiske variable. $\underline{\underline{I_n}}$ betegner her $n \times n$ enhedsmatricen. Men der gælder hvis og kun, for som vi nu skal vise, er uafhængighed og ukorellerethed det samme for simultant normalt fordelte variable.

N 3 *Hvis* $\underline{Z}$ *er en flerdimensional normalfordelt stokastisk vektor, er vilkårlige marginaler* $(Z_{n_1}, \ldots Z_{n_k})$ *og* $(Z_{m_1}, \ldots Z_{m_l})$ *uafhængige hvis og kun hvis*

$$Cov(Z_{n_i}, Z_{m_j}) = 0 \qquad \text{for alle } i = 1, \ldots, k \text{ og } j = 1, \ldots, l.$$

Hvis $\underline{X} \sim N_n(\underline{\mu}_1, \underline{\underline{\sigma}}_1)$ *og* $\underline{Y} \sim N_m(\underline{\mu}_2, \underline{\underline{\sigma}}_2)$ *er uafhængige, er* $(\underline{X}, \underline{Y}) \sim N_{n+m}(\underline{\mu}, \underline{\underline{\sigma}})$*, hvor*

$$\underline{\mu} = (\underline{\mu}_1, \underline{\mu}_2) \;\text{ og }\; \underline{\underline{\sigma}} = \begin{pmatrix} \underline{\underline{\sigma}}_1 & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{\sigma}}_2 \end{pmatrix}.$$

Beviset, der består i en gentagen anvendelse af ækvivalensen mellem uafhængighed og faktorisering af den karakteristiske funktion, d.v.s. Kf 7, overlades til læseren.

N 4 $\underline{X} \sim N_n(\underline{\mu}, \underline{\underline{\sigma}})$ *er absolut kontinuert, hvis og kun hvis* $\underline{\underline{\sigma}}$ *er invertibel, og i givet fald er en tæthed givet ved*

$$\underline{x} \mapsto \frac{1}{\sqrt{(2\pi)^n \, \det \underline{\underline{\sigma}}}}\; exp\left(-\frac{1}{2} \sum_{1 \le i,j \le n} (x_i - \mu_i)\, \underline{\underline{\sigma}}^{-1}(i,j)\, (x_j - \mu_j)\right) \qquad \underline{x} \in \mathbf{R}^n.$$

Bevis. Hvis $\underline{\underline{\sigma}}$ ikke er invertibel, findes der et $\underline{t} \in \mathbf{R}^n \setminus \{\underline{0}\}$, så at

$$Var(\underline{t} \cdot \underline{X}) = \underline{t} \cdot \underline{\underline{\sigma}} \cdot \underline{t}^t = 0.$$

Der findes derfor en konstant $c \in \mathbf{R}$, så $\underline{t} \cdot \underline{X} = c$ P-n.o. D.v.s.

$$P(\underline{X} \in A(\underline{t}, c)) = 1,$$

hvor $A(\underline{t}, c) = \{\underline{x} \in \mathbf{R}^n \,|\, \underline{t} \cdot \underline{x} = c\}$, hvilket er uforenelig med absolut kontinuitet, da ethvert ægte affint underrum i $\mathbf{R}^n$ har Lebesgue mål 0.
Hvis omvendt $\underline{\underline{\sigma}}$ er invertibel, kan den ifølge vel kendt teori skrives på formen

$$\underline{\underline{\sigma}} = T \cdot \underline{\underline{I_n}} \cdot T^t \quad \text{for } T : \mathbf{R}^n \to \mathbf{R}^n \text{ lineær bijektion.}$$

D.v.s. ifølge N 2

$$\underline{X} \sim \underline{\mu} + T(\underline{U}) \quad \text{hvor} \quad \underline{U} \sim N_n(\underline{0}, \underline{\underline{I_n}}).$$

Resten følger nu som tidligere vist af den lineære transformationssætning, da koordinatvariablene $U_1, \ldots, U_n$ i $\underline{U}$ er uafhængige $N(0,1)$-variable, og $\underline{U}$ derfor har tæthed

$$\underline{x} \mapsto (2\pi)^{-n/2} \, exp\,(-\|\underline{x}\|^2/2\,) = (2\pi)^{-n/2} \, exp\,(-\frac{1}{2} \sum_{i=1}^{n} x_i^2\,). \quad \diamond$$

De store tals love I.
Betegnelsen *De store tals love* dækker over et utal af resultater angående den asymptotiske opførsel af *empiriske gennemsnit*, d.v.s. variable af formen

$$\frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{eller mere generelt} \quad \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i),$$

med henblik på konvergens P-n.o. eller i sandsynlighed for $n \to \infty$. $(X_n)_{n\geq 1}$ er her en følge af stokastiske variable og $(\mu_n)_{n\geq 1}$ en reel talfølge. Der findes tilsvarende resultater for stokastiske vektorer $(\underline{X}_n)_{n\geq 1}$ og vektorer $(\underline{\mu}_n)_{n\geq 1}$. Hvis X_i'erne har endelig middelværdi, vælges μ_i normalt som middelværdien $E[X_i]$, og der er i denne situation dermed tale om normerede centrerede partialsummer. Resultaterne opdeles i to kategorier, idet der skelnes mellem *stærke* og *svage* love. En stærk lov er her et udsagn, der sikrer konvergens P-n.o. i modsætning til en svag lov, som vedrører konvergens i sandsynlighed. Da konvergens n.o. som bekendt medfører konvergens i sandsynlighed, giver enhver stærk lov anledning til en tilsvarende svag lov. Det absolut vigtigste resultat indenfor emnet er flg. klassiske stærke lov ofte omtalt som en af sandsynlighedsteoriens tre perler.

Kolmogorov's Store tals lov.
Hvis $(X_n)_{n\geq 1}$ er en følge af uafhængige identisk fordelte stokastiske variable med endelig middelværdi μ, konvergerer

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \quad P\text{-n.o. og i } L^1(P).$$

Da $E[X_n] = \mu$ for alle $n \geq 1$ kan påstanden ækvivalent formuleres som

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i]) \to 0 \quad P\text{-n.o. og i } L^1(P).$$

Resultatet spiller en enorm rolle i sandsynlighedsteorien, da det dukker naturligt op i mange sammenhænge. Men det er også af mere fundamental betydning for den moderne sandsynlighedsteori, d.v.s. Kolmogorov-modellen. For kunne et sådant resultat ikke vises, ville modellen simpelt hen være ubrugelig. Endvidere fremhæver det betydningen af det indførte middelværdibegreb, for som resultatet viser, konvergerer den empiriske middelværdi mod den teoretiske, hvis denne eksisterer, uanset hvilken fordeling der end er tale om.
I bestræbelserne på at bevise ovenstående sætning er der udviklet mange særdeles værdifulde teknikker, som udover at tjene deres oprindelige formål har muliggjort mange udvidelser af resultatet. Vi skal i det følgende beskæftige os med en lille del af denne omfattende teori, men det er vigtigt hele tiden at have ovenstående hovedresultat i tankerne.

Flg. spørgsmål fra den reelle analyse er tydeligvis af interesse:

Hvornår er $\lim_n \frac{1}{n}\sum_{i=1}^n a_i = 0$ for en given reel talfølge $(a_n)_{n\geq 1}$?,

d.v.s. hvornår konvergerer a_n mod 0 i Cecaro middel? Udover at vise at dette som bekendt gælder, hvis $a_n \to 0$ i sædvanlig forstand, indeholder sektion 4.9 to i denne forbindelse interessante resultater. Først og fremmest det såkaldte *Kronecker Lemma*, d.v.s. implikationen

$$\sum_{n=1}^{\infty} a_n/b_n \text{ konvergent i } \mathbf{R} \Rightarrow \lim_n \frac{1}{b_n}\sum_{i=1}^n a_i = 0,$$

hvor $0 < b_n < b_{n+1} \uparrow \infty$. Tilfældet $b_n \equiv n$ er specielt interessant.
Derudover vises også, at hvis a_n'erne enten er opad eller nedad begrænsede, så er

$$\lim_n \frac{1}{n}\sum_{i=1}^n a_i = 0 \quad \text{hvis} \quad \lim_n \frac{1}{[\lambda]^n}\sum_{i=1}^{[\lambda]^n} a_i = 0 \quad \text{for } \lambda > 1.$$

Til senere brug bemærkes, at det er nok, det holder for $\lambda = 1 + 1/k \;\; k \geq 1$.

Med baggrund heri åbner der sig derfor to mulige bevismetoder for ovenstående sætning. Enten kan den omformuleres til et spørgsmål om konvergens i $\mathbf{R}$ P-n.o. af den uendelige række

$$\sum_{n=1}^{\infty} (X_n - \mu)/n,$$

eller man kan først studere

$$\frac{1}{n}\sum_{i=1}^n (X_i - \mu)$$

langs med hurtigt voksende delfølger af formen $([\lambda^n])_{n\geq 1}$ for et $\lambda > 1$, og dernæst herudfra forhåbentligt deducere den ønskede konvergens for hele følgen.

Tilfældet, hvor X_i'erne er uafhængige, er af speciel interesse. Af afgørende betydning er her flg. resultat, som viser, at for summer af uafhængige variable er n.o.-konvergens det samme som konvergens i sandsynlighed.

Konvergens af summer af uafhængige variable.
Lad $(Z_n)_{n\geq 1}$ betegne en følge af uafhængige variable. Da gælder

$$\lim_n \sum_{i=1}^n Z_i \text{ er summabel } P\text{-n.o.} \Leftrightarrow \sum_{n=1}^{\infty} Z_n \text{ konvergent i sandsynlighed,}$$

hvor 'summabel P-n.o.' betyder at $\sum_{n=1}^{\infty} Z_n(\omega)$ er konvergent i $\mathbf{R}$ for P-n.a. ω.

Korollar *For alle $p > 0$ gælder*

$$\sum_{n=1}^{\infty} Z_n \;\text{ konvergent i } L^p(P) \Rightarrow \sum_{n=1}^{\infty} Z_n \;\text{ er summabel } P\text{-n.o.}$$

Korollaret, der er en umiddelbar konsekvens af sætningen, da konvergens i L^p medfører konvergens i sandsynlighed, er interessant, fordi konvergens i L^p ofte er simpelt at eftervise. F.eks. deduceres uden problemer flg. stærke lov.

De store tals lov (L^2-udgave).
Lad $(X_n)_{n\geq 1}$ betegne en følge af uafhængige kvadratisk integrable stokastiske variable. Da gælder

$$\sum_{n=1}^{\infty} Var(X_n)/n^2 < \infty \Rightarrow \frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i]) \to 0 \quad P\text{-n.o. og i } L^2(P).$$

Bevis. Uafhængigheden bevirker, at $(X_n - E[X_n])_{n\geq 1}$ udgør en orthogonal følge i L^2, og da

$$\|X_n - E[X_n]\|_2^2 = Var(X_n) \text{ for alle } n \geq 1$$

fås af Pythagoras, d.v.s. Lemma 13, at

$$\sum_{n=1}^{\infty} Var(X_n)/n^2 < \infty \Rightarrow \sum_{n=1}^{\infty}(X_n - E[X_n])/n \text{ konvergerer i } L^2(P).$$

Konvergensen P-n.o. følger nu af ovenstående korollar samt Kroneckers Lemma, og da

$$E[\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i)\right)^2] = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)$$

følger konvergensen i L^2 også af Kronecker Lemmaet. $\diamond$

Bemærkning. Da beviset udnytter begrebet orthogonalitet, er det på ingen måde klart, at resultatet kan generaliseres til andre eksponenter $\alpha > 1$. Men som vi senere skal se, er dette dog muligt.

Bevis for resultatet angående konvergens af summer af uafhængige variable:
Lad for ethvert $n \geq 1$ S_n betegne $\sum_{i=1}^{n} Z_i$ og lad S betegne grænsevariablen, d.v.s. $S_n \to S$ i sandsynlighed. Der findes derfor en delfølge $(n_k)_{k\geq 1}$, så at $S_{n_k} \to S$ n.o. for $k \to \infty$. Definer for $k \geq 1$

$$M_k := \max_{n_{k-1} < l \leq n_k} |S_l - S_{n_{k-1}}|,$$

hvor $n_0 = 0$ og $S_0 := 0$. Ifølge Ottaviani's ulighed gælder for alle $\epsilon > 0$ og $k \geq 1$

$$P(M_k > 2\epsilon) \cdot (1 - \max_{n_{k-1} < l \leq n_k} P(|S_{n_k} - S_l| > \epsilon)) \leq P(|S_{n_k} - S_{n_{k-1}}| > \epsilon).$$

Men da $|S_{n_k} - S_{n_{k-1}}| \to 0$ P-n.o. er mængden

$$\{k \geq 1 \,|\, |S_{n_k}(\omega) - S_{n_{k-1}}(\omega)| > \epsilon\}$$

endelig for P-n.a. ω, og da $|\, S_{n_k} - S_{n_{k-1}}|$'erne yderligere er uafhængige, følger af Det andet Borel - Cantelli Lemma at

$$\sum_{k=1}^{\infty} P(|S_{n_k} - S_{n_{k-1}}| > \epsilon) < \infty.$$

Ved brug af trekantsuligheden fås, da $S_n \to S$ i sandsynlighed, at

$$\max_{n_{k-1} < l \leq n_k} P(|S_{n_k} - S_l| > \epsilon) \leq 2 \cdot \sup_{l > n_{k-1}} P(|S - S_l| > \epsilon/2) \to_{k \to \infty} 0,$$

og indsættes dette ses, at

$$\sum_{k=1}^{\infty} P(M_k > 2\epsilon) < \infty \quad \text{for alle } \epsilon > 0,$$

og dermed $M_k \to 0$ P-n.o. ifølge Det første Borel - Cantelli Lemma. For n.a. ω konvergerer altså for $k \to \infty$

$$S_{n_k}(\omega) \to S(\omega) \text{ og } M_k(\omega) \to 0 \text{ og dermed } \lim_l S_l(\omega) = S(\omega),$$

for med $k(l)$ bestemt ved $n_{k(l)-1} < l \leq n_{k(l)}$ gælder for alle l uligheden

$$\begin{aligned} |S_l(\omega) - S(\omega)| &\leq |S_l(\omega) - S_{n_{k(l)-1}}(\omega)| + |S_{n_{k(l)-1}}(\omega) - S(\omega)| \\ &\leq M_{k(l)}(\omega) + |S_{n_{k(l)-1}}(\omega) - S(\omega)|. \quad \diamond \end{aligned}$$

Bevis for Kolmogorov's Store tals lov:
Lad $(X_n)_{n \geq 1}$ betegne en følge af uafhængige identisk fordelte stokastiske variable med endelig middelværdi μ. Vi skal vise, at

$$\frac{1}{n} \sum_{i=1}^{n} X_i \to \mu \quad P\text{-n.o.}$$

Som netop vist, findes der et relevant resultat i det kvadratisk integrable tilfælde, men da vi her kun antager integrabilitet, får vi brug for den såkaldte *trunkerings-teknik*. De enkelte variable skrives her som en sum af to i h.h.t. flg. ide

$$X_n = X_n' + X_n'' \text{ hvor } X_n' := X_n \cdot \mathbf{1}_{[-a_n, a_n]}(X_n) \text{ og } X_n'' := X_n - X_n' = X_n \cdot \mathbf{1}_{\{|X_n| > a_n\}}$$

for et passende valg af positive reelle tal a_n. Da X_n'erne er forudsat integrable sættes $a_n = n$, idet der da gælder

$$\sum_{n=1}^{\infty} P(X_n'' \neq 0) = \sum_{n=1}^{\infty} P(|X_n| > n) = \sum_{n=1}^{\infty} P(|X_1| > n) < \infty.$$

Ved brug af Det første Borel-Cantelli Lemma fås derfor at

$$P(\,\exists\, n \geq 1 :\ X_i^{"} = 0 \ \ i \geq n) = 1$$

og dermed

$$\frac{1}{n}\sum_{i=1}^{n} X_i^{"} \to 0 \ P\text{-n.o.},$$

og da

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n} X_i^{'} + \frac{1}{n}\sum_{i=1}^{n} X_i^{"},$$

mangler vi kun at eftervise at første led konvergerer P-n.o. mod μ. Men da

$$Var(X_n^{'}) \leq E[X_n^{'2}] = E[X_n^2 \cdot \mathbf{1}_{[-n,n]}(X_n)] = E[X_1^2 \cdot \mathbf{1}_{[-n,n]}(X_1)],$$

d.v.s. $Var(X_n^{'}) \leq E[X_1^2, |X_1| \leq n\,]$, og der findes en konstant $C \in \mathbf{R}_+$, så at

$$\sum_{n=1}^{\infty} E[X_1^2, |X_1| \leq n\,]/n^2 = E[X_1^2 \cdot \sum_{n,\, n \geq |X_1| \vee 1} 1/n^2] \leq C \cdot E[\,|X_1|] < \infty,$$

følger af ovenstående L^2-resultat, at

$$\frac{1}{n}\sum_{i=1}^{n} X_i^{'} - \frac{1}{n}\sum_{i=1}^{n} E[X_i^{'}] = \frac{1}{n}\sum_{i=1}^{n} (X_i^{'} - E[X_i^{'}]) \ \to\ 0 \quad P\text{-n.o.}$$

Resten er nu let, thi ifølge Lebesgue's Sætning konvergerer

$$E[X_n^{'}] = E[X_1 \cdot \mathbf{1}_{[-n,n]}(X_1)] \ \to_{n\to\infty}\ \mu,$$

og derfor

$$\frac{1}{n}\sum_{i=1}^{n} E[X_i^{'}] \ \to_{n\to\infty}\ \mu,$$

da konvergens i sædvanlig forstand medfører konvergens i Cecaro-middel. Konvergensen i L^1 følger umiddelbart ved at kombinere konvergensen P-n.o. med Sætning 6, da

$$\{(X_n - E[X_n])\,|\, n \geq 1\} \text{ og dermed } \{\,\frac{1}{n}\sum_{i=1}^{n} (X_i - E[X_i])\,|\, n \geq 1\}$$

er uniformt integrable. $\diamond$

Trunkeringsteknikken kan på lignende vis bruges til at vise flg. generalisation af Komogorov's Store tals lov. Bemærk at den ændrede integrabilitetsantagelse afspejler sig i valget af trunkeringskonstant.

Marcinkiewicz-Zygmond's Store tals lov.
Lad $1 \le q < 2$ være givet og lad $(Y_n)_{n\ge 1}$ betegne en følge af uafhængige identisk fordelte variable med endelig q'te moment. Da gælder

$$\frac{1}{n^{1/q}}\sum_{i=1}^{n}(Y_i - E[Y_i]) \to 0 \quad P\text{-n.o. og i } L^q.$$

Bevis. Da $q = 1$ allerede er klaret, betragter vi et $1 < q < 2$, og ved eventuelt at se på $Y_n - E[Y_n]$ i stedet for Y_n, kan vi antage, at den fælles middelværdi er lig 0. Skriv

$$Y_j = Y_j^{'} + Y_j^{"} = (Y_j^{'} - E[Y_j^{'}]) + Y_j^{"} + E[Y_j^{'}]$$

hvor

$$Y_j^{'} = Y_j \cdot \mathbf{1}_{\{|Y_j| < j^{1/q}\}} \ \text{ og } \ Y_j^{"} = Y_j \cdot \mathbf{1}_{\{|Y_j| \ge i^{1/q}\}}.$$

Ifølge Kronecker Lemma vil

$$n^{-1/q}\sum_{j=1}^{n} Y_j \to 0 \ P\text{-n.o. hvis } \sum_{j=1}^{\infty} Y_j/j^{1/q} \text{ er } P\text{-summabel.}$$

Det er derfor nok at vise, at flg. tre rækker hver for sig konvergerer P-n.o.

$$\sum_{j=1}^{\infty}(Y_j^{'} - E[Y_j^{'}]) \cdot j^{-1/q}, \quad \sum_{j=1}^{\infty} Y_j^{"} \cdot j^{-1/q} \ \text{ og } \ \sum_{j=1}^{\infty} E[Y_j^{'}] \cdot j^{-1/q}.$$

Leddene i den første sum er uafhængige, centrerede og har endelig varians. Ifølge L^2-udgaven af de Store tals lov er rækken derfor P-summabel, hvis summen af varianserne er endelig, d.v.s. hvis

$$\sum_{j=1}^{\infty} E[\,(Y_j^{'} - E[Y_j^{'}])^2] \cdot j^{-2/q} \le \sum_{j=1}^{\infty} E[Y_j^{'\,2}] \cdot j^{-2/q} < \infty.$$

Men dette gælder, da der findes en konstant $r_q > 0$ kun afhængig af q, så at

$$\sum_{i=1}^{\infty} E[Y_j^{'\,2}] \cdot j^{-2/q} = E[\,Y_1^2 \cdot \sum_{j:\, j > |Y_1|^q} j^{-2/q}] \le r_q\, E[\,Y_1^2 \cdot |Y_1|^{-q(2/q-1)}] = r_q\, E[\,|Y_1|^q].$$

Konvergensen af række nr. to følger som ovenfor af Borel-Cantelli Lemmaet, thi da $E[\,|Y_1|^q] < \infty$ er

$$\sum_{j=1}^{\infty} P(Y_j^{''} \neq 0) = \sum_{j=1}^{\infty} P(|Y_j| \ge j^{1/q}) = \sum_{j=1}^{\infty} P(|Y_1|^q \ge j) < \infty.$$

Hvad angår den sidste række bemærkes først, at da Y_j'erne har middelværdi 0 er

$$E[Y_j^{'}] = -E[Y_j^{"}] = -E[\,Y_j \cdot \mathbf{1}_{\{|Y_j| \ge j^{1/q}\}}] = -E[\,Y_j \cdot \mathbf{1}_{\{|Y_j|^q \ge j\}}],$$

d.v.s. vi skal vise, at

$$\sum_{j=1}^{\infty} E[\,|Y_1| \cdot \mathbf{1}_{\{|Y_1|^q \geq j\}}] \cdot j^{-1/q} < \infty.$$

Men dette følger af, at der findes endnu en konstant $\tilde{r}_q$ kun afhængig af q, så at

$$\sum_{j=1}^{\infty} E[\,|Y_1| \cdot \mathbf{1}_{\{|Y_1|^q \geq j\}}] \cdot j^{-1/q} = E[\,|Y_1| \cdot \sum_{1 \leq j \leq |Y_1|^q} j^{-1/q}] \leq \tilde{r}_q\, E[\,|Y_1| \cdot |Y_1|^{-q(1/q-1)}].$$

D.v.s. den betragtede sum er domineret af $\tilde{r}_q \cdot E[\,|Y_1|^q]$ og dermed endelig. Beviset for konvergensen i L^q udsættes til senere. $\diamond$

Som det fremgår af beviset spiller implikationen

uafhængighed $\Rightarrow$ ukorrellerethed d.v.s. orthogonalitet

en væsentlig rolle i beviset for ovenstående L^2-udgave af de store tals lov. I det næste resultat tages udgangspunkt i ukorrellerethed i stedet for uafhængighed.

De store tals lov (L^2-udgave, supplement).
Lad $(X_n)_{n\geq 1}$ betegne en følge af ukorrellerede kvadratisk integrable stokastiske variable så at

$$\sum_{n=1}^{\infty} Var(X_n)/n^2 < \infty$$

Da gælder, idet $\hat{X}_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i])$ for $n \geq 1$,

1) $\hat{X}_n \to 0$ i sandsynlighed og $L^2(P)$.

2) $\hat{X}_{[\lambda^n]} \to 0$ P-n.o. for $\lambda > 1$.

3) $\hat{X}_n \to 0$ P-n.o. hvis

$$P\left(\sup_n (X_n - E[X_n]) < \infty \text{ eller } \inf_n (X_n - E[X_n]) > -\infty\right) = 1.$$

Bevis. For nemheds skyld skrives μ_n i stedet for $E[X_n]$. Da $(X_n - \mu_n)_{n\geq 1}$ pr. antagelse er parvis orthogonale i $L^2(P)$ fås for ethvert $n \geq 1$ af Pythagoras, at

$$E[\hat{X}_n^2] = \frac{1}{n^2}\sum_{j=1}^{n} E[(X_j - \mu_j)^2] = \frac{1}{n^2}\sum_{j=1}^{n} Var(X_j) \to_{n\to\infty} 0,$$

hvor konvergensen følger af antagelsen og Kronecker Lemmaet. D.v.s. $\hat{X}_n \to 0$ i $L^2(P)$ og dermed også i sandsynlighed.

For ethvert $\lambda > 1$ har vi tilsvarende

$$\sum_{n=1}^{\infty} E[\hat{X}^2_{[\lambda^n]}] = \sum_{n=1}^{\infty} \frac{1}{[\lambda^n]^2} \sum_{j=1}^{[\lambda^n]} Var(X_j)$$

$$= \sum_{j=1}^{\infty} \left(Var(X_j) \sum_{n:\,[\lambda^n]\geq j}^{\infty} \frac{1}{[\lambda^n]^2} \right) \leq C_\lambda \sum_{j=1}^{\infty} Var(X_j)/j^2 < \infty,$$

hvor C_λ er en konstant kun afhængig af λ. For ethvert $\lambda > 1$ er

$$E[\sum_{n=1}^{\infty} \hat{X}^2_{[\lambda^n]}]$$

dermed endelig, hvoraf 2) let følger f.eks. ved brug af Monoton konvergens.
Ifølge 2) er

$$P(\hat{X}_{\nu_k(n)} \to 0 \text{ for alle } k \geq 1) = 1,$$

hvor $\nu_k(n) = [(1 + 1/k)^n]$ for alle $n, k \geq 1$. Kombineres dette med antagelserne gælder derfor for P-n.a. ω, at

$$\lim_n \hat{X}_{\nu_k(n)}(\omega) = 0 \ k \geq 1 \text{ samt } -\infty < \inf_j (X_j(\omega) - \mu_j) \text{ eller } \sup_j (X_j(\omega) - \mu_j) < \infty,$$

hvilket i h.h.t. en tidligere overvejelse netop betyder, at $\lim_n \hat{X}_n = 0$ P-n.o. $\diamond$

Ved at udnytte punkt 3) i det netop viste resultat kan det vises, at Kolmogorov's store tals lov stadig gælder selvom uafhængighed erstattes med parvis uafhængighed. Da denne generalisation yderst sjældent er relevant, vil vi lade den ligge. Den interesserede læser kan finde argumenterne i Hoffmann's bog sektionerne 4.11 og 4.12.

De store tals love II.
Som allerede nævnt adskiller eksponenten 2 sig fra andre eksponenter. Men som vi nu skal se, kan man i det uafhængige tilfælde ved hjælp af den såkaldte *symmetriseringsteknik* vise lignende resultater for alle eksponenter $\alpha > 0$.
En væsentlig brik i teorien er flg. resultat normalt kaldet *Khinchine's Ulighed*:

Khinchine's Ulighed.
Lad $(\epsilon_i)_{i\geq 1}$ betegne en følge af uafhængige Bernoulli variable. Da findes der for alle $\alpha > 0$ positive konstanter c_α og *C_α kun afhængig af α, så at*

$$c_\alpha \cdot (\sum_{j=1}^{n} b_j^2)^{\alpha/2} \leq E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha] \leq C_\alpha \cdot (\sum_{j=1}^{n} b_j^2)^{\alpha/2}$$

for alle $n \geq 1$ og alle reelle talfølger $(b_j)_{j\geq 1}$.
Bevis del I. Ifølge Jensen's ulighed er $\alpha \mapsto E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha]^{1/\alpha}$ voksende for ethvert $n \geq 1$ og alle reelle talfølger $(b_n)_{n\geq 1}$, og da

$$E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^2] = E[\left(\sum_{j=1}^{n} b_j \cdot \epsilon_j\right)^2] = \sum_{j=1}^{n} b_j^2$$

ses, at

$$E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha] \leq E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^2]^{\alpha/2} = (\sum_{j=1}^{n} b_j^2)^{\alpha/2} \quad \text{for } \alpha \leq 2$$

og

$$(\sum_{j=1}^{n} b_j^2)^{\alpha/2} = E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^2]^{\alpha/2} \leq E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha] \quad \text{for } \alpha \geq 2.$$

D.v.s. 1 kan bruges som C_α for $0 < \alpha \leq 2$ og som c_α for $\alpha \geq 2$. Følgen $(b_j)_{j\geq 1}$ givet ved $b_1 = 1$ og $b_j = 0$ ellers viser, at der i begge tilfælde er tale om den optimale konstant. $\diamond$

De resterende tilfælde er tæt forbundne for har vi bestemt C_α for $\alpha > 2$ gælder ifølge Cauchy-Schwarz's ulighed for $0 < \alpha < 2$, at

$$\sum_{j=1}^{n} b_j^2 = E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^2] = E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^{2-\alpha/2} \cdot |\sum_{j=1}^{n} b_j \cdot \epsilon_j|^{\alpha/2}] \leq$$

$$E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^{4-\alpha}]^{1/2} \cdot E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha]^{1/2} \leq C_{4-\alpha}^{1/2} \cdot (\sum_{j=1}^{n} b_j^2)^{1-\alpha/4} \cdot E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha]^{1/2},$$

som efter forkortning viser, at

$$(\sum_{j=1}^{n} b_j^2)^{\alpha/4} \leq C_{4-\alpha}^{-1/2} \cdot E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha]^{1/2}.$$

D.v.s. $C_{4-\alpha}^{-1}$ kan bruges som c_α i intervallet $0 < \alpha < 2$.

Bestemmelsen af C_α for $\alpha > 2$ er mere kompliceret, specielt er bestemmelsen af den optimale værdi yderst vanskeligt. Hoffmann viser i sektion 4.30, at $C_\alpha = 2^{\alpha/2} \cdot \alpha \cdot \Gamma(\alpha/2)$ kan bruges. Konstanten er ikke optimal ligesom den værdi, vi nu vil bestemme ved brug af teorien om betingede middelværdier.

Bevis del II. Lad $\alpha > 2$ og $n \geq 1$ være givet og lad $X_1, \ldots, X_n$ betegne uafhængige $N(0,1)$-fordelte stokastiske variable. Definer

$$\mathcal{B}_i := \sigma(\{X_i > 0\}) \;\; i = 1, \ldots, n \;\; \text{og} \;\; \mathcal{B} := \sigma(\bigcup_{i=1}^{n} \mathcal{B}_i).$$

Ifølge regneregler for betingede middelværdier gælder for ethvert i, da X_i'erne er symmetriske og $P(X_i > 0)$ derfor lig $1/2$ for alle i, at

$$E[X_i \,|\, \mathcal{B}] = E[X_i \,|\, \mathcal{B}_i] = \rho \cdot (\mathbf{1}_{\{X_i > 0\}} - \mathbf{1}_{\{X_i \leq 0\}}) \;\; P\text{-n.o.},$$

hvor

$$\rho := 2 \cdot E[X_i \,|\, X_i > 0] = -2 \cdot E[X_i \,|\, X_i \leq 0] = \sqrt{2/\pi}.$$

D.v.s. $E[X_1 \,|\, \mathcal{B}]/c, \ldots, E[X_n \,|\, \mathcal{B}]/\rho$ er uafhængige Bernoulli variable, og for ethvert valg af konstanter $b_1, \ldots, b_n$ gælder derfor

$$\rho^\alpha \cdot E[\,|\sum_{j=1}^{n} b_j \cdot \epsilon_j|^\alpha] = E[\,|\sum_{j=1}^{n} b_j \cdot E[X_j \,|\, \mathcal{B}]\,|^\alpha] = E[\,|\,E[\sum_{j=1}^{n} b_j \cdot X_j \,|\, \mathcal{B}]\,|^\alpha]$$

$$\leq E[\,|\sum_{j=1}^{n} b_j \cdot X_j|^\alpha] = E[\,|N(0, \sum_{j=1}^{n} b_j^2\,)|^\alpha] = (\,\sum_{j=1}^{n} b_j^2\,)^{\alpha/2} \cdot E[\,|N(0,1)|^\alpha].$$

D.v.s. $C_\alpha := (\pi/2)^{\alpha/2} \cdot E[\,|N(0,1)|^\alpha] = \pi^{(\alpha-1)/2} \cdot \Gamma((\alpha+1)/2)$ kan bruges. $\diamond$

Korollar 1 *Lad $Z_1, \ldots, Z_n$ betegne uafhængige symmetriske stokastiske variable. For ethvert $\alpha > 0$ gælder*

$$E[\,|\sum_{k=1}^{n} Z_k|^\alpha\,] \leq C_\alpha \cdot E[\,(\,\sum_{k=1}^{n} |Z_k|^2\,)^{\alpha/2}] \leq C_\alpha \cdot n^{\beta(\alpha)} \sum_{k=1}^{n} E[\,|Z_k|^\alpha\,],$$

hvor C_α er konstanten fra Khinchine's ulighed, og $\beta(\alpha) = (\alpha/2 - 1)^+$, d.v.s. $\beta(\alpha) = 0$ for $0 < \alpha \leq 2$ og $\beta(\alpha) = \alpha/2 - 1$ for $\alpha > 2$.

Bevis. Lad $\alpha > 0$ være givet og lad $\epsilon_1, \ldots, \epsilon_n$ betegne uafhængige Bernoulli variable, så at $(Z_1, \ldots, Z_n)$ og $(\epsilon_1, \ldots, \epsilon_n)$ er uafhængige. Da Z_i'erne er symmetriske og uafhængige, gælder for ethvert valg af $+$ og $-$, at

$$(Z_1, \ldots, Z_n) \sim (\pm Z_1, \ldots, \pm Z_n)$$

Ved nu at udnytte Loven om total sandsynlighed fås derfor, at

$$(Z_1, \ldots, Z_n) \sim (\epsilon_1 \cdot Z_1, \ldots, \epsilon_n \cdot Z_n),$$

og idet

$$H_\alpha(a_1, \ldots, a_n) := E[\,|\sum_{k=1}^{n} \epsilon_k \cdot a_k|^\alpha\,] \quad a_1,\ \ldots,\ a_n \in \mathbf{R}$$

følger ved brug af Fubini's Sætning, at

$$E[\,|\sum_{k=1}^{n} Z_k|^\alpha\,] = E[\,|\sum_{k=1}^{n} \epsilon_k \cdot Z_k|^\alpha\,] = E[H_\alpha(Z_1, \ldots, Z_n)],$$

og dermed ifølge Khinchine' s Ulighed, at

$$E[\,|\sum_{k=1}^{n} Z_k|^\alpha\,] \leq C_\alpha \cdot E[\,(\sum_{k=1}^{n} |Z_k|^2\,)^{\alpha/2}].$$

Korollarets sidste ulighed følger nu ved for $0 < \alpha \leq 2$ at udnytte, at $x \mapsto x^{\alpha/2}$ er subadditiv, og for $\alpha > 2$ at benytte Jensen's Ulighed. $\diamond$

Ved brug af formel (4.6.5), d.v.s. uligheden

$$E[\,|Z + Y|^p] \geq E[\,|Z + E[Y]\,|^p]$$

for $p \geq 1$ og uafhængige Z og Y, hvor Y har endelig middelværdi,

giver dette anledning til flg. generelle ulighed, hvor konstanterne C_α og $\beta(\alpha)$ har samme mening som i Korollar 1.

Korollar 2 *Lad $Z_1, \ldots, Z_n$ betegne uafhængige stokastiske variable med endelig middelværdi μ_k for $k = 1, \ldots, n$. Da gælder for ethvert $\alpha > 0$*

$$E[\,|\sum_{k=1}^{n}(Z_k - \mu_k)|^\alpha\,] \leq 2^\alpha \cdot C_\alpha \cdot n^{\beta(\alpha)} \sum_{k=1}^{n} E[\,|Z_k - \mu_k|^\alpha\,].$$

Bevis. Da uligheden er triviel for $\alpha \leq 1$, idet $x \mapsto x^\alpha$ er voksende og subadditiv på $\mathbf{R}_+$, betragtes et $\alpha > 1$. Lad $Y_1, \ldots, Y_n$ være en uafhængig kopi af $Z_1, \ldots, Z_n$, d.v.s.

$$(Y_1, \ldots, Y_n) \text{ og } (Z_1, \ldots, Z_n) \text{ er uafhængige og identisk fordelte.}$$

Da $Z_1 - Y_1, \ldots, Z_n - Y_n$ derfor er uafhængige og symmetriske, fås af Korollar 1

$$E[\,|\sum_{k=1}^{n}(Z_k - Y_k)|^\alpha\,] \leq C_\alpha \cdot n^{\beta(\alpha)} \sum_{k=1}^{n} E[\,|Z_k - Y_k|^\alpha\,].$$

Men da $Z_k \sim Y_k$ for alle $k \geq 1$ og dermed

$$E[\,|Z_k - Y_k|^\alpha\,] \leq 2^{\alpha-1} \cdot (E[\,|Z_k - \mu_k|^\alpha\,] + E[\,|Y_k - \mu_k|^\alpha\,]) = 2^\alpha \cdot E[\,|Z_k - \mu_k|^\alpha\,],$$

følger påstanden af formel (4.6.5), idet

$$E[\,|\sum_{k=1}^{n}(Z_k-Y_k)|^\alpha\,] = E[\,|\sum_{k=1}^{n}(Z_k-\mu_k)-\sum_{k=1}^{n}(Y_k-\mu_k)|^\alpha\,] \geq E[\,|\sum_{k=1}^{n}(Z_k-\mu_k)|^\alpha\,],$$

da $\sum_{k=1}^{n}(Y_k-\mu_k)$ har middelværdi 0 og $\alpha > 1$. ◇

Korollar 2 giver umiddelbart anledning til flg. L^α-konvergensresultater for summer af uafhængige stokastiske variable. C_α og $\beta(\alpha)$ har samme mening som ovenfor og

$$\gamma(\alpha) := \alpha \text{ hvis } 0 < \alpha \leq 2 \text{ og } \gamma(\alpha) := \alpha/2+1 \text{ hvis } \alpha > 2.$$

Korollar 3 *Lad $(X_n)_{n\geq 1}$ betegne uafhængige stokastiske variable med endelig middelværdi μ_n for $n \geq 1$. Da gælder for ethvert $0 < \alpha \leq 2$, at*

$$\sum_{n=1}^{\infty} E[\,|X_n-\mu_n|^\alpha\,] < \infty \Rightarrow \sum_{n=1}^{\infty}(X_n-\mu_n) \text{ eksisterer i } L^\alpha,$$

og for alle $\alpha > 0$, at

$$\sum_{n=1}^{\infty} E[\,|X_n-\mu_n|^\alpha\,]/n^{\gamma(\alpha)} < \infty \Rightarrow \frac{1}{n}\sum_{k=1}^{n}(X_k-\mu_k) \to 0 \text{ i } L^\alpha.$$

Bevis. For at vise første del, er det nok at vise, at højresiden udgør en Cauchy følge i L^α. Men dette følger umiddelbart af antagelsen og Korollar 2, idet denne viser, at for $0 < \alpha \leq 2$ er

$$E[\,|\sum_{k=n}^{m}(X_k-\mu_k)|^\alpha\,] \leq 2^\alpha C_\alpha \sum_{k=n}^{m} E[\,|X_k-\mu_k|^\alpha\,]$$

for alle $1 \leq n \leq m$. Hvad angår andel del fås af Korollar 2, at

$$E[\,|\frac{1}{n}\sum_{k=1}^{n}(X_k-\mu_k)|^\alpha\,] \leq \frac{2^\alpha C_\alpha}{n^\alpha}\cdot n^{\beta(\alpha)} \sum_{k=1}^{n} E[\,|X_k-\mu_k|^\alpha\,]$$

og dermed, da $\gamma(\alpha) = \alpha - \beta(\alpha)$,

$$E[\,|\frac{1}{n}\sum_{k=1}^{n}(X_k-\mu_k)|^\alpha\,] \leq \frac{2^\alpha C_\alpha}{n^\gamma(\alpha)} \sum_{k=1}^{n} E[\,|X_k-\mu_k|^\alpha\,]$$

for alle $n \geq 1$, hvoraf påstanden følger ved brug af Kronecker's lemma. ◇

Vi kan nu formulere og bevise en L^α-version af de store tals lov for et generelt $\alpha > 0$. For $\alpha < 2$ er der tale om en direkte oversættelse af L^2-udgaven, hvorimod momentbetingelsen er ændret lidt i intervallet $\alpha > 2$.

De store tals lov (L^α-udgave).
Lad $(X_n)_{n\geq 1}$ *betegne en følge af uafhængige stokastiske variable med endelig middelværdi* μ_n *for* $n \geq 1$. *Da gælder for* $\alpha \leq 2$

$$\sum_{n=1}^{\infty} E[\,|X_n - \mu_n|^\alpha\,]/n^\alpha < \infty \Rightarrow \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i) \to 0 \quad P\text{-n.o. og i } L^\alpha(P).$$

og for $\alpha > 2$

$$\sum_{n=1}^{\infty} E[\,|X_n - \mu_n|^\alpha\,]/n^{(1+\alpha/2)} < \infty \Rightarrow \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i) \to 0 \quad P\text{-n.o. og i } L^\alpha(P).$$

Bevis. Lad $\alpha \leq 2$ være givet. Ifølge Korollar 3 punkt 1 har vi

$$\sum_{n=1}^{\infty} E[\,|X_n - \mu_n|^\alpha\,]/n^\alpha < \infty \Rightarrow \sum_{n=1}^{\infty}(X_n - \mu_n)/n \;\text{ konvergent i } L^\alpha.$$

Rækken konvergerer derfor også i sandsynlighed og dermed P-n.o. ifølge sætningen angående konvergens af summer af uafhængige variable. Kronecker Lemmaet giver derfor umiddelbart, at

$$\hat{X}_n := \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu_k) \;\to\; 0 \quad P\text{-n.o.}$$

og da Korollar 3 punkt 2 sikrer, at $\hat{X}_n \to 0$ i L^α, er tilfældet $\alpha \leq 2$ klaret.
Betragt dernæst et $\alpha > 2$. L^α-konvergensen af $(\hat{X}_n)_{n\geq 1}$ følger igen af Korollar 3 punkt 2. Hvad angår konvergensen P-n.o. udnyttes som i beviset for sætningen om konvergens af summer af uafhængige variable, at det er nok at vise, at $\hat{X}_{2^n} \to 0$ og $M_n \to 0$ P-n.o., hvor for $n \geq 1$

$$M_n := \max_{2^n < k \leq 2^{n+1}} |\hat{X}_k - \hat{X}_{2^n}| \leq |\hat{X}_{2^n}| + \frac{1}{2^n} \max_{2^n < k \leq 2^{n+1}} |\sum_{j=2^n+1}^{k}(X_j - \mu_j)|.$$

Hertil er det som bekendt nok at vise, at

$$\sum_{n=1}^{\infty} E[\,|\hat{X}_{2^n}|^\alpha] < \infty \quad \text{og} \quad \sum_{n=1}^{\infty} E[M_n^\alpha] < \infty.$$

Udnyttes Korollar 2 ses, at der findes en konstant C kun afhængig af α, så at

$$\sum_{n=1}^{\infty} E[\,|\hat{X}_{2^n}|^\alpha] \leq C\sum_{n=1}^{\infty}\frac{2^{n(\alpha/2-1)}}{2^{n\alpha}}\sum_{j=1}^{2^n} E[\,|X_j - \mu_j|^\alpha\,]$$

$$= C\sum_{j=1}^{\infty} E[\,|X_j - \mu_j|^\alpha\,] \sum_{n:\,2^n \geq j} \frac{1}{2^{n(\alpha/2+1)}}$$

$$\leq 2\,C \sum_{j=1}^{\infty} E[\,|X_j-\mu_j|^{\alpha}\,]/j^{1+\alpha/2} < \infty.$$

$\sum_{n=1}^{\infty} E[M_n^{\alpha}]$ er derfor også endelig, hvis

$$\sum_{n=1}^{\infty} \frac{1}{2^{n\alpha}} \cdot E[\max_{2^n<k\leq 2^{n+1}} |\sum_{j=2^n+1}^{k} (X_j-\mu_j)|^{\alpha}] < \infty.$$

Men ifølge Korollaret til Ottaviani's Ulighed er dette tilfældet, hvis

$$\sum_{n=1}^{\infty} \frac{1}{2^{n\alpha}} \cdot E[\,|\sum_{j=2^n+1}^{2^{n+1}} (X_j-\mu_j)|^{\alpha}] < \infty,$$

og dermed ifølge Korollar 2 hvis

$$\sum_{n=1}^{\infty} \frac{2^{n(\alpha/2-1)}}{2^{n\alpha}} \sum_{j=2^n+1}^{2^{n+1}} E[\,|X_j-\mu_j|^{\alpha}\,] < \infty.$$

Men dette sikres præcist af antagelsen da

$$\sum_{n=1}^{\infty} \frac{1}{2^{n(\alpha/2+1)}} \sum_{j=2^n+1}^{2^{n+1}} E[\,|X_j-\mu_j|^{\alpha}\,] \leq 2^{1+\alpha/2} \sum_{j=1}^{\infty} E[\,|X_j-\mu_j|^{\alpha}\,]/j^{1+\alpha/2} < \infty. \quad \diamond$$

Symmetriseringsteknikken sikrer også den manglende L^q-konvergens i Marcinkiewicz-Zygmond's Store tals lov. Thi lad for $1 < q < 2$ situationen være som i sætningen.
Først reduceres til det symmetriske tilfælde. For hvis $(\tilde{Y}_j)_{j\geq 1}$ er en uafhængig kopi af $(Y_j)_{j\geq 1}$, d.v.s. en følge af indbyrdes uafhængige variable, som er uafhængige af $(Y_j)_{j\geq 1}$, men med samme fordeling, fås af formel (4.6.5), at

$$E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y_j|^q] \leq E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y_j - \frac{1}{n^{1/q}}\sum_{j=1}^{n} \tilde{Y}_j|^q] = E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n}(Y_j-\tilde{Y}_j)|^q],$$

og da $(Y_j-\tilde{Y}_j)_{j\geq 1}$'erne er uafhængige, symmetriske og identisk fordelte, er det, hvad konvergens i q-middel angår, nok at betragte det symmetriske tilfælde. Vi vil derfor i det videre forløb yderligere antage, at Y_i'erne er symmetriske.
Betragt for et givet $k \geq 1$ opsplitningen

$$Y_i = Y'_{k,i} + Y''_{k,i} \quad \text{hvor } Y'_{k,i} := Y_i \cdot \mathbf{1}_{\{|Y_i|\leq k\}} \quad \text{og} \quad Y''_{k,i} := Y_i \cdot \mathbf{1}_{\{|Y_i|>k\}}.$$

$(Y'_{k,i})_{i\geq 1}$ og $(Y''_{k,i})_{i\geq 1}$ er da begge uafhængige, symmetriske og identisk fordelte. Ifølge Korollar 2 ovenfor findes der derfor en konstant C kun afhængig af q, så at

$$E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y''_{k,j}|^q] \leq C \cdot E[\,|Y''_{k,1}|^q] = C \cdot E[\,|Y_1|^q,\ |Y_1| > k],$$

hvilket viser, at

$$\sup_n E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y''_{k,j}|^q]$$

kan gøres så lille som ønsket ved at vælge k stor nok. L^q-konvergensen vil derfor være vist, hvis vi for givet k kan vise, at

$$\lim_n E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y'_{k,j}|^q\,] = \lim_n E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y_j \cdot \mathbf{1}_{\{|Y_j|\leq k\}}|^q\,] = 0.$$

Da $q < 2$ er det nok at vise konvergens i L^2, og da summanderne er uafhængige centrede kvadratisk integrable variable, følger dette af Pythagoras, da

$$E[\,|\frac{1}{n^{1/q}}\sum_{j=1}^{n} Y_j \cdot \mathbf{1}_{\{|Y_j|<k\}}|^2] = \frac{1}{n^{2/q}}\sum_{j=1}^{n} E[Y_j^2 \cdot \mathbf{1}_{\{|Y_j|<k\}}] \leq \frac{n \cdot k^2}{n^{2/q}} \rightarrow_{n\rightarrow\infty} 0$$

for ethvert $k \geq 1$. $\diamond$

Fordelingskonvergens.
Lad i det følgende (S, d) betegne et separabelt metrisk rum. Vigtige eksempler er $\mathbf{R}^n$ $n \geq 1$ eller mere generelt delmængder heraf udstyret med den Euklidiske metrik. Lad endvidere $(X_n)_{n\geq 1}$ og X betegne stokastiske funktioner med værdier i S, d.v.s. $(\mathcal{F}, \mathcal{B}(S))$-målelige funktioner fra Ω ind i S, hvor $(\Omega, \mathcal{F}, P)$ er et sandsynlighedsfelt, hvorpå alle omtalte variable tænkes defineret. I analogi med det reelle tilfælde indføres flg. konvergensbegreb.

Definition. $X_n \to X$ i sandsynlighed hvis $\lim_n P(d(X_n, X) > \epsilon) = 0$ for alle $\epsilon > 0$.

Bemærkning. Da separabiliteten af S sikrer, at $\mathcal{B}(S \times S) = \mathcal{B}(S) \otimes \mathcal{B}(S)$, sikrer kontinuiteten af $(x, y) \mapsto d(x, y)$, at $\{d(X_n, X) > \epsilon\}$ er en hændelse, og den kan som sådan tilordnes en sandsynlighed.

På den reelle akse er betingelsen for konvergens i sandsynlighed

$$\lim_n P(\,|X_n - X| > \epsilon) = 0 \quad \text{for alle } \epsilon > 0,$$

hvilket, som tidligere vist (se noternes Lemma 12), er ækvivalent med at

$$\lim_n E[\,|X_n - X| \wedge 1\,] = 0.$$

Denne ækvivalens generaliserer ordret til det almene tilfælde, d.v.s.

$$X_n \to X \text{ i sandsynlighed} \Leftrightarrow \lim_n E[\,d(X_n, X) \wedge 1\,] = 0.$$

Ved brug heraf fås som i det reelle tilfælde.

Fk 1 *$X_n \to X$ i sandsynlighed $\Rightarrow$ $X_{n_k} \to X$ P-n.o. for en delfølge $(n_k)_{k\geq 1}$.*

Bevis. Da $\lim_n E[\,d(X_n, X) \wedge 1\,] = 0$ kan vi vælge en delfølge $(n_k)_{k\geq 1}$, så at

$$\sum_{k=1}^{\infty} E[\,d(X_{n_k}, X) \wedge 1\,] < \infty \quad \text{og derfor } d(X_{n_k}, X) \wedge 1 \to 0 \;\; P\text{-n.o.},$$

hvilket netop betyder at $X_{n_k} \to X$ P-n.o. ◇

En anden vigtig konsekvens er følgende.

Fk 2 *Lad (T, δ) betegne endnu et separabelt metrisk rum og lad $f : S \to T$ være en kontinuert funktion. Da gælder*

$$X_n \to X \text{ i sandsynlighed} \Rightarrow f(X_n) \to f(X) \text{ i sandsynlighed}$$

Bevis. Vi skal vise $\lim_n E[\,\delta(f(X_n), f(X)) \wedge 1\,] = 0$. Antag at det ikke gælder, d.v.s. antag

$$\exists\, \delta > 0\, \exists\, (n_k)_{k\geq 1} : \; E[\,\delta(f(X_{n_k}), f(X)) \wedge 1\,] > \delta \;\; \text{for alle } k.$$

Men dette fører til en modstrid, da

$$X_{n_k} \to X \text{ i s.s.} \Rightarrow \exists\, (k_l)_{l\geq 1}\ X_{n_{k_l}} \to X \ P\text{-n.o.} \Rightarrow f(X_{n_{k_l}}) \to f(X) \ P\text{-n.o.}$$

$$\Rightarrow \delta(f(X_{n_{k_l}}), f(X))) \wedge 1 \to 0 \ P\text{-n.o.} \Rightarrow E[\,\delta(f(X_{n_{k_l}}), f(X))) \wedge 1\,] \to 0. \ \diamond$$

Kontinuiteten udnyttes i $\Rightarrow$ nummer to, og da det her er nok, at f er kontinuert i $X(\omega)$ for n.a. ω, kan antagelsen svækkes til, at f er kontinuert P_X-n.o. Vi har derfor flg. skærpelse.

Fk 2a *Lad* (T, δ) *betegne endnu et separabelt metrisk rum og lad* $f : S \to T$ *være en Borel funktion, som er kontinuert* P_X*-n.o. Da gælder*

$$X_n \to X \text{ i sandsynlighed} \Rightarrow f(X_n) \to f(X) \text{ i sandsynlighed}$$

Specialtilfældet $T = S$ og δ en metrik, der er ækvivalent med d, viser, idet den identiske afbildning er kontinuert både som afbildning

$$(S, d) \to (S, \delta) \ \text{ og } \ (S, \delta) \to (S, d),$$

at konvergens i sandsynlighed ikke afhænger af den eksplicit valgte metrik, blot vi holder os indenfor klassen af ækvivalente metrikker. Dette udnyttes f.eks. i følgende bevis.

Fk 3 *Idet* $S \times S$ *udstyres med en produktmetrik gælder*

$$X_n \to X \text{ og } Y_n \to Y \text{ i sandsynlighed} \Leftrightarrow (X_n, Y_n) \to (X, Y) \text{ i sandsynlighed.}$$

Bevis. $\Leftarrow$ følger af Fk 2, da projektionsafbildningerne er kontinuerte, og $\Rightarrow$ fås, da

$$d_1((x_1, y_1), (x_2, y_2)) := d(x_1, y_1) + d(x_2, y_2)$$

er en produktmetrik, umiddelbart af uligheden

$$\tilde{d}((X_n, Y_1), (X, Y)) \wedge 1 \leq d(X_n, X) \wedge 1 + d(Y_n, Y) \wedge 1. \ \diamond$$

Sætning 6 og Fk 2 viser endvidere, at for alle $f \in \mathrm{b}C(S)$ gælder

$$X_n \to X \text{ i sandsynlighed} \Rightarrow f(X_n) \to f(X) \text{ i sandsynlighed}$$

$$\Rightarrow f(X_n) \to f(X) \text{ i } L^1(P) \Rightarrow E[f(X_n)] \to_{n\to\infty} E[f(X)] = \int f\, dP_X.$$

Med udgangspunkt heri indføres den såkaldte *konvergens i fordeling*, i Hoffman's bog betegnet $\xrightarrow{\sim}$, i h.h.t. flg. definition.

Definition. En følge $(X_n)_{n\geq 1}$ af stokastiske funktioner med værdier i S siges at *konvergere i fordeling* mod μ, et Borel sandsynlighedsmål på S, hvis

$$E[f(X_n)] \to_{n\to\infty} \int f\, d\mu \ \text{ for alle } f \in \mathrm{b}C(S).$$

Dette betegnes i givet fald $X_n \xrightarrow{\sim} \mu$. Hvis $\mu = P_X$ for en stokastisk funktion X med værdier i S skrives også $X_n \xrightarrow{\sim} X$, og man taler om konvergens i fordeling mod X. I følge den lille transformationssætning gælder altså, at

$$X_n \xrightarrow{\sim} X \;\Leftrightarrow\; E[f(X_n)] \to_{n\to\infty} E[f(X)] \;\text{ for alle } f \in \mathrm{b}C(S).$$

Det ovenstående kan derfor formuleres som implikationen (se (5.4.4)).

Fk 4 $X_n \to X$ i sandsynlighed $\Rightarrow$ $X_n \xrightarrow{\sim} X$.

Grænsemålet for en konvergent følge er tydeligvis entydigt bestemt, d.v.s.

$$X_n \xrightarrow{\sim} \mu \;\text{ og }\; X_n \xrightarrow{\sim} \nu \;\Rightarrow\; \mu = \nu$$

da

$$X_n \xrightarrow{\sim} \mu \text{ og } X_n \xrightarrow{\sim} \nu \;\Rightarrow\; \int f\,d\mu = \int f\,d\mu \;\; f \in \mathrm{b}C(S) \;\Rightarrow\; \mu = \nu.$$

Derimod kan vi sagtens have, at $X_n \xrightarrow{\sim} X$ og $X_n \xrightarrow{\sim} Y$, selv om X og Y er vidt forskellige. Men der gælder dog flg. implikation

$$X_n \xrightarrow{\sim} X \;\text{ og }\; X_n \xrightarrow{\sim} Y \;\Rightarrow\; P_X = P_Y.$$

Før vi ser nærmere på konvergens i fordeling knyttes et par kommentarer til definitionen. Da kontinuitet i metriske rum svarer til følgekontinuitet bevares $C(S)$ og dermed konvergens i fordeling under overgang til en ækvivalent metrik. Endvidere ses ved opsplitning i positiv og negativ del, at det er nok at eftervise definitionsbetingelsen for $f \in \mathrm{b}C(S)_+$. Men for ethvert $f \in \mathrm{b}C(S)_+$ med $0 \le f \le M$ er

$$M = f + (M - f) \;\text{ og } f,\, M - f \in \mathrm{b}C(S)_+,$$

og da konvergensen er oplagt for konstante funktioner, sikrer integralet linearitet, samt flg. simple implikation angående positive reelle tal (a_n), (b_n), a og b:

$$\liminf_n a_n \ge a,\; \liminf_n b_n \ge b \text{ og } \lim_n (a_n + b_n) = a + b \;\Rightarrow\; \lim_n a_n = a,$$

at

$$X_n \xrightarrow{\sim} \mu \Leftrightarrow \liminf_n E[f(X_n)] \ge \int f\,d\mu \;\text{ for alle } f \in \mathrm{b}C(S)_+.$$

og dermed også

$$X_n \xrightarrow{\sim} X \;\Leftrightarrow\; \liminf_n E[f(X_n)] \ge E[f(X)] \;\text{ for alle } f \in \mathrm{b}C(S)_+.$$

Da $f \wedge n \uparrow f$ og $f \wedge n \in \mathrm{b}C(S)_+$ for $f \in C(S)_+$ kan $\Rightarrow$ ved brug af Monoton konvergens skærpes til

$$X_n \xrightarrow{\sim} \mu \;\;(\,X\,) \;\Rightarrow\; \liminf_n E[f(X_n)] \ge \int f\,d\mu \;\;(\,E[f(X)]\,) \;\text{ for alle } f \in C(S)_+.$$

Tilsvarende overvejelser viser endvidere, at vi som i Fk 2a kan udvide til funktioner, som kun er kontinuerte n.o.

Fk 5

$$X_n \overset{\sim}{\to} \mu \;\Rightarrow\; E[f(X_n)] \to_{n\to\infty} \int f\, d\mu$$

for ethvert $f \in \mathrm{bM}(S)$*, som er kontinuert* μ*-n.o.*

Bevis. Ved som ovenfor at opsplitte i positiv og negativ del samt dernæst studere f og $M - f$, hvor $0 \leq f \leq M$, indses, at det er nok at vise

$$\liminf_n \; E[f(X_n)] \geq \int f\, d\mu$$

for et givet $f \in \mathrm{bM}(S)_+$, som er kontinuert μ-n.o. Men for ethvert $g \in \mathrm{bM}(S)_+$ definerer (se Hoffmann side 561 nederst)

$$g_k(x) := \inf\{k \wedge g(y) + k \cdot d(x,y) \,|\, y \in S\} \qquad x \in S,\ k \geq 1$$

funktioner opfyldende: $g_k(x) \to_{k\to\infty} g(x)$, hvis g er kontinuert i x, samt

$$g_k \in \mathrm{b}C(S)_+ \;\text{ og }\; 0 \leq g_k \leq g_{k+1} \leq g \quad k \geq 1.$$

Da f er kontinuert μ-n.o. konvergerer $f_k \uparrow f$ μ-n.o., hvor $(f_k)_{k\geq 1}$ er konstrueret ud fra f, som netop beskrevet. Heraf følger ved brug af Monoton konvergens, at

$$\liminf_n \; E[f(X_n)] \geq \sup_k \; \liminf_n \; E[f_k(X_n)] = \sup_k \int f_k\, d\mu = \int f\, d\mu. \;\; \diamond$$

Kriterier for konvergens i fordeling.

Portmanteau Sætning I.
Lad (S, d) betegne et separabelt metrisk rum og lad μ være et Borel sandsynlighedsmål på S samt $(X_n)_{n\geq 1}$ en følge af stokastiske funktioner med værdier i S. Flg. udsagn er da ækvivalente.

1) $X_n \overset{\sim}{\to} \mu$

2) $\int_S g\, d\mu \leq \liminf_n E[g(X_n)]$ for alle $g \in \mathrm{b}Lip(S, d)_+$

3) $\mu(G) \leq \liminf_n P(X_n \in G)$ for alle $G \subseteq S$ åben

4) $\mu(F) \geq \limsup_n P(X_n \in F)$ for alle $F \subseteq S$ lukket.

$Lip(S, d)_+ := \{f \in C(S)_+ \mid \exists\, M > 0 : |f(x) - f(y)| \leq M \cdot d(x, y) \;\; x, y \in S\}$.

Bemærk at modsat $C(S)$ og $C(S)_+$ afhænger $Lip(S, d)_+$ eksplicit af metrikken d.

Bevis. Da 1) $\Rightarrow$ 2) er en direkte konsekvens af definitionen, og ækvivalensen mellem 3) og 4) følger ved overgang til komplementær mængden, vises kun 2) $\Rightarrow$ 3) $\Rightarrow$ 1). Antag 2) og lad G være en given åben delmængde af S. Definer

$$g_k = (k \cdot d(\cdot, G^c)) \wedge 1 \quad \text{for alle } k \geq 1.$$

Pr. konstruktion har vi $g_k \uparrow \mathbf{1}_G$, og ved brug af trekantsuligheden ses, at

$$|g_k(x) - g_k(y)| \leq k \cdot d(x, y) \quad x, y \in S,$$

d.v.s. $g_k \in \mathrm{b}Lip(S, d)_+$. Ifølge 2) og Monoton konvergens gælder derfor

$$\mu(G) = \sup_k \int_S g_k\, d\mu \leq \sup_k \lim_n E[g_k(X_n)]$$

$$\leq \liminf_n E[\mathbf{1}_G(X_n)] \leq \liminf_n P(X_n \in G).$$

Antag 3). Som før nævnt er det nok at vise, at for $f \in \mathrm{b}C(S)_+$ er

$$\int_S f\, d\mu \leq \liminf_n E[f(X_n)].$$

Lad derfor $f \in \mathrm{b}C(S)_+$ være valgt. For ethvert n har vi

$$E[f(X_n)] = \int_0^\infty P(f(X_n) > t)\, dt = \int_0^\infty P(X_n \in \{f > t\})\, dt$$

og tilsvarende for $\int f\, d\mu$, og da $\{f > t\}$ er åben fås af Fatou's lemma, at

$$\int_0^\infty \mu(f > t)\, dt \leq \int_0^\infty \liminf_n P(f(X_n) > t)\, dt \leq \liminf_n \int_0^\infty P(f(X_n) > t)\, dt,$$

hvilket giver den ønskede ulighed. $\diamond$

Til enhver Borel mængde B tilordnes mængderne

$$B^\circ := \{x \in B \mid \exists\, \epsilon > 0 : b(x,\epsilon) \subseteq B\} \text{ og } \overline{B} := \{x \in S \mid \forall\, \epsilon > 0 : b(x,\epsilon) \cap B \neq \emptyset\}.$$

(Hoffmann bruger betegnelserne $int(B)$ og $cl(B)$) D.v.s. $B^\circ \subseteq B \subseteq \overline{B}$ med lighedstegn til venstre hvis og kun hvis B er åben, og lighedstegn til højre hvis og kun hvis B er lukket. B° kaldes *det indre* af B og er den største åbne mængde indeholdt i B, og $\overline{B}$ kaldes *aflukningen* af B og er den mindste lukkede mængde, der indeholder B. $bd(B) := \overline{B} \backslash B^\circ$ kaldes *randen* af B. Ækvivalensen mellem 1), 3) og 4) kan med denne notation formuleres som

Korollar.
$X_n \xrightarrow{\sim} \mu$ hvis og kun hvis

$$\mu(B^\circ) \leq \liminf_n P(X_n \in B) \leq \limsup_n P(X_n \in B) \leq \mu(\overline{B}) \quad \forall B \in \mathcal{B}(S).$$

Specielt gælder $X_n \xrightarrow{\sim} \mu \;\Rightarrow\; \lim_n P(X_n \in B) = \mu(B)$ hvis $\mu(bd(B)) = 0$.

Vi vil nu se, hvad dette betyder i tilfældet $S = \mathbf{R}$. Benyttes korollaret på mængder af formen $B = (-\infty, x\,]$, fås, da $B^\circ = (-\infty, x\,[$ og $\overline{B} = B$, at

$$X_n \xrightarrow{\sim} \mu \;\Rightarrow\; \mu((-\infty, x\,[\,) \leq \liminf_n F_n(x) \leq \limsup_n F_n(x) \leq \mu((-\infty, x\,]\,)$$

og dermed, da $\mu(\{x\}) = \mu((-\infty, x\,]\,) - \mu((-\infty, x\,[\,)$,

$$X_n \xrightarrow{\sim} \mu \;\Rightarrow\; \lim_n F_n(x) = \mu((-\infty, x\,]\,) \quad \text{hvis } \mu(\{x\}) = 0.$$

F_n er her fordelingsfunktionen for X_n. Men som vi nu skal se, karakteriserer dette konvergens i fordeling på $\mathbf{R}$. Der gælder nemlig flg. resultat.

Konvergens i fordeling på R.
Lad $(X_n)_{n\geq 1}$ betegne en følge af stokastiske variable og lad μ betegne et Borel sandsynlighedsmål på $\mathbf{R}$. Idet F_n er fordelingsfunktionen for X_n og F_μ funktionen $x \mapsto \mu((-\infty, x\,]\,)$ er flg. punkter ækvivalente

a) $X_n \xrightarrow{\sim} \mu$

b) $F_\mu(x-) \leq \liminf_n F_n(x) \leq \limsup_n F_n(x) \leq F_\mu(x) \quad x \in \mathbf{R}$

c) $\lim_n F_n(x) = F_\mu(x)$ hvis $\mu(\{x\}) = 0$

d) $\lim_n F_n(x) = F_\mu(x)$ for $x \in \mathrm{D}$ hvor D er tæt i $\mathbf{R}$

e) $\liminf_n P(a < X_n < b) \geq \mu(\,]\,a, b\,[\,)$ for alle $-\infty < a < b < \infty$.

Bemærkning. Opskrives udsagnet i tilfældet, hvor $\mu = P_X$ for en stokastisk variabel X, fremkommer resultatet i sektion 5.9.

Bevis. Vi mangler kun at vise, at d) $\Rightarrow$ e) $\Rightarrow$ a). Lad derfor $a < b$ betegne givne reelle tal. Da D er tæt i $\mathbf{R}$, findes der følger $(a_k)_{k\geq 1}$ og $(b_k)_{k\geq 1}$ af elementer i D, så at

$$a < a_k < b_k < b \text{ og } a_k \downarrow a \text{ og } b_k \uparrow b.$$

For alle $n, k \geq 1$ gælder derfor

$$P(a < X_n < b) \geq P(a_k < X_n \leq b_k) = F_n(b_k) - F_n(a_k) \to_{n\to\infty} F_\mu(b_k) - F_\mu(a_k).$$

D.v.s.

$$\liminf_n P(a < X_n < b) \geq \sup_k (F_\mu(b_k) - F_\mu(a_k)) = \mu(\,]\,a, b\,[\,)$$

og dermed d) $\Rightarrow$ e). For at vise den manglende implikation lader vi $G \subseteq \mathbf{R}$ betegne en begrænset åben mængde. Som vist i Appendiks C findes der højst tællelig mange parvis disjunkte intervaller $(\,]\,a_i, b_i\,[\,)_{i\geq 1}$, så at $G = \bigcup_{i\geq 1}]a_i, b_i[$. Under antagelse af e) gælder derfor for ethvert $k \geq 1$

$$\sum_{j=1}^{k} \mu(\,]\,a_j, b_j\,[\,) \leq \sum_{j=1}^{k} \liminf_n P(a_j < X_n < b_j)$$

$$\leq \liminf_n \sum_{j=1}^{k} P(a_j < X_n < b_j) \leq \liminf_n P(X_n \in G),$$

d.v.s.

$$\mu(G) = \sup_k \sum_{j=1}^{k} \mu(\,]\,a_j, b_j\,[\,) \leq \liminf_n P(X_n \in G).$$

Lad dernæst G betegne en vikårlig åben mængde. Da

$$G_k := G \cap]-k, k\,[\, \uparrow G \quad \text{for } k \to \infty$$

fås af det netop viste

$$\liminf_n P(X_n \in G) \geq \sup_k \liminf_n P(X_n \in G_k) \geq \sup_k \mu(\in G_k) = \mu(G).$$

Implikationen e) $\Rightarrow$ a) følger nu af Portmanteau sætningen. $\diamond$

Regneregler for konvergens i fordeling.

Portmanteau Sætning II.
Lad (S, d) *og* $(T\,\delta)$ *betegne separable metriske rum og lad* $(X_n)_{n\geq 1}$ *og* X *h.h.v.* $(Y_n)_{n\geq 1}$ *og* Y *betegne stokastiske funktioner med værdier i* S *h.h.v.* T*. Da gælder*

1) $X_n \stackrel{\sim}{\to} X \;\Rightarrow\; f(X_n) \stackrel{\sim}{\to} f(X)$ *for enhver* $f : S \to T$*, som er Borel målelig og kontinuert* P_X*-n.o.*

2) $X_n \stackrel{\sim}{\to} X$ og X degenereret $\Rightarrow$ $X_n \to X$ i sandsynlighed.

3) $X_n \stackrel{\sim}{\to} X$, $Y_n \stackrel{\sim}{\to} Y$ og Y degenereret $\Rightarrow$ $(X_n, Y_n) \stackrel{\sim}{\to} (X, Y)$.

4) $X_n \stackrel{\sim}{\to} X$, $Y_n \stackrel{\sim}{\to} Y$ og X_n og Y_n uafhængige $n \geq 1$ $\Rightarrow$ $(X_n, Y_n) \stackrel{\sim}{\to} P_X \otimes P_Y$.

Bevis. For ethvert $g \in \mathrm{b}C(T)$ er sammensætningen $g \circ f$ Borel målelig og kontinuert P_X-n.o. Ifølge Fk 5 gælder derfor

$$E[g(f(X_n))] = E[g \circ f(X_n)] \to_{n\to\infty} \int g \circ f \, dP_X = \int g \, dP_{f(X)},$$

hvilket viser 1). I 2) antages $P(X = a) = 1$. Da $x \mapsto d(x, a) \wedge 1 \in \mathrm{b}C(S)$ fås

$$E[d(X_n, X) \wedge 1] = E[d(X_n, a) \wedge 1] \to_{n\to\infty} E[d(X, a) \wedge 1] = d(a, a) \wedge 1 = 0,$$

d.v.s. 2) er også vist. I 3) antages $P(Y = a) = 1$. Definer

$$\tilde{d}_1((x_1, y_1), (x_2, y_2)) := d(x_1, x_2) \wedge 1 + \delta(y_1, y_2) \wedge 1 \quad x_1, x_2 \in S,\ y_1, y_2 \in T.$$

$\tilde{d}_1$ er da en produktmetrik og for et vilkårligt element $g \in Lip(S \times T, \tilde{d})_+$ gælder

$$\begin{aligned} |\, E[g(X_n, Y_n)] - E[g(X, Y)] \,| &= |\, E[g(X_n, Y_n)] - E[g(X, a)] \,| \\ &\leq E[\,|g(X_n, Y_n) - g(X_n, a)|\,] + |\, E[g(X_n, a)] - E[g(X, a)] \,| \\ &\leq M \cdot E[\delta(Y_n, a) \wedge 1] + |\, E[g(X_n, a)] - E[g(X, a)] \,| \to_{n\to\infty} 0 \end{aligned}$$

da $x \mapsto g(x, a) \in \mathrm{b}C(S)$ og $Y_n \to a$ i sandsynlighed. Påstanden følger derfor af den første Portmanteu sætning.

Det generelle bevis for 4) gennemgås ikke, men det vigtige specialtilfælde, hvor $S = \mathbf{R}^n$ og $T = \mathbf{R}^m$, behandles senere i forbindelse med Kontinuitetssætningen.

Kontinuitetssætningen for karakteristiske funktioner.
Fra Mat11 vides, at en punktfølge $(x_k)_{k\geq 1}$ i $\mathbf{R}^n$ er konvergent, hvis og kun hvis

a) $(x_k)_{k\geq 1} \subseteq K$, hvor $K \subseteq \mathbf{R}^n$ er lukket og begrænset,

b) $L((x_k)_{k\geq 1})$ består præcist af et punkt.

$L((x_k)_{k\geq 1})$ betegner her mængden af *limespunkter* eller *fortætningspunkter* for følgen $(x_k)_{k\geq 1}$, d.v.s.

$$L((x_k)_{k\geq 1}) := \{x \in \mathbf{R}^n \,|\, \exists\, (k_l)_{l\geq 1} \text{ delfølge} : x_{k_l} \to x\}$$

$$(\text{tænk over dette}) = \{x \in \mathbf{R}^n \,|\, b(x,r) \cap \{x_k \,|\, k \geq 1\} \neq \emptyset \text{ for alle } r > 0\}.$$

I gængse matematiske termer udtrykkes dette ved at sige, at i det metriske rum $\mathbf{R}^n$ udstyret med den sædvanlige metrik, er en mængde kompakt, hvis og kun hvis den er lukket og begrænset. Kompakthed defineres nemlig generelt som følger.

Definition. Lad (S,d) betegne et metrisk rum. En delmængde $K \subseteq S$ siges at være *kompakt*, hvis $L((x_k)_{k\geq 1}) \cap K \neq \emptyset$ for enhver punktfølge $(x_k)_{k\geq 1}$ i K.

I ethvert metrisk rum (S,d) gælder der nu flg. generalisation af ovenstående:

En punktfølge $(x_k)_{k\geq 1}$ i S er konvergent, hvis og kun hvis flg. punkter er opfyldte

a) $(x_k)_{k\geq 1} \subseteq K$, hvor $K \subseteq S$ er kompakt,

b) $L((x_k)_{k\geq 1})$ består præcist af et punkt.

Bevis. Antag $x_k \to x$. Da $x_{k_l} \to x$ for enhver delfølge $(k_l)_{l\geq 1}$ er x det eneste limespunkt, d.v.s. b) gælder. a) er ligeledes opfyldt, hvis K er kompakt, hvor

$$K := \{x_k, x \,|\, k \geq 1\}.$$

Lad derfor $(y_k)_{k\geq 1} \subseteq K$ være givet. Da $b(x,r)^c \cap \{x_k \,|\, k \geq 1\}$ er endelig for alle $r > 0$, da $x_k \to x$, og

$$x \notin L((y_k)_{k\geq 1}) \;\Rightarrow\; \exists\, r > 0 : b(x,r) \cap \{y_k \,|\, k \geq 1\} = \emptyset,$$

er enten $x \in L((y_k)_{k\geq 1})$ eller der findes et k_0, så at $y_k = x_{k_0}$ for uendelig mange k. Men da det sidste medfører, at $x_{k_0} \in L((y_k)_{k\geq 1})$ er 'kun hvis'-delen vist .

Den anden implikation vises ved at vise, at $x_k \to x$, hvor x er det entydigt givne limespunkt. Antag derfor at dette ikke gælder, d.v.s.

$$\exists\, r > 0 \;\exists\, (k_l)_{l\geq 1} \text{ delfølge} : \; x_{k_l} \notin b(x,r) \text{ for alle } l \geq 1.$$

Kompaktheden sikrer, at $(x_{k_l})_{l\geq 1} \subseteq K$ har mindst et limespunkt $\tilde{x}$. Men da

$$L((x_{k_l})_{l\geq 1}) \subseteq L((x_k)_{k\geq 1})$$

må $\tilde{x} = x$, hvilket er umuligt, da $x_{k_l} \notin b(x,r)$ for alle $l \geq 1$. $\diamond$

Da kompakthed er defineret ved hjælp af konvergens af følger, er begrebet invariant under overgang mellem ækvivalente metrikker. Ligeledes viser en simpel overvejelse, at enhver kompakt mængde er lukket, d.v.s. specielt en Borel mængde. Kompakthed bevares endvidere under endelig forening og vilkårlig gennemsnit, d.v.s. i ethvert metrisk rum gælder

a) K_1 og K_2 kompakte $\Rightarrow$ $K_1 \cup K_2$ kompakt.

b) K_i kompakte for alle $i \in I \Rightarrow \bigcap_{i\in I} K_i$ kompakt.

Det er normalt ikke let at afgøre, om en mængde er kompakt, men endelige mængder samt, som netop vist, mængder af formen $\{x, (x_n)_{n\geq 1}\}$, hvor $x_n \to x$, er altid kompakte. Normalt er kompakte mængder meget 'små', men i $\mathbf{R}^n$ og mere generelt i polske rum (et metrisk rum (S, d) kaldes et polsk rum, hvis der findes en ækvivalent metrik (måske d selv), som er fuldstændig og separabel) findes der 'store' kompakte mængder. Mere præcist gælder:

I) *For ethvert Borel sandsynlighedsmål μ på $\mathbf{R}^n$ gælder*

$$\forall \epsilon \; \exists K \subseteq \mathbf{R}^n \text{ kompakt} : \; \mu(K^c) < \epsilon.$$

I') *For ethvert Borel sandsynlighedsmål μ på et polsk rum (S, d) gælder*

$$\forall \epsilon \; \exists K \subseteq S \text{ kompakt} : \; \mu(K^c) < \epsilon.$$

Bevis for I). Da $b(\underline{0}, k) \uparrow \mathbf{R}^n$ for $k \to \infty$ konvergerer $\mu(b(\underline{0}, k)) \uparrow 1$, og da $\overline{b(\underline{0}, k)}$ som nævnt er kompakt for alle k, følger I) umiddelbart. $\diamond$

Den samme teknik kan ikke bruges for I'), da et generelt polsk rum ikke er en tællelig og dermed en tællelig voksende foreningsmægde af kompakte mængder. Men ved brug af en alternativ karakterisation af kompakthed i polske rum kan I') dog vises. Beviset findes i afsnittet "Mål på metriske rum".

I) og I') viser, at det såkaldte stramhedsbegreb, som vi nu vil indføre, er et kompakthedsbegreb for konvergens i fordeling for stokastiske funktioner med værdier i et givet polsk rum (S, d). Men da vi i dette kursus kun skal betragte tilfældet $S = \mathbf{R}^n$, vil jeg i det følgende udelukkende koncentrere mig om dette tilfælde. Men begrebets betydning og konsekvenser overføres uændret til et vilkårligt polsk rum, men ikke til de generelle endog ikke-separable metriske rum, som Hoffmann arbejder med i sin bog.

Definition. En familie af sandsynlighedsmål $\{\mu_i \,|\, i \in I\}$ på $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ siges at være *stram* (*tight*), hvis

$$\forall \epsilon > 0, \, \exists K \subseteq \mathbf{R}^n \text{ kompakt} : \; \sup_{i\in I} \mu_i(K^c) < \epsilon,$$

og en familie af n-dimensionale stokastiske vektorer $(\underline{X}_i)_{i\in I}$ siges at være *stram*, hvis mængden af fordelingsmål $\{P_{\underline{X}_i} \,|\, i \in I\}$ udgør en stram familie, d.v.s. hvis

$$\forall \epsilon > 0, \, \exists K \subseteq \mathbf{R}^n \text{ kompakt} : \; \sup_i P(\underline{X}_i \notin K) < \epsilon.$$

Set i lyset af den simple struktur af de kompakte mængder i $\mathbf{R}^n$ kan den sidste betingelse naturligvis ækvivalent formuleres som

$$\forall \epsilon > 0, \exists\, r > 0 : \quad \sup_i P(\|\underline{X}_i\| > r) < \epsilon,$$

hvor $\|\underline{x}\|^2 := \sum_{i=1}^n x_i^2$ for alle $\underline{x} \in \mathbf{R}^n$. Markov's ulighed sikrer derfor flg. tilstrækkelige betingelse.

En familie af n-dimensionale stokastiske vektorer $(\underline{X}_i)_{i\in I}$ *er stram, hvis*

$$\sup_i E[\, \|\underline{X}_i\|^\alpha] < \infty \quad \text{for et } \alpha > 0.$$

I) viser, at enhver et-punktsmængde og derfor enhver endelig familie af mål eller stokastiske vektorer er stram. Denne observation betyder, at en følge $(\underline{X}_i)_{i\geq 1}$ er stram, hvis man kan vise, at

$$\forall \epsilon > 0, \exists\, K \subseteq \mathbf{R}^n \text{ kompakt} : \quad \limsup_i P(\underline{X}_i \notin K) \leq \epsilon.$$

eller ækvivalent

$$\forall \epsilon > 0, \exists\, r > 0 : \quad \limsup_i P(\|\underline{X}_i\| > r) \leq \epsilon.$$

Dette udnyttes blandt andet i flg. 'stramheds'-kriterium, som vil indgå i beviset for Kontinuitetssætningen.

En følge af n-dimensionale stokastiske vektorer $(\underline{X}_k)_{k\geq 1}$ *er stram hvis*

$$\forall \epsilon > 0, \exists\, a > 0 : \liminf_k E[e^{-a\cdot\|\underline{X}_k\|^2}] > 1 - \epsilon.$$

Bevis. Lad $\epsilon > 0$ være givet. Vælg $a > 0$, så at $\liminf_k E[e^{-a\cdot\|\underline{X}_k\|^2}] > 1 - \epsilon/2$. Da $x \mapsto 1 - e^{-ax^2}$ er voksende på $\mathbf{R}_+$ gælder ifølge Markov's ulighed at

$$P(\|\underline{X}_k\| > 1/\sqrt{a}) = P(1 - e^{-a\cdot\|\underline{X}_k\|^2} > 1 - e^{-1}) \leq \frac{1 - E[e^{-a\cdot\|\underline{X}_k\|^2}]}{1 - e^{-1}}$$

for alle $k \geq 1$, og da $1 - e^{-1} > 1/2$ har vi derfor pr. antagelse, at

$$\limsup_k P(\|\underline{X}_k\| > 1/\sqrt{a}) < 2(1 - (1 - \epsilon/2)) = \epsilon,$$

hvilket, som nævnt, medfører stramhed. $\diamond$

Det næste resultat viser, at stramhed er det kompakthedsbegreb vedrørende konvergens i fordeling for stokastiske vektorer, som vi har påstået. Der gælder et tilsvarende resultat i ethvert polsk rum, men den generelle sætning, der er kendt under navnet **Prokhorov's Sætning**, er noget mere kompliceret.

Helly-Bray's Sætning.
Enhver stram følge $(\underline{X}_k)_{k\geq 1}$ af n-dimensionale stokastiske vektorer har mindst et limespunkt, d.v.s. der findes en delfølge $(k_l)_{l\geq 1}$ og et Borel sandsynlighedsmål μ på $\mathbf{R}^n$, så at $\underline{X}_{k_l} \stackrel{\sim}{\to} \mu$.
En følge af n-dimensionale stokastiske vektorer $(\underline{X}_k)_{k\geq 1}$ konvergerer i fordeling, hvis og kun hvis den er stram og har præcist et limespunkt.

Første del vises kun i tilfældet $n = 1$. Tilfældet $n > 1$ klares på 'næsten' samme måde, men er dog mere kompliceret, både hvad angår opskrivning og indhold.

Bevis. Lad for ethvert $k \geq 1$ F_k betegne fordelingsfunktionen for X_k. Da mængden af rationale tal er tællelig, kan vi, da fordelingsfunktioner kun antager værdier i $[0,1]$, ved successiv udtynding konstruere en delfølge $(\sigma(k))_{k\geq 1}$, så at

$$G(r) := \lim_k F_{\sigma(k)}(r) \quad \text{eksisterer for alle } r \in \mathbf{Q}.$$

Definitionen viser umiddelbart, at for vilkårlige rationale tal $r_1 < r_2$ er

$$0 \leq G(r_1) \leq G(r_2) \leq 1;$$

og da stramhed af $(X_k)_{k\geq 1}$ i $\mathbf{R}$ kan formuleres som

$$\forall\, \epsilon > 0,\, \exists\, r > 0 : \sup_k F_k(-r) \leq \epsilon \quad \text{og} \quad \inf_k F_k(r) \geq 1 - \epsilon$$

ses endvidere, at

$$\lim_{r\to-\infty} G(r) = 0 \quad \text{og} \quad \lim_{r\to\infty} G(r) = 1.$$

Definer

$$F(x) := \inf_{r>x,\, r\in Q} G(r) \quad x \in \mathbf{R}.$$

Simpel reel analyse viser, at F er ikke-aftagende og højrekontinuert, samt at

$$\lim_{x\to-\infty} F(x) = 0 \quad \text{og} \quad \lim_{x\to\infty} F(x) = 1.$$

Lebesgue-Stieltjes målet λ_F hørende til F er derfor et sandsynlighedsmål, da

$$\lambda_F(\mathbf{R}) = \lim_{r\to\infty} \lambda_F(\,]-r, r\,]\,) = \lim_{r\to\infty}(F(r) - F(-r)) = 1,$$

og vi vil nu vise, at $X_{\sigma(k)} \stackrel{\sim}{\to} \lambda_F$. Da $F(x) = \lambda_F(\,]-\infty, x\,]\,)$ for alle $x \in \mathbf{R}$, er det ifølge formel (5.9.2) hertil nok at vise, at

$$F(x-) \leq \liminf_k F_{X_{\sigma(k)}}(x) \leq \limsup_k F_{X_{\sigma(k)}}(x) \leq F(x) \quad \text{for alle } x \in \mathbf{R}.$$

Betragt $x \in \mathbf{R}$. For alle $m \geq 1$ og rationale tal r, så at $x - 1/m < r < x$, har vi

$$F(x - 1/m) \leq G(r) = \lim_k F_{\sigma(k)}(r) \leq \liminf_k F_{X_{\sigma(k)}}(x),$$

og dermed $F(x-) = \sup_k F(x - 1/m) \leq \liminf_k F_{X_{\sigma(k)}}(x)$. Tilsvarende har vi for alle rationale tal $r > x$, at

$$\limsup_k F_{X_{\sigma(k)}}(x) \leq \lim_k F_{\sigma(k)}(r) = G(r)$$

og dermed den anden ulighed, da $F(x) = \inf_{r>x,\, r\in Q} G(r)$ pr. definition.
Anden del vises i den generelle $\mathbf{R}^n$-situation. Antag at $(\underline{X}_k)_{k\geq 1}$ konvergerer i fordeling, og lad μ betegne grænsemålet. D.v.s. μ er Borel sandsynlighedsmål på $\mathbf{R}^n$, og $\underline{X}_k \stackrel{\sim}{\to} \mu$. Enhver delfølge konvergerer derfor også mod μ, d.v.s. mængden af limespunkter består udelukkende af μ. Som vist i 5.3 har vi

$$\lim_k P(\|\underline{X}_k\| \geq r) = \mu(b(\underline{0}, r)^c)$$

for ethvert r for hvilket $\mu(bd(b(\underline{0}, r))) = 0$, d.v.s. alle pånær tællelig mange r. Specielt kan $\mu(b(\underline{0}, r)^c)$ gøres så lille som ønsket ved at vælge r stor nok og dermed

$$\forall \epsilon > 0,\ \exists\, r > 0 : \quad \limsup_k P(\|\underline{X}_k\| > r) \leq \epsilon,$$

d.v.s. $(\underline{X}_k)_{k\geq 1}$ er stram. Lad nu omvendt μ betegne det entydige limespunkt. Hvis $(\underline{X}_k)_{k\geq 1}$ ikke konvergerer i fordeling mod μ, findes der pr. definition af konvergens i fordeling, en funktion $f \in \mathrm{b}C(\mathbf{R}^n)$, et $\epsilon > 0$ og en delfølge $(k_l)_{l\geq 1}$, så at

$$|\, E[\, f(\underline{X}_{k_l})\,] - \int f\, d\mu\, | > \epsilon$$

for alle $l \geq 1$. Følgen $(\underline{X}_{k_l})_{l\geq 1}$ er igen stram, og det eksisterende limespunkt er ifølge entydigheden μ. Men dette er tydeligvis umuligt. $\diamond$

Korollar. *En stram følge af n-dimensionale stokastiske vektorer* $(\underline{X}_k)_{k\geq 1}$ *konvergerer i fordeling, hvis* $\lim_k \varphi_{\underline{X}_k}(\underline{t})$ *eksisterer for alle* $\underline{t} \in \mathbf{R}^n$.

Bevis. Antag at μ_1 og μ_2 er limespunkter, d.v.s. der findes delfølger $(k(l))_{l\geq 1}$ og $(k(m))_{m\geq 1}$, så at

$$\underline{X}_{k(l)} \stackrel{\sim}{\to} \mu_1 \quad \text{og} \quad \underline{X}_{k(m)} \stackrel{\sim}{\to} \mu_2.$$

Da

$$\underline{t} \mapsto \cos \underline{a} \cdot \underline{t} \quad \text{og} \quad \underline{t} \mapsto \sin \underline{a} \cdot \underline{t}$$

er kontinuerte og begrænsede for ethvert $\underline{a} \in \mathbf{R}^n$, betyder dette, at

$$\int e^{i\underline{a}\cdot\underline{t}}\, \mu_1(d\underline{a}) = \lim_l \varphi_{\underline{X}_{k(l)}}(\underline{t}) = \lim_k \varphi_{\underline{X}_k}(\underline{t}) = \lim_m \varphi_{\underline{X}_{k(m)}}(\underline{t}) = \int e^{i\underline{a}\cdot\underline{t}}\, \mu_2(d\underline{a})$$

for alle $\underline{t} \in \mathbf{R}^n$, d.v.s. $\hat{\mu_1} = \hat{\mu_2}$. Ifølge Entydighedssætningen for karakteristiske funktioner er μ_1 og μ_2 derfor identiske, og der er derfor højst et og dermed på grund af stramhed præcist et limespunkt. $\diamond$

Hovedsætningen om fordelingskonvergens i $\mathbf{R}^n$, der er kendt under navnet *Kontinuitetssætningen for karakteristiske funktioner*, viser, at konvergens i fordeling i

$\mathbf{R}^n$ er ækvivalent med punktvis konvergens af de tilhørende karakteristiske funktioner. Thi da enhver karakteristisk funktion er kontinuert, gælder for givne n-dimensionale stokastiske vektorer $(\underline{X}_k)_{k\geq 1}$ og $\underline{X}$, at

$$\underline{X}_k \stackrel{\sim}{\to} \underline{X} \Leftrightarrow \lim_k \varphi_{\underline{X}_k}(\underline{t}) = \varphi_{\underline{X}}(\underline{t}) \quad \text{for alle } \underline{t} \in \mathbf{R}^n,$$

eller tilsvarende for ethvert Borel sandsynlighedsmål μ på $\mathbf{R}^n$

$$\underline{X}_k \stackrel{\sim}{\to} \mu \Leftrightarrow \lim_k \varphi_{\underline{X}_k}(\underline{t}) = \hat{\mu}(\underline{t}) \quad \text{for alle } \underline{t} \in \mathbf{R}^n.$$

Specielt ses at $\underline{X}_k \stackrel{\sim}{\to} \underline{X}$ i $\mathbf{R}^n$ $\Leftrightarrow$ $\underline{t} \cdot \underline{X}_k \stackrel{\sim}{\to} \underline{t} \cdot \underline{X}$ i $\mathbf{R}$ for alle $\underline{t} \in \mathbf{R}^n$.

Kontinuitetssætningen for karakteristiske funktioner.
Lad $(\underline{X}_k)_{k\geq 1}$ betegne n-dimensionale stokastiske vektorer, så at $\lim_k \varphi_{\underline{X}_k}(\underline{t}) = \gamma(\underline{t})$ for alle $\underline{t} \in \mathbf{R}^n$, hvor $\gamma : \mathbf{R}^n \to C$ er kontinuert i $\underline{0}$. Da er $(\underline{X}_k)_{k\geq 1}$ stram, og der findes et Borel sandsynlighedsmål μ på $\mathbf{R}^n$, så at $\underline{X}_k \stackrel{\sim}{\to} \mu$. Endvidere er $\gamma(\underline{t}) = \hat{\mu}(\underline{t})$ for alle $\underline{t} \in \mathbf{R}^n$.

Bevis. Da den sidste påstand er umiddelbar, udestår kun implikationen

$$\gamma \text{ kontinuert i } \underline{0} \;\Rightarrow\; (\underline{X}_k)_{k\geq 1} \text{ stram.}$$

Ifølge Kf 1 er $|\gamma(\underline{t})| \leq \gamma(\underline{0}) = 1$ for alle $\underline{t} \in \mathbf{R}^n$. D.v.s. hvis $U_1, \ldots, U_n$ er uafhængige $N(0,2)$-fordelte stokastiske variable og $\underline{U} := (U_1, \ldots, U_n)$, sikrer γ's kontinuitet i punktet $\underline{0}$, at der til et givent ϵ findes et $a > 0$, så at

$$|\, E[\gamma(a\,\underline{U})]\,| > 1 - \epsilon,$$

thi for $a_n \to 0$ konvergerer $\gamma(a_n\,\underline{U})$ mod $\gamma(\underline{0}) = 1$ P-n.o. domineret af 1. Men da

$$\varphi_{\underline{U}}(\underline{t}) = \prod_{i=1}^{n} \varphi_{U_i}(t_i) = \prod_{i=1}^{n} exp(-t_i^2) = e^{-\|\underline{t}\|^2} \quad \text{for alle } \underline{t} \in \mathbf{R}^n,$$

fås af antagelserne, Lebesgue's Sætning samt formel (4.18.4), d.v.s. Cf 4, at

$$E[\gamma(a\,\underline{U})] = \lim_k E[\varphi_{\underline{X}_k}(a\,\underline{U})] = \lim_k E[\varphi_{\underline{U}}(a\,\underline{X}_k)] = \lim_k E[\, e^{-a^2 \cdot \|\underline{X}_k\|^2}].$$

D.v.s. $\liminf_k E[e^{-a^2 \cdot \|\underline{X}_k\|^2}] > 1 - \epsilon$, og $(\underline{X}_k)_{k\geq 1}$ er derfor stram. $\diamond$

Korollar *Lad $\underline{X}$ og $(\underline{X}_k)_{k\geq 1}$ samt $\underline{Y}$ og $(\underline{Y}_k)_{k\geq 1}$ betegne h.h.v. n og m-dimensionale stokastiske vektorer, så at $\underline{X}_k$ og $\underline{Y}_k$ er uafhængige for alle k. Da gælder*

$$\underline{X}_k \stackrel{\sim}{\to} \underline{X} \;\text{ og }\; \underline{Y}_k \stackrel{\sim}{\to} \underline{Y} \Rightarrow (\underline{X}_k, \underline{Y}_k) \stackrel{\sim}{\to} P_{\underline{X}} \otimes P_{\underline{Y}}.$$

Bevis. Da

$$\varphi_{(\underline{X}_k, \underline{Y}_k)}(\underline{t}, \underline{s}) = \varphi_{\underline{X}_k}(\underline{t}) \cdot \varphi_{\underline{Y}_k}(\underline{s}) \to_{k\to\infty} \varphi_{\underline{X}}(\underline{t}) \cdot \varphi_{\underline{Y}}(\underline{s})$$

for alle $(\underline{t}, \underline{s}) \in \mathbf{R}^n \times \mathbf{R}^m = \mathbf{R}^{n+m}$ og

$$(\underline{t}, \underline{s}) \mapsto \varphi_{\underline{X}}(\underline{t}) \cdot \varphi_{\underline{Y}}(\underline{s})$$

er kontinuert, sikrer Kontinuitetssætningen, at $((\underline{X}_n, \underline{Y}_n))_{n \geq 1}$ konvergerer i fordeling. Identifikationen af grænsemålet følger dernæst af Entydighedssætningen, da

$$(\underline{t}, \underline{s}) \mapsto \varphi_{\underline{X}}(\underline{t}) \cdot \varphi_{\underline{Y}}(\underline{s})$$

ifølge Kf 5 er den Fouriertransformerede for sandsynlighedsmålet $P_{\underline{X}} \otimes P_{\underline{Y}}$. $\diamond$

Den Centrale Grænseværdisætning.
Kontinuitetssætningen er det ideelle værktøj til at undersøge flg. konvergensproblem ofte omtalt som et CLT-problem (Central Limit Theorem).

Givet en følge af uafhængige stokastiske variable $(X_n)_{n\geq 1}$ findes der da talfølger $(a_n)_{n\geq 1} \subseteq (0,\infty)$ og $(b_n)_{n\geq 1} \subseteq \mathbf{R}$, så at

$$\frac{\sum_{k=1}^n X_k - b_n}{a_n} \overset{\sim}{\to} \text{ ikke degenereret fordeling ?}$$

Udtrykt ved hjælp af Kontinuitetssætningen er dette ækvivalent med at spørge om eksistensen af $(a_n)_{n\geq 1} \subseteq (0,\infty)$ og $(b_n)_{n\geq 1} \subseteq \mathbf{R}$, så at

$$\lim_n e^{it\,b_n/a_n} \cdot \prod_{k=1}^n \varphi_{X_k}(t/a_n) \ \to_n \ \psi(t) \quad t \in \mathbf{R},$$

hvor ψ er en karakteristisk funktion for ikke degenereret fordeling. Formuleret på denne sidste måde, er det naturligt også at lade φ_{X_k}'erne variere med n, d.v.s. i stedet for at tage udgangspunkt i en følge af uafhængige stokastiske variable, starter vi med et såkaldt *uafhængigt trekantsskema* $(X_{nj})_{1\leq j\leq n}$, d.v.s.

$$\begin{gathered} X_{11} \\ X_{21},\ X_{22} \\ \cdots\cdots \\ X_{n1},\ X_{n2},\ \ldots,\ X_{nn} \\ \cdots\cdots \end{gathered}$$

hvor $X_{n1}, X_{n2}, \ldots, X_{nn}$ er uafhængige for alle n. En følge af uafhængige variable $(X_n)_{n\geq 1}$ giver trivielt anledning til et uafhængigt trekantsskema af flg. form

$$\begin{gathered} X_1 \\ X_1,\ X_2 \\ \cdots\cdots \\ X_1,\ X_2,\ \ldots,\ X_n \\ \cdots\cdots \end{gathered}$$

Problemet er nu om der findes konstanter $(a_n)_{n\geq 1} \subseteq (0,\infty)$ og $(b_n)_{n\geq 1} \subseteq \mathbf{R}$ eller retterere $(m_{nk})_{1\leq k\leq n} \subseteq \mathbf{R}$ så at

$$\frac{\sum_{k=1}^n X_{nk} - \sum_{k=1}^n m_{nk}}{a_n} = \frac{\sum_{k=1}^n (X_{nk} - m_{nk})}{a_n}$$

konvergerer i fordeling mod en ikke degeneret fordeling, eller ækvivalent om

$$\prod_{k=1}^n \varphi_{X_{nk}-m_{nk}}(t/a_n) = exp(-\frac{it}{a_n}\sum_{k=1}^n m_{nk}) \cdot \prod_{k=1}^n \varphi_{X_{nk}}(t/a_n) \ \to \ \psi(t)$$

for alle $t \in \mathbf{R}$, hvor ψ er en karakteristisk funktion for ikke degenereret fordeling.
For at undgå at et enkelt led i summen dominerer det hele, restrikterer man sig normalt til situationer, hvor den såkaldte *Uan*-betingelse, d.v.s.

$$\lim_n \sup_{1 \le k \le n} P(|\, X_{nk} - m_{nk} \,|/a_n \ge \epsilon) = 0 \quad \text{for alle } \epsilon > 0,$$

er opfyldt. Under denne antagelse kan man vise, at det kun er bestemte fordelinger, der kan fremkomme som grænseværdi på ovennævnte måde. F.eks. Normale og Poisson fordelinger, men derimod f.eks. ikke Binomial og Uniforme fordelinger. Helt præcist er de eneste mulige grænsefordelinger under uan-betingelsen de såkaldte *uendelig delelige fordelinger* (se Hoffmann sektion 5.28).
De historisk vigtigste resultater er de *normale CLT*, d.v.s. udsagn som sikrer, at grænsefordelingen eksisterer og er en normal fordeling. Vi skal udelukkende beskæftige os med denne type, idet vi skal omtale tre normale CLT-resultater samt diskutere deres indbyrdes relationer. Som det fremgår er udgangspunktet i de tre situationer et uafhængigt trekantsskema, hvor de indgående variable alle har endeligt andet moment, og konstanterne vælges i h.h.t.

$$m_{nk} := E[X_{nk}] \quad 1 \le k \le n \quad \text{og} \quad a_n := \sqrt{\sum_{k=1}^{n} Var(X_{nk})} \quad n \ge 1.$$

I stedet for a_n benyttes betegnelsen s_n, d.v.s. $s_n^2 = \sum_{k=1}^{n} Var(X_{nk}) \quad n \ge 1$.

Den Centrale Grænseværdisætning, klassisk udgave.
Lad $(X_n)_{n \ge 1}$ betegne en iid-følge, hvor den fælles fordeling har endelig middelværdi μ og varians $\sigma^2 > 0$. Da konvergerer

$$U_n := \frac{1}{\sqrt{n\,\sigma^2}} \sum_{j=1}^{n} (X_j - \mu) \xrightarrow{\sim} N(0,1).$$

Bevis. Lad φ betegne den karakteristiske funktion for $X_1 - \mu$. Regnereglerne for karakteristiske funktioner viser, at for alle $t \in \mathbf{R}$ og $n \ge 1$ er

$$\varphi_{U_n}(t) = \varphi(\, t/\sqrt{n\,\sigma^2}\,)^n = \left(1 - \frac{n\,(1 - \varphi(\, t/\sqrt{n\,\sigma^2}\,)\,)}{n}\right)^n.$$

Men da $E[X_1 - \mu] = 0$ og $Var(X_1 - \mu) = \sigma^2$ fås af et vist Taylorudviklings resultat for karakteristiske funktioner, at

$$\varphi_{U_n}(t) \to_{n \to \infty} e^{-t^2/2} \quad t \in \mathbf{R}$$

og dermed ifølge Kontinuitetssætningen det ønskede resultat. $\diamond$

Lyapounov's Sætning.
Lad $\{X_{nj} \,|\, 1 \le j \le n\}$ betegne et uafhængigt trekantsskema, så at X_{nj}'erne har endelig første og andet moment. Sæt

$$\mu_{nj} = E[X_{nj}],\ \sigma_{nj}^2 = Var(X_{nj}) \ \text{ og } \ s_n := \sqrt{\sigma_{n1}^2 + \cdots + \sigma_{nn}^2}$$

for $n \ge 1$ og antag, at $s_n > 0$ for alle n. Hvis Lyapounov's betingelse er opfyldt, d.v.s. hvis

$$\exists\, \alpha > 2 : \quad \lim_{n\to\infty} \frac{1}{s_n^\alpha} \sum_{j=1}^n E[\,|X_{nj} - \mu_{nj}|^\alpha] = 0,$$

konvergerer

$$U_n := \frac{1}{s_n} \sum_{j=1}^n (X_{nj} - \mu_{nj}) \xrightarrow{\sim} N(0,1).$$

Lyapounov's betingelse er specielt opfyldt, hvis $s_n \to_{n\to\infty} \infty$ og $X_{nj} - \mu_{nj}$'erne er uniformt begrænsede af en konstant M. For i denne situation gælder

$$\frac{1}{s_n^3} \sum_{j=1}^n E[\,|X_{nj} - \mu_{nj}|^3] \le \frac{M}{s_n^3} \sum_{j=1}^n E[\,|X_{nj} - \mu_{nj}|^2] = \frac{M}{s_n} \to_{n\to\infty} 0.$$

I beviset kan vi uden tab af generalitet antage, at Lyapounov's betingelse er opfyldt for et α i intervallet $]2,3]$. Thi betragt for givet $n \ge 1$ produktrummet $\Omega \times E_n$, hvor $E_n := \{1, \ldots, n\}$, udstyret med produkt σ-algebraen $\mathcal{F} \times 2^{E_n}$. Definer

$$\mu_n(A) := \sum_{j=1}^n E[\,\mathbf{1}_A(\,\cdot\,, j) \cdot \left|\frac{Z_{nj}}{s_n}\right|^2] \qquad A \in \mathcal{F} \times 2^{E_n},$$

hvor $Z_{nj} := X_{nj} - \mu_{nj}$ for alle n og j. μ_n er da et sandsynlighedsmål, og ifølge korollaret til Jensen's ulighed gælder derfor for ethvert $\delta \ge 1$, at

$$\frac{1}{s_n^3} \sum_{j=1}^n E[\,|Z_{nj}|^3] = \sum_{j=1}^n E[\left|\frac{Z_{nj}}{s_n}\right| \cdot \left|\frac{Z_{nj}}{s_n}\right|^2] = \int f \, d\mu_n$$

$$\le (\int f^\delta \, d\mu_n)^{1/\delta} \le \left(\sum_{j=1}^n E[\left|\frac{Z_{nj}}{s_n}\right|^\delta \cdot \left|\frac{Z_{nj}}{s_n}\right|^2] \right)^{1/\delta} = \left(\frac{1}{s_n^{2+\delta}} \sum_{j=1}^n E[\,|Z_{nj}|^{2+\delta}] \right)^{1/\delta},$$

hvor

$$f(\omega, j) := \left|\frac{Z_{nj}(\omega)}{s_n}\right| \quad \text{for alle } (\omega, j) \in \Omega \times E_n.$$

Bevis for Lyapounov's Sætning. Ifølge Jensen's ulighed er

$$\max_{1\le j\le n} \frac{\sigma_{nj}^2}{s_n^2} = \max_{1\le j\le n} E[\Big(\frac{X_{nj} - \mu_{nj}}{s_n}\Big)^2] \le \left(\frac{1}{s_n^\alpha} \sum_{j=1}^n E[\,|X_{nj} - \mu_{nj}|^\alpha] \right)^{2/\alpha},$$

d.v.s. Lyapounov's betingelse sikrer, at $\max_{1\le j\le n} \sigma_{nj}^2/s_n^2 \to 0$ for $n \to \infty$, og da $E[(X_{nj}-\mu_{nj})/s_n] = 0$ fås af tillægget til formel (4.18.7), at

$$|\,\varphi_{(X_{nj}-\mu_{nj})/s_n}(t) - 1 + \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}\,| \le |t|^\alpha \, \frac{E[\,|X_{nj}-\mu_{nj}|^\alpha]}{s_n^\alpha}.$$

Ifølge et simpelt induktionsargument gælder for ethvert sæt af komplekse tal $a_1, \ldots, a_n$ og $b_1, \ldots, b_n$ med længde højst 1, at

$$|\prod_{j=1}^n a_j - \prod_{j=1}^n b_j| \le \sum_{j=1}^n |a_j - b_j|,$$

og benyttes dette for givet t og n så stor, at $t^2 \cdot \max_{1\le j\le n} \sigma_{nj}^2/s_n^2 < 1$, fås at

$$|\,\varphi_{U_n}(t) - \prod_{j=1}^n (1 - \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2})\,| = |\,\prod_{j=1}^n \varphi_{(X_{nj}-\mu_{nj})/s_n}(t) - \prod_{j=1}^n (1 - \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2})\,|$$

$$\le \sum_{j=1}^n |\,\varphi_{(X_{nj}-\mu_{nj})/s_n}(t) - 1 + \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}\,| \le |t|^\alpha \sum_{j=1}^n \frac{E[\,|X_{nj}-\mu_{nj}|^\alpha]}{s_n^\alpha}.$$

D.v.s. Lyapounov's betingelse viser, at

$$\lim_n \varphi_{U_n}(t) = \lim_n \prod_{j=1}^n (1 - \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}),$$

og ifølge Kontinuitetssætningen er det derfor nok at vise, at

$$\lim_n \prod_{j=1}^n (1 - \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}) = e^{-t^2/2}$$

for ethvert t. Ved brug af logaritmefunktionen ses, at dette er ækvivalent med

$$\sum_{j=1}^n \log(1 - \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}) + \frac{t^2}{2} = \sum_{j=1}^n \left(\log(1 - \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}) + \frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2}\right) \to 0,$$

hvilket følger af uligheden

$$|\log(1-x) + x| \le 2x^2 \quad \text{for } |x| \le \frac{1}{2}$$

da

$$\sum_{j=1}^n (\frac{t^2}{2}\frac{\sigma_{nj}^2}{s_n^2})^2 \le \frac{t^4}{4} \cdot \max_{1\le j\le n} \frac{\sigma_{nj}^2}{s_n^2} \sum_{j=1}^n \frac{\sigma_{nj}^2}{s_n^2} = \frac{t^4}{4} \cdot \max_{1\le j\le n} \frac{\sigma_{nj}^2}{s_n^2} \to_{n\to\infty} 0. \qquad \diamond$$

Lindeberg's Sætning.
Lad $\{X_{nj} \,|\, 1 \leq j \leq n\}$ betegne et uafhængigt trekantsskema, så at X_{nj}'erne har endelig første og andet moment. Sæt

$$\mu_{nj} = E[X_{nj}],\ \sigma_{nj}^2 = Var(X_{nj}) \ \text{ og } \ s_n := \sqrt{\sigma_{n1}^2 + \cdots + \sigma_{nn}^2}$$

for $n \geq 1$ og antag, at $s_n > 0$ for alle n. Hvis Lindeberg's betingelse er opfyldt, d.v.s. hvis

$$\lim_{n\to\infty} s_n^{-2} \int_{\{|X_{nj}-\mu_{nj}|>\epsilon s_n\}} (X_{nj} - \mu_{nj})^2 \, dP = 0 \qquad \text{for alle } \epsilon > 0$$

konvergerer

$$U_n := \frac{1}{s_n} \sum_{j=1}^{n} (X_{nj} - \mu_{nj}) \xrightarrow{\sim} N(0,1).$$

Bevis. Ifølge Kontinuitetssætningen er det nok at vise, at for ethvert $t \in \mathbf{R}$ gælder

$$\varphi_{U_n}(t) = \prod_{j=1}^{n} \varphi_{\tilde{X}_{nj}}(t) \to_{n\to\infty} e^{-t^2/2},$$

hvor $\tilde{X}_{nj} = (X_{nj} - \mu_{nj})/s_n$. Lad $t \in \mathbf{R}$ være givet. Ved at bruge en allerede benyttet vurdering af afstanden mellem to produkter af komplekse tal, hvis faktorer har længde højst 1, fås, da $\sum_{j=1}^{n} E[\tilde{X}_{nj}^2] = 1$ for alle n, at

$$|\varphi_{U_n}(t) - e^{-t^2/2}| = |\prod_{j=1}^{n} \varphi_{\tilde{X}_{nj}}(t) - \prod_{j=1}^{n} e^{-\frac{t^2}{2}\cdot E[\tilde{X}_{nj}^2]}| \leq \sum_{j=1}^{n} |\varphi_{\tilde{X}_{nj}}(t) - e^{-\frac{t^2}{2}\cdot E[\tilde{X}_{nj}^2]}|.$$

Uligheden $|e^{-x} - 1 + x| \leq x^2$ for $x \geq 0$ viser, at for alle j og n er

$$|\varphi_{\tilde{X}_{nj}}(t) - e^{-\frac{t^2}{2}\cdot E[\tilde{X}_{nj}^2]}| \leq |\varphi_{\tilde{X}_{nj}}(t) - 1 + \frac{t^2}{2} \cdot E[\tilde{X}_{nj}^2]| + \left(\frac{t^2}{2} \cdot E[\tilde{X}_{nj}^2]\right)^2,$$

d.v.s.

$$|\varphi_{U_n}(t) - e^{-t^2/2}| \leq \sum_{j=1}^{n} |\varphi_{\tilde{X}_{nj}}(t) - 1 + \frac{t^2}{2} \cdot E[\tilde{X}_{nj}^2]| + \frac{t^4}{4} \sum_{j=1}^{n} E[\tilde{X}_{nj}^2]^2,$$

og vi skal derfor blot vise, at de to summer konvergerer mod 0 hver for sig. Hvad angår den sidste, har vi da $\sum_{j=1}^{n} E[\tilde{X}_{nj}^2] = 1$, at for ethvert $\epsilon > 0$ og $n \geq 1$ er

$$\sum_{j=1}^{n} E[\tilde{X}_{nj}^2]^2 \leq \max_{1\leq j\leq n} E[\tilde{X}_{nj}^2] \leq \epsilon^2 + \sum_{j=1}^{n} E[\tilde{X}_{nj}^2,\, |\tilde{X}_{nj}| > \epsilon],$$

hvilket sammen med Lindeberg's betingelse viser den ønskede konvergens.
I forbindelse med den første sum benyttes uligheden

$$| e^{iy} - 1 - iy + y^2/2 | \leq y^2 \wedge |y|^3/6 \qquad y \in \mathbf{R}.$$

Da $\tilde{X}_{nj}$'erne har middelværdi 0, fås heraf for alle $1 \leq j \leq n$ og $\epsilon > 0$, at

$$|\varphi_{\tilde{X}_{nj}}(t) - 1 + t^2/2 \cdot E[\tilde{X}_{nj}^2] \,| = |E[e^{it\tilde{X}_{nj}} - 1 - it\tilde{X}_{nj} + t^2/2 \cdot \tilde{X}_{nj}^2] \,|$$

$$\leq t^3/6 \cdot E[\,|\tilde{X}_{nj}|^3,\ |\tilde{X}_{nj}| \leq \epsilon] + t^2 \cdot E[\tilde{X}_{nj}^2,\ |\tilde{X}_{nj}| > \epsilon]$$

$$\leq \epsilon \cdot t^3/6 \cdot E[\tilde{X}_{nj}^2] + t^2 \cdot E[\tilde{X}_{nj}^2,\ |\tilde{X}_{nj}| > \epsilon],$$

d.v.s.

$$|\sum_{j=1}^{n} E[e^{it\tilde{X}_{nj}} - 1 + t^2/2 \cdot \tilde{X}_{nj}^2] \,| \leq \epsilon \cdot t^3/6 + \sum_{j=1}^{n} E[\tilde{X}_{nj}^2,\ |\tilde{X}_{nj}| > \epsilon],$$

hvilket igen via Lindeberg's betingelse viser den ønskede konvergens. D.v.s.

$$|\varphi_{U_n}(t) - e^{-t^2/2}| \leq \sum_{j=1}^{n} |\varphi_{\tilde{X}_{nj}}(t) - 1 + \frac{t^2}{2} \cdot E[\tilde{X}_{nj}^2] \,| \to_{n\to\infty} 0,$$

og Lindeberg's Sætning er dermed vist. $\diamond$

Lyapounov's betingelse medfører Lindeberg's betingelse.
Bevis. Lad $\epsilon > 0$ være givet. For alle $1 \leq j \leq n$ og $\alpha > 2$ har vi, at

$$s_n^{-2} \int_{\{|X_{nj}-\mu_{nj}| > \epsilon s_n\}} (X_{nj} - \mu_{nj})^2 \, dP = \int_{\{\frac{|X_{nj}-\mu_{nj}|}{s_n} > \epsilon\}} \frac{(X_{nj} - \mu_{nj})^2}{s_n^2} \, dP$$

$$\leq \epsilon^{2-\alpha} \int \frac{|X_{nj} - \mu_{nj}|^\alpha}{s_n^\alpha} \, dP \leq \epsilon^{2-\alpha} \cdot s_n^{-\alpha} \cdot E[\,|X_{nj} - \mu_{nj}|^\alpha],$$

og dermed for alle $\epsilon > 0,\ n \geq 1$ og $\alpha > 2$

$$s_n^{-2} \sum_{j=1}^{n} \int_{\{|X_{nj}-\mu_{nj}| > \epsilon s_n\}} (X_{nj} - \mu_{nj})^2 \, dP \leq \epsilon^{2-\alpha} \cdot s_n^{-\alpha} \sum_{j=1}^{n} E[\,|X_{nj} - \mu_{nj}|^\alpha]. \ \diamond$$

Lindeberg's betingelse er opfyldt i iid-tilfældet
Lad μ og $\sigma^2 > 0$ betegne den fælles middelværdi og varians. Da $s_n^2 = n \cdot \sigma^2$ fås dermed for ethvert $\epsilon > 0$

$$s_n^{-2} \sum_{j=1}^{n} \int_{\{|X_{nj}-\mu_{nj}| > \epsilon s_n\}} (X_{nj} - \mu_{nj})^2 \, dP = \sigma^{-2} \int_{\{|X_1-\mu| > \sqrt{n}\cdot\epsilon\cdot\sigma\}} (X_1 - \mu)^2 \, dP,$$

som ifølge Monoton konvergens konvergerer mod 0 for $n \to \infty$. $\diamond$

Regnereglerne for konvergens i fordeling viser, at for stokastiske variable $(X_n)_{n\geq 1}$ og $(Y_n)_{n\geq 1}$ gælder

$$X_n \xrightarrow{\sim} N(0,\sigma^2) \;\text{ og }\; Y_n \to 0 \;\text{ s.s. } \Rightarrow\; X_n + Y_n \xrightarrow{\sim} N(0,\sigma^2)$$

for ethvert $\sigma^2 \in [0,\infty)$. Denne ide kan ved at opsplitte variablene i en 'lille' og en 'stor' del udnyttes til at bevise normale CLT-sætninger under svagere integrabilitetsbetingelser end ovenfor anvendt.
Opsplitningen sker ofte ved hjælp af en Borel funktion $q : \mathbf{R} \to \mathbf{R}$, idet vi skriver

$$X_{nj} = q(X_{nj}) + (X_{nj} - q(X_{nj})),$$

hvor første del spiller rollen som den lille del og den anden den store del. Et eksempel på et sådant resultat er følgende.

En normal CLT-sætning uden eksistens af momenter.
Lad $\{X_{nj} \,|\, 1 \leq j \leq n\}$ betegne et uafhængigt trekantsskema og $q : \mathbf{R} \to \mathbf{R}$ en Borel funktion, så at $m_{nj} = E[q(X_{nj})]$ og $\sigma_{nj}^2 = Var(q(X_{nj}))$ eksisterer og er endelige for alle $1 \leq j \leq n$. Antag endvidere at
i)

$$\lim_{n\to\infty} \sum_{j=1}^{n} m_{nj} = 0 \quad \text{og} \quad \lim_{n\to\infty} \sum_{j=1}^{n} \sigma_{nj}^2 = \sigma^2$$

ii)

$$q(0) = 0 = \lim_{x\to 0} \frac{x - q(x)}{x^2}$$

iii)

$$\lim_{n\to\infty} \left(\sum_{j=1}^{n} \int_{\{|X_{nj}|>\epsilon\}} (1 + q(X_{nj})^2)\, dP \right) = 0 \quad \forall\, \epsilon > 0$$

for et $0 \leq \sigma^2 < \infty$. Da konvergerer

$$U_n := \sum_{j=1}^{n} (X_{nj} - m_{nj}) \xrightarrow{\sim} N(0,\sigma^2).$$

Bevis. For ethvert n er

$$\sum_{j=1}^{n} (X_{nj} - m_{nj}) = \sum_{j=1}^{n} (X_{nj} - q(X_{nj})) + \sum_{j=1}^{n} (q(X_{nj}) - m_{nj}),$$

og ifølge de ovenstående bemærkninger er det derfor nok at vise

A) $\sum_{j=1}^{n} (q(X_{nj}) - m_{nj}) \xrightarrow{\sim} N(0,\sigma^2)$.

B) $\sum_{j=1}^{n} (X_{nj} - q(X_{nj})) \to 0$ i sandsynlighed.

Vi viser først A). Hvis $\sigma^2 = 0$ ses, at

$$E[(\sum_{j=1}^{n}(q(X_{nj}) - m_{nj}))^2] = \sum_{j=1}^{n}\sigma_{nj}^2 \to 0,$$

d.v.s.

$$\sum_{j=1}^{n}(q(X_{nj}) - m_{nj}) \to 0 \ \text{ i } L^2(P)$$

og derfor også konvergens i fordeling mod $\delta_0 = N(0,0)$. Hvis derimod $\sigma^2 > 0$, skriver vi for n så stor at $s_n^2 = \sum_{j=1}^{n}\sigma_{nj}^2 > 0$,

$$\sum_{j=1}^{n}(q(X_{nj}) - m_{nj}) = \sigma \cdot \frac{s_n}{\sigma} \cdot \frac{1}{s_n}\sum_{j=1}^{n}(q(X_{nj}) - m_{nj}),$$

hvoraf A) følger, hvis

$$\frac{1}{s_n}\sum_{j=1}^{n}(q(X_{nj}) - m_{nj}) \xrightarrow{\sim} N(0,1).$$

Dette vises ved at eftervise, at det uafhængige trekantsskema $\{q(X_{nj})\}_{1\le j\le n}$ opfylder Lindeberg's betingelse holder. Lad $\epsilon > 0$ være givet. Da

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{\{|q(X_{nj})-m_{nj}|>\epsilon\cdot s_n\}}(q(X_{nj}) - m_{nj})^2\,dP$$

$$\le \frac{2}{s_n^2}\sum_{j=1}^{n}\int_{\{|q(X_{nj})-m_{nj}|>\epsilon\cdot s_n\}} q(X_{nj})^2\,dP + \frac{2}{s_n^2}\sum_{j=1}^{n}m_{nj}^2$$

er det ifølge i) nok at vise, at

$$\lim_{n\to\infty}\sum_{j=1}^{n}\int_{\{|q(X_{nj})-m_{nj}|>\epsilon\cdot\sigma^2/2\}} q(X_{nj})^2\,dP = 0.$$

Men for n stor og dermed ifølge i) $\sup_{1\le j\le n}|m_{nj}|$ lille er for alle $1 \le j \le n$

$$\{|q(X_{nj}) - m_{nj}| > \epsilon\cdot\sigma^2/2\} \subseteq \{|q(X_{nj})| > \epsilon\cdot\sigma^2/4\} \subseteq \{|X_{nj}| > \tilde{\epsilon}\},$$

hvor $\tilde{\epsilon} > 0$ er valgt i h.h.t. ii), så at $|q(x)| > \epsilon\cdot\sigma^2/4 \Rightarrow |x| > \tilde{\epsilon}$. Påstanden og dermed A) følger nu umiddelbart af iii).

Hvad angår B) skrives for alle $\epsilon > 0$ og alle n

$$\sum_{j=1}^{n}(X_{nj} - q(X_{nj})) = \sum_{j=1}^{n}(X_{nj} - q(X_{nj}))\cdot\mathbf{1}_{\{|X_{nj}|\le\epsilon\}} + \sum_{j=1}^{n}(X_{nj} - q(X_{nj}))\cdot\mathbf{1}_{\{|X_{nj}|>\epsilon\}}$$

Da $\lim_n \sum_{j=1}^n P(|X_{nj}| > \epsilon) = 0$ for ethvert $\epsilon > 0$ ifølge iii) konvergerer

$$\sum_{j=1}^n (X_{nj} - q(X_{nj})) \cdot \mathbf{1}_{\{|X_{nj}|>\epsilon\}} \to_{n\to\infty} 0 \quad \text{i sandsynlighed,}$$

for ethvert $\epsilon > 0$. Et argument baseret på trekantsuligheden i en metrik, der svarer til konvergens i sandsynlighed, sikrer derfor, at det er nok at vise, at

$$\limsup_n E[\,|\sum_{j=1}^n (X_{nj} - q(X_{nj})) \cdot \mathbf{1}_{\{|X_{nj}|\leq\epsilon\}}|\,] \to 0 \quad \text{for } \epsilon \to 0.$$

For $\epsilon > 0$ så lille, at $|q(x) > |x|/2$ for $0 < |x| \leq \epsilon$ (her benyttes ii)) gælder nu

$$|\sum_{j=1}^n (X_{nj} - q(X_{nj})) \cdot \mathbf{1}_{\{|X_{nj}|\leq\epsilon\}}| \leq \sum_{j=1}^n \frac{|X_{nj} - q(X_{nj})|}{q(X_{nj})^2} \cdot q(X_{nj})^2 \cdot \mathbf{1}_{\{0<|X_{nj}|\leq\epsilon\}}$$

$$\leq M_\epsilon \sum_{j=1}^n q(X_{nj})^2 \cdot \mathbf{1}_{\{0<|X_{nj}|\leq\epsilon\}} \quad \text{hvor} \quad M_\epsilon := \sup_{0<|x|\leq\epsilon} \frac{|x - q(x)|}{q(x)^2}.$$

D.v.s. for alle n er

$$E[\,|\sum_{j=1}^n (X_{nj} - q(X_{nj})) \cdot \mathbf{1}_{\{|X_{nj}|\leq\epsilon\}}|\,]$$

domineret af

$$M_\epsilon \cdot \sum_{j=1}^n E[q(X_{nj})^2] = M_\epsilon \cdot \sum_{j=1}^n (\sigma_{nj}^2 + m_{nj}^2),$$

som ifølge i) konvergerer mod $M_\epsilon \cdot \sigma^2$ for $n \to \infty$. Men da

$$\sup_{0<|x|\leq\epsilon} \frac{|x - q(x)|}{q(x)^2} = \sup_{0<|x|\leq\epsilon} \frac{|x - q(x)|}{x^2} \cdot \frac{x^2}{q(x)^2} \leq 4 \cdot \sup_{0<|x|\leq\epsilon} \frac{|x - q(x)|}{x^2},$$

konvergerer M_ϵ mod 0 for $\epsilon \to 0$, hvilket viser B) og dermed sætningen. $\diamond$

Betingede middelværdier.
Som optakt til emnet betingede middelværdier vender vi kort tilbage til begrebet σ-algebraer. Allerede ved introduktionen af et sandsynlighedsfelt $(\Omega, \mathcal{F}, P)$, omtaltes elementerne i hændelsessystemet $\mathcal{F}$, som de mængder vi var interesseret i og derfor 'kendte'. Dette synspunkt er udgangspunktet for den måde, vi i det følgende skal betragte σ-algebraer eller rettere del σ-algebraer i et givent måleligt rum $(\Omega, \mathcal{F})$. Pr. definition er en del σ-algebra $\mathcal{B}$ i $\mathcal{F}$ en σ-algebra i Ω, hvis elementer alle er ligger i $\mathcal{F}$, d.v.s. $\mathcal{B} \subseteq \mathcal{F}$. En sådan del σ-algebra $\mathcal{B}$ tænkes at modellere en *informationsmængde* i den forstand, at elementerne i $\mathcal{B}$ er de mængder, vi kender, d.v.s. vi kan afgøre, om de indtræffer eller ej; og en variabel X siges derfor at være *observerbar* på baggrund af informationsmængden $\mathcal{B}$, hvis den er $\mathcal{B}$-målelig. Stabilitetskravene til en σ-algebra passer godt til informationssynspunktet, thi kender vi en hændelse A, så kender vi jo også A^c, og kender vi A_i for ethvert i, så er det nærliggende at sige, at vi også kender $\bigcup_{i=1}^{\infty} A_i$, thi denne indtræffer jo præcist, hvis mindst en af A_i'erne indtræffer. I dette 'informationssprog' opfattes σ-algebraen $\{\emptyset, \Omega\}$ som svarende til ingen information i modsætning til $\mathcal{F}$, som tolkes som fuld information.
Ofte vil informationsmængden være givet ved, at vi kender værdierne af en eller flere stokastiske variable $X_1, \ldots, X_n$, d.v.s. vi kan afgøre, om hændelserne

$$\{(X_1, \ldots, X_n) \in A\} \qquad A \in \mathcal{B}(\mathbf{R}^n)$$

indtræffer eller ej. Dette svarer til, at informationsmængden er σ-algebraen frembragt at X_i'erne, d.v.s. $\sigma(X_1, \ldots, X_n)$. Denne er klart en del σ-algebra i $\mathcal{F}$, da X_i'erne er stokastiske variable. Dens størrelsen afhænger af, hvor komplicerede $X_1, \ldots, X_n$ er, og som vist i Faktoriseringssætningen (se Øvelse 8 eller sektion 6.4), er enhver variabel, som er $\sigma(X_1, \ldots, X_n)$-målelig, d.v.s. kendt på basis af denne information, en funktion af $(X_1, \ldots, X_n)$, og dermed på formen

$$\varphi(X_1, \ldots, X_n) \text{ hvor } \varphi : \mathbf{R}^n \to \mathbf{R} \text{ er Borel målelig.}$$

Begrebet betinget middelværdi af en variabel X indføres nu som en formalisering af det 'bedste skøn' over X på basis af en given informationsmængde $\mathcal{B}$. Antag først at $\mathcal{B}$ er på formen $\sigma(A_1, \ldots, A_n)$, hvor $A_1, \ldots, A_n$ udgør en målelig partition af Ω, d.v.s. A_i'erne ligger i $\mathcal{F}$, er parvis disjunkte og udfylder Ω. Da enhver reel $\mathcal{B}$-målelig variabel er på formen $\sum_i \alpha_i \cdot \mathbf{1}_{A_i}$, er angivelsen af det bedste skøn over en variabel X givet $\mathcal{B}$ derfor ækvivalent med en beregningsformel for α_i'erne. En sådan beregningsformel afhænger naturligvis helt af, hvordan vendingen 'bedste skøn' tolkes, men i forlængelse af forståelsen af begrebet middelværdi som en form for midling af værdierne forekommer det nærliggende at vedtage, at α_i'erne skal bestemmes ved formlen

$$\alpha_i^X := \frac{1}{P(A_i)} \int_{A_i} X \, dP \;\text{ hvis } P(A_i) > 0, \text{ og } \alpha_i^X := 0 \;\text{ ellers.}$$

Den foreslåede beregningsformel kræver en vis integrabilitet af X, og vi vil her nøjes med at betragte variable X i $L^1(P)$, thi i så fald er $\sum_i \alpha_i^X \cdot \mathbf{1}_{A_i}$ en vel defineret $\mathcal{B}$-målelig variabel, der igen ligger i $L^1(P)$ og opfylder (husk at ethvert $A \in \mathcal{B}$ er en foreningsmængde af visse af A_i'erne)

$$\int_A X\,dP = \int_A \sum_i \alpha_i^X \cdot \mathbf{1}_{A_i}\,dP \quad \text{for alle } A \in \mathcal{B}.$$

Dette giver nu anledning til flg. generelle eksistensspørgsmål:

Givet en del σ-algebra $\mathcal{B}$ og en variabel $X \in L^1(P)$ findes der da en variabel $X_{\mathcal{B}}$ i $L^1(P,\mathcal{B}) := \{Y \in L^1(P) \,|\, Y \text{ er } \mathcal{B} \text{ målelig}\}$, som opfylder

$$\int_A X\,dP = \int_A X_{\mathcal{B}}\,dP \quad \text{for alle } A \in \mathcal{B},$$

og i hvor høj grad er en sådan entydigt bestemt ?

Det sidste er det letteste, for hvis $X_{\mathcal{B}}$ og $\tilde{X}_{\mathcal{B}}$ er to elementer i $L^1(P,\mathcal{B})$, der begge opfylder integralbetingelsen, så er $\{X_{\mathcal{B}} \neq \tilde{X}_{\mathcal{B}}\}$ ifølge Proposition 5 en P-nulmængde, d.v.s.

$$X_{\mathcal{B}} = \tilde{X}_{\mathcal{B}} \quad P\text{-n.o.}$$

Denne entydighed op til P-nulmængder betyder, at det har mening at bruge notationen $E[X \,|\, \mathcal{B}]$ om enhver variabel, der opfylder ovenstående krav, og en sådan kaldes *en betinget middelværdi* af X givet $\mathcal{B}$. D.v.s. en betinget middelværdi af en variabel X er observerbar m.h.t. den betragtede informations σ-algebra, og dens integral m.h.t. P over ethvert element i informationsmængden er lig integralet af X over den samme mængde.

For at vise eksistensen vil vi først studere kvadratisk integrable variable. Da underrummet $L^2(P,\mathcal{B})$ i $L^2(P)$ er lukket under konvergens i kvadratisk middel, findes der ifølge Projektionssætningen til ethvert $X \in L^2(P)$ et $P_{\mathcal{B}}X \in L^2(P,\mathcal{B})$, så at

$$X - P_{\mathcal{B}}X \in L^2(P,\mathcal{B})^{\perp};$$

og da $\mathbf{1}_A \in L^2(P,\mathcal{B})$ for alle $A \in \mathcal{B}$ er $X - P_{\mathcal{B}}X$ og $\mathbf{1}_A$ derfor orthogonale for alle $A \in \mathcal{B}$, d.v.s.

$$\int_A X\,dP = \int_A X_{\mathcal{B}}\,dP \quad \text{for alle } A \in \mathcal{B}.$$

$P_{\mathcal{B}}X$ opfylder altså de krav, der stilles til en betinget middelværdi, og der gælder yderligere flg. lemma.

Lemma BM 1 *For alle X, $Y \in L^2(P)$ er*

$$P_{\mathcal{B}}X - P_{\mathcal{B}}Y = P_{\mathcal{B}}(X - Y) \quad P\text{-n.o.} \;\text{ og }\; E[\,|P_{\mathcal{B}}X|\,] \leq E[\,|X|\,].$$

Bevis. Første del følger umiddelbart af, at $L^2(P,\mathcal{B})^\perp$ er stabil under addition. Hvad angår anden del defineres for et vilkårligt X i $L^2(P)$ $A_+ := \{P_\mathcal{B}X \geq 0\}$. D.v.s. A_+ og A_+^c ligger i $\mathcal{B}$, og der gælder derfor

$$E[\,|P_\mathcal{B}X|\,] = \int_{A_+} P_\mathcal{B}X\,dP - \int_{A_+^c} P_\mathcal{B}X\,dP =$$

$$\int_{A_+} X\,dP - \int_{A_+^c} X\,dP \leq \int_{A_+} |X|\,dP - \int_{A_+^c} |X|\,dP = E[\,|X|\,]. \qquad \diamond$$

Eksistensen af den betingede middelværdi for variable i $L^1(P)$ er nu let.

Thi vælg for et $X \in L^1(P)$ en følge $(X_n)_{n\geq 1} \subseteq L^2(P)$, så at $X_n \to X$ i $L^1(P)$. $X_n := X \cdot \mathbf{1}_{\{|X|\leq n\}}$ for $n \geq 1$ kan f.eks. bruges. Da

$$E[\,|\,E[X_n\,|\,\mathcal{B}\,] - E[X_m\,|\,\mathcal{B}\,]\,|\,] = E[\,|\,E[X_n - X_m\,|\,\mathcal{B}\,]\,|\,] \leq E[\,|X_n - X_m|\,]$$

ifølge Lemma 17, udgør $(E[X_n\,|\,\mathcal{B}\,])_{n\geq 1}$ en Cauchy-følge i $L^1(P,\mathcal{B})$, og den konvergerer derfor i P-middel mod et element i $L^1(P,\mathcal{B})$. Denne grænseværdi $X_\mathcal{B}$ er en betinget middelværdi af X givet $\mathcal{B}$, thi benyttes implikationen

$$Z_n \to Z \text{ i } L^1(P) \;\Rightarrow\; \int_A Z_n\,dP \to \int_A Z\,dP \;\text{ for alle } A \in \mathcal{F}$$

på følgerne $(X_n)_{n\geq 1}$ og $(E[X_n\,|\,\mathcal{B}\,])_{n\geq 1}$ og hændelser $A \in \mathcal{B}$ fås

$$\int_A X\,dP = \lim_n \int_A X_n\,dP = \lim_n \int_A E[X_n\,|\,\mathcal{B}\,]\,dP = \int_A X_\mathcal{B}\,dP.$$

Eksistens og entydighed af betingede middelværdier er hermed vist, d.v.s.

Proposition Bm 1 *For enhver del σ-algebra $\mathcal{B}$ i $\mathcal{F}$ findes der til ethvert $X \in L^1(P)$ et element $E[X\,|\,\mathcal{B}\,] \in L^1(P,\mathcal{B})$, så at*

$$\int_A X\,dP = \int_A E[X\,|\,\mathcal{B}\,]\,dP \;\text{ for alle } A \in \mathcal{B},$$

og $E[X\,|\,\mathcal{B}\,]$ er entydigt bestemt P-n.o. D.v.s. op til P-nulmængder findes der netop en variabel Y, som opfylder

a) Y er $\mathcal{B}$-målelig og integrabel.

b) $E[X,\,A] = E[Y,\,A]$ for alle $A \in \mathcal{B}$.

Vi vil nu for en vilkårlig del σ-algebra $\mathcal{B}$ i $\mathcal{F}$ nærmere studere afbildningen

$$L^1(P) \ni X \mapsto E[X\,|\,\mathcal{B}\,] \in L^1(P,\mathcal{B}).$$

Afbildningen, der kaldes *betinget middelværdi dannelse* m.h.t. $\mathcal{B}$, har en række vigtige egenskaber, hvoraf de fleste er angivet nedenfor. Sektionerne 6.8, 6.9 og

6.10 indeholder endnu flere, men de er alle simple konsekvenser af de nedenstående. Det er dog værd at understrege, at Hoffmann betragter variable i $L(P)$, hvor vi her kun ser på elementer i $L^1(P)$. Men igen er udvidelsen ikke vanskelig.

Betinget middelværdi dannelse bevarer middelværdi og er lineær og voksende (eller rettere P-lineær og P-voksende), d.v.s. for $X, Y \in L^1(P)$ og $a \in \mathbf{R}$ gælder

Bm 0 $E[\,E[X \,|\, \mathcal{B}\,]\,] = E[X]$.

Bm 1 $E[X + Y \,|\, \mathcal{B}\,] = E[X \,|\, \mathcal{B}\,] + E[Y \,|\, \mathcal{B}\,]$ og $E[aX \,|\, \mathcal{B}\,] = a \cdot E[X \,|\, \mathcal{B}\,]$ P-n.o.

Bm 2 $P(X \geq Y) = 1 \Rightarrow P(E[X \,|\, \mathcal{B}\,] \geq E[Y \,|\, \mathcal{B}\,]) = 1$ og tilsvarende med $>$.

Bevis. Antag $P(X \geq Y) = 1$. For alle $A \in \mathcal{B}$ er

$$\int_A E[X \,|\, \mathcal{B}\,]\, dP = \int_A X\, dP \geq \int_A Y\, dP = \int_A E[Y \,|\, \mathcal{B}\,]\, dP,$$

hvorefter $P(E[X \,|\, \mathcal{B}\,] \geq E[Y \,|\, \mathcal{B}\,]) = 1$ følger af Proposition 5. Betragtes specielt

$$A = \{E[X \,|\, \mathcal{B}\,] = E[Y \,|\, \mathcal{B}\,]\} \in \mathcal{B}$$

er de to yderpunkter ens. Men under antagelsen $P(X > Y) = 1$ har vi for alle $B \in \mathcal{F}$, at

$$\int_B X\, dP = \int_B Y\, dP \Rightarrow P(B) = 0.$$

D.v.s. $P(X > Y) = 1 \Rightarrow P(E[X \,|\, \mathcal{B}\,] > E[Y \,|\, \mathcal{B}\,]) = 1$ $\diamond$

Endvidere bevarer afbildningen 'konstanter' idet

Bm 3 $X = c$ P-n.o. $\Rightarrow$ $E[X \,|\, \mathcal{B}\,] = c$ P-n.o. for ethvert $c \in \mathbf{R}$.

Kombineres Bm 2 og 3 fås for $X \in L^1(P)$:

Bm 4 $P(X \in I) = 1 \Rightarrow P(E[X \,|\, \mathcal{B}\,] \in I) = 1$ for ethvert interval $I \subseteq \mathbf{R}$.

Bevis. Antag f.eks. $I = [a, b]$ de øvrige intervaltyper går analogt. Ved brug af Bm 2 og 3 har vi

$$P(X \in I) = 1 \Rightarrow P(X \leq b) = P(X \geq a) = 1 \Rightarrow$$

$$P(E[X \,|\, \mathcal{B}\,] \leq b) = P(E[X \,|\, \mathcal{B}\,] \geq a) = 1 \Rightarrow P(E[X \,|\, \mathcal{B}\,] \in [a, b]) = 1. \quad \diamond$$

Bm 2 og 4 sikrer flg. variant af Jensen's ulighed for betingede middelværdier.

Bm 5 Lad $\varphi : \mathbf{R} \to \mathbf{R}$ være en Borel funktion, som er konveks på et åbent interval $I \subseteq \mathbf{R}$. Lad $X \in L^1(P)$ være givet, så at $P(X \in I) = 1$ og $\varphi(X) \in L^1(P)$, da er

$$\varphi(E[X \,|\, \mathcal{B}\,]) \leq E[\varphi(X) \,|\, \mathcal{B}\,] \quad P\text{-n.o.}$$

Bevis. Da φ er konveks på det åbne interval I, eksisterer der, som vist i appendiks A, en følge $(l_n)_{n \geq 1}$ af affine funktioner, så at

$$l_n \leq \varphi \text{ på } I \text{ og } \varphi(x) = \sup_n l_n(x) \text{ for alle } x \in I,$$

d.v.s. specielt $\varphi(X) = l_n(X)$ $P-$n.o. for alle $n \geq 1$, da $P(X \in I) = 1$. Da $\mathcal{N}_P$ er stabil under tællelig forening, gælder derfor ifølge Bm 1, 2 og 3, at

$$E[\varphi(X) \,|\, \mathcal{B}\,] \geq \sup_n E[l_n(X) \,|\, \mathcal{B}\,] = \sup_n l_n(E[X \,|\, \mathcal{B}\,]) \quad P\text{-n.o.}$$

Men ifølge Bm 4 er $P(E[X \,|\, \mathcal{B}\,] \in I) = 1$ og dermed

$$\sup_n l_n(E[X \,|\, \mathcal{B}\,]) = \varphi(E[X \,|\, \mathcal{B}\,]) \quad P\text{-n.o.} \ \diamond$$

Bruges Bm 5 på funktionerne $x \mapsto |x|^p$ for $p \geq 1$ fås for alle $p \geq 1$

$$E[X \,|\, \mathcal{B}\,] \in L^p(P) \text{ hvis } X \in L^p(P) \text{ og } \|\, E[X \,|\, \mathcal{B}\,] \,\|_p \leq \|X\|_p \text{ d.v.s.}$$

Bm 6 $X \mapsto E[X \,|\, \mathcal{B}\,]$ er en lineær kontraktion i $L^p(P)$ for ethvert $p \geq 1$.

Bm 2 og 6 medfører flg. konvergensresultater for variable $(X_n)_{n\geq 1}$ og X i $L^1(P)$.

Bm 7 $X_n \uparrow (\downarrow) X$ P-n.o.$\Longrightarrow E[X_n \,|\, \mathcal{B}\,] \uparrow (\downarrow) E[X \,|\, \mathcal{B}]$ P-n.o. og i P-middel.
Bevis. Antag $X_n \uparrow X$ P-n.o. og dermed $X_n \to X$ i P-middel. Ifølge Bm 2 gælder derfor P-n.o.

$$E[X_n \,|\, \mathcal{B}\,] \uparrow \sup_n E[X_n \,|\, \mathcal{B}\,] \leq E[X \,|\, \mathcal{B}\,],$$

d.v.s. specielt at $E[X_n \,|\, \mathcal{B}\,] \to \sup_n E[X_n \,|\, \mathcal{B}\,]$ i P-middel. Men ifølge Bm 6 gælder også $E[X_n \,|\, \mathcal{B}\,] \to E[X \,|\, \mathcal{B}\,]$ i P-middel, d.v.s.

$$\sup_n E[X_n \,|\, \mathcal{B}\,] = E[X \,|\, \mathcal{B}\,] \ \ P\text{-n.o.} \quad \diamond$$

Bm 8 Hvis $X_n \geq 0$ og $X \leq \liminf_n X_n$ P-n.o., er

$$E[X \,|\, \mathcal{B}\,] \leq \liminf_n E[X_n \,|\, \mathcal{B}\,] \quad P\text{-n.o.}$$

Specielt $E[\liminf_n X_n \,|\, \mathcal{B}\,] \leq \liminf_n E[X_n \,|\, \mathcal{B}\,]$ P-n.o. hvis $\liminf_n X_n \in L^1(P)$.
Bevis. Definer $Y_n := \inf_{k\geq n} X_k$ for $n \geq 1$. Da

$$X \wedge Y_n \uparrow X \wedge \liminf_n X_n = X \ P\text{-n.o. og } \ Y_n,\, X \wedge Y_n \in L^1(P) \text{ for alle } n,$$

fås af Bm 2 og 7, at

$$E[X \,|\, \mathcal{B}\,] = \sup_n E[X \wedge Y_n \,|\, \mathcal{B}\,] \leq \sup_n E[Y_n \,|\, \mathcal{B}\,]$$

$$\leq \sup_n \inf_{k\geq n} E[X_k \,|\, \mathcal{B}\,] = \liminf_n E[X_n \,|\, \mathcal{B}\,] \ \ P\text{- n.o.} \ \diamond$$

Bm 9 Hvis $X_n \to X$ P-n.o. og $|X_n| \le Y$ P-n.o. for et $Y \in L^1(P)$ konvergerer

$$E[X_n \,|\, \mathcal{B}\,] \;\to\; E[X \,|\, \mathcal{B}\,] \quad P\text{-n.o. og i } P\text{-middel.}$$

Bevis. Da $X_n \to X$ i P-middel følger P-middel konvergensen af 6), og konvergensen P-n.o. følger ved at bruge Bm 8 på følgerne $(Y - X_n)_{n\ge1}$ og $(Y + X_n)_{n\ge1}$. Detaljerne overlades til læseren. ◇

De to næste resultater viser, at $\mathcal{B}$-målelige variable behandles som 'konstanter'.

Bm 10 $E[U \,|\, \mathcal{B}\,] = U$ P-n.o. for ethvert $U \in L^1(P, \mathcal{B})$; og for $X \in L^1(P)$ og $\mathcal{B}$-målelige variable U_1 og U_2 gælder

$$P(U_1 \le X \le U_2) = 1 \;\Rightarrow\; P(U_1 \le E[X \,|\, \mathcal{B}\,] \le U_2) = 1.$$

Bevis. Første er del er umiddelbar og overlades til læseren. Da

$$P(X \wedge n \le U_2 \wedge n) = 1 \;\text{ og }\; |U_2 \wedge n| \le n + |X| \in L^1(P)$$

for ethvert $n \ge 1$ fås af Bm 2, at

$$E[X \wedge n \,|\, \mathcal{B}] \le E[U_2 \wedge n \,|\, \mathcal{B}] = U_2 \wedge n \le U_2 \quad P\text{-n.o.}$$

Lader vi nu $n \to \infty$ fås ved brug af Bm 7, at $P(E[X \,|\, \mathcal{B}\,] \le U_2) = 1$. Den anden halvdel følger tilsvarende. ◇

Bm 11 $E[U \cdot X \,|\, \mathcal{B}\,] = U \cdot E[X \,|\, \mathcal{B}\,]$ P-n.o., for U $\mathcal{B}$-målelig og X, $U \cdot X \in L^1(P)$.

Bevis. Da $U \cdot X \in L^1(P)$ medfører, at $U^{\pm} \cdot X^{\pm}$ alle ligger i $L^1(P)$, sikrer lineariteten, at det er nok at betragte ikke-negative U og X, og da målelighedsbetingelserne klart er opfyldte, udestår blot at vise, at

$$\int_B U \cdot E[X \,|\, \mathcal{B}\,]\, dP = \int_B U \cdot X \, dP \;\text{ for alle } B \in \mathcal{B}.$$

Men dette følger ved brug af Standardbeviset ved først at antage, at U er en indikatorfunktion, dernæst en simpel funktion, hvorefter man går til grænsen ved hjælp af Monoton konvergens. Detaljerne overlades til læseren. ◇

De næste fire egenskaber er af en lidt anden natur. De første tre omhandler betingning med uafhængig information, og den sidste er reglen om *successiv betingning*. X betegner her et vilkårligt element i $L^1(P)$ og $\mathcal{B}_1$ endnu en del σ-algebra i $\mathcal{F}$. Bemærk at Bm 12 er et specialtilfælde af Bm 13.

Bm 12 $E[X \,|\, \mathcal{B}\,] = E[X]$ P-n.o. hvis X og $\mathcal{B}$ er uafhængige.

Bm 13 $E[X \,|\, \sigma(\mathcal{B} \cup \mathcal{B}_1)\,] = E[X \,|\, \mathcal{B}\,]$ P-n.o. hvis $(X, \mathcal{B})$ og $\mathcal{B}_1$ er uafhængige.

Bm 14 $E[H(X, Y) \,|\, \mathcal{B}\,] = \tilde{H}(Y)$ P-n.o. hvis X og $\mathcal{B}$ er uafhængige og $Y \in \mathrm{M}(\mathcal{B})$ og $\tilde{H}(y) := E[H(X, y)]$, hvor $H : \mathbf{R}^2 \to \mathbf{R}$ er begrænset og Borel målelig.

Bm 15 $E[\,E[X\,|\,\mathcal{B}\,]\,|\,\mathcal{B}_1] = E[X\,|\,\mathcal{B}_1\,]$ P-n.o. hvis $\mathcal{B}_1 \subseteq \mathcal{B}$.

Bevis. Bm 12 og 15 overlades til læseren. I Bm 14 viser et linearitetsargument, at det er nok at se på ikke-negative H. Da måleligheden og integrabiliteten følger af Tonelli's sætning, mangler vi kun at vise, at for givet $B \in \mathcal{B}$ er

$$\int_A \tilde{H}(Y)\,dP = \int_A H(X,Y)\,dP.$$

Men lader vi Z betegne den stokastiske variabel $\mathbf{1}_A$ fås ved gentagen brug af den lille transformationssætning og Tonelli's sætning, at

$$\int_A \tilde{H}(Y)\,dP = \int \tilde{H}(Y)\cdot \mathbf{1}_A\,dP = \int \tilde{H}(y)\cdot z\,P_{Y,Z}(dy\,dz)$$

$$= \int \{\int H(x,y)\cdot z\,P_X(dx)\}\,P_{(Y,Z)}(dy\,dz) = \int H(x,y)\cdot z\,P_X \otimes P_{(Y,Z)}(dx\,dy\,dz)$$

$$= \int H(x,y)\cdot z\,P_{(X,Y,Z)}(dx\,dy\,dz) = \int H(X,Y)\cdot \mathbf{1}_A\,dP = \int_A H(X,Y)\,dP$$

Hvad angår Bm 13, viser et nyt linearitetsargument, at vi kan og vil antage, at X er ikke-negativ. Da både måleligheden og integrabiliteten igen er oplagt, mangler vi kun at vise, at

$$\int_C X\,dP = \int_C E[X\,|\,\mathcal{B}\,]\,dP \quad \text{for alle } C \in \sigma(\mathcal{B}\cup\mathcal{B}_1).$$

Men da begge sider definerer et endeligt mål på $\sigma(\mathcal{B}\cup\mathcal{B}_1)$ med samme masse $E[X]$, behøver vi kun at vise ligheden for C af formen $A \cap B$, hvor $A \in \mathcal{B}$ og $B \in \mathcal{B}_1$, thi mængden af disse er stabil under endelig gennemsnit og frembringer $\sigma(\mathcal{B}\cup\mathcal{B}_1)$. Men for $A \in \mathcal{B}$ og $B \in \mathcal{B}_1$ gælder ifølge den antagede uafhængighed, at

$$\int_{A\cap B} E[X\,|\,\mathcal{B}\,]\,dP = \int E[X\,|\,\mathcal{B}]\cdot \mathbf{1}_A\cdot \mathbf{1}_B\,dP = P(B)\int E[X\,|\,\mathcal{B}\,]\cdot \mathbf{1}_A\,dP$$

$$= P(B)\int X\cdot \mathbf{1}_A\,dP = \int X\cdot \mathbf{1}_A\cdot \mathbf{1}_B\,dP = \int_{A\cap B} X\,dP. \quad \diamond$$

Som det fremgår af Bm 4 og 6 afbilder enhver betinget middelværdi mængder af formen

$$\{X \in L^1(P)\,|\,P(|X| \le M) = 1\} \text{ og } \{X \in L^1(P)\,|\,E[\,|X|] \le M\} \text{ hvor } 0 < M < \infty$$

ind i sig selv. Øvelse 17 medfører derfor flg. vigtige egenskab, hvor $(\mathcal{B}_n)_{n\ge 1}$ er del σ-algebraer i $\mathcal{E}$ og $\mathcal{H}$ en delmængde af $L^1(P)$.

Bm 16 $\mathcal{H}$ uniformt integrabel $\Rightarrow$ $\{\,E[X\,|\,\mathcal{B}_n\,]\,|\,X \in \mathcal{H},\ n \ge 1\}$ unif. integrabel.

Tilfældet hvor $\mathcal{H}$ består af en enkelt variabel, er specielt vigtigt, d.v.s.

Bm 16' $\{ E[X \mid \mathcal{B}_n] \mid n \geq 1\}$ er uniformt integrabel for alle $X \in L^1(P)$.

Bevis. Lad $X \in L^1(P)$ være givet. Da $\{|E[X \mid \mathcal{B}_n]| \geq K\} \in \mathcal{B}_n$ for alle n og K fås af regnereglerne for betingede middelværdier, at

$$\int_{\{|E[X \mid \mathcal{B}_n]| \geq K\}} |E[X \mid \mathcal{B}_n]| \, dP \leq \int_{\{|E[X \mid \mathcal{B}_n]| \geq K\}} E[\,|X| \mid \mathcal{B}_n] \, dP$$

$$= \int_{\{|E[X \mid \mathcal{B}_n]| \geq K\}} |X| \, dP \leq \int_{\{E[\,|X| \mid \mathcal{B}_n] \geq K\}} |X| \, dP$$

hvilket giver det ønskede, da $\{|X|\}$ er uniformt integrabel og

$$P(\{E[\,|X| \mid \mathcal{B}_n] \geq K\}) \leq E[\,E[\,|X| \mid \mathcal{B}_n]]/K \leq E[\,|X|]/K \to_{K \to \infty} 0. \quad \diamond$$

Bemærk at Bm 16 kunne være vist på samme måde.

Vi afslutter med at se lidt nærmere på tilfældet, hvor $\mathcal{B} = \sigma(Y)$ for en målelig variabel Y med værdier i et måleligt rum $(E, \mathcal{E})$. Her skrives normalt $E[\,\cdot \mid Y]$ i stedet for $E[\,\cdot \mid \sigma(Y)]$. Ifølge faktoriseringssætningen findes der til ethvert X i $L^1(P)$ en funktion $\varphi \in \mathrm{M}(\mathcal{E})$, generelt afhængig af både X, Y og P, så at

$$E[X \mid Y] = \varphi(Y),$$

og da

$$\int_B \varphi \, dP_Y = \int_{\{Y \in B\}} \varphi(Y) \, dP = \int_{\{Y \in B\}} X \, dP = \int x \cdot \mathbf{1}_B(y) \, P_{(X,Y)}(dx \, dy)$$

for alle $B \in \mathcal{E}$, ses dels, at φ er entydigt bestemt P_Y-n.o., samt at der gælder

Bm 17 Lad X og Z betegne elementer i $L^1(P)$, så at $(X, Y) \sim (Z, Y)$ og φ et element i $\mathrm{M}(\mathcal{E})$. Da er

$$E[X \mid Y] = \varphi(Y) \;\; P\text{-n.o.} \Leftrightarrow E[Z \mid Y] = \varphi(Y) \;\; P\text{-n.o.}.$$

Men hvordan bestemmer man et φ, der passer til et givent $X \in L^1(P)$? Problemet behandles i sektion 6.11, som blandt andet indeholder flg. resultat.

Bm 18 Lad (X, Y) betegne en absolut kontinuert 2-dimensional stokastisk vektor med tæthed $(x, y) \mapsto f(x, y)$ m.h.t. det plane Lebesgue mål. Definer

$$f_2(y) := \int_{\mathbf{R}} f(u, y) \, du \;\; \text{og} \;\; f_{X|Y}(x|y) := \frac{f(x, y)}{f_2(y)} \cdot \mathbf{1}_{\{f_2 > 0\}}(y) \;\; \text{for } x, y \in \mathbf{R}.$$

Da gælder for enhver begrænset Borel funktion $\psi : \mathbf{R}^2 \to \mathbf{R}$, at

$$E[\psi(X, Y) \mid Y] = \tilde{\psi}(Y) \; P\text{-n.o., hvor } \; \tilde{\psi}(y) := \int_{\mathbf{R}} \psi(x, y) \cdot f_{X|Y}(x|y) \, dx \;\; y \in \mathbf{R}.$$

Da f_2 er en tæthed for Y, er $P_Y(f_2 > 0) = 1$. Resultatet gælder uændret for ubegrænsede ψ, hvis blot $E[\,|\psi(X,Y)|] < \infty$, dog skal $\tilde{\psi}(y)$ sættes lig 0 på mængden

$$\{y \in \mathbf{R} \,|\, \int_{\mathbf{R}} |\psi(x,y)| \cdot f_{X|Y}(x|y)\,dx = \infty\}.$$

Denne er igen en P_Y-nulmængde, idet

$$\int_{\mathbf{R}} \{\int_{\mathbf{R}} |\psi(x,y)| \cdot f_{X|Y}(x|y)\,dx\}\,P_Y(dy) = \int_{\mathbf{R}} \{\int_{\mathbf{R}} |\psi(x,y)| \cdot f_{X|Y}(x|y)\,dx\}\,f_2(y)\,dy$$

$$= \int_{\mathbf{R}} \{\int_{\mathbf{R}} |\psi(x,y)| \cdot \mathbf{1}_{\{f_2>0\}}(y) \cdot f(x,y)\,dx\}\,dy = \int_{\mathbf{R}^2} |\psi(x,y)|\,P_{(X,Y)}(dx\,dy)$$

$$= E[\,|\psi(X,Y)|] < \infty.$$

En anden situation, hvor problemet umiddelbart lader sig løse, er flg. Beviset overlades til læseren.

Bm 19 Lad Y betegne en diskret stokastisk variabel, og lad $(y_n)_{n\geq 1}$ være en nummerering af den højst tællelige mængde $Sp(Y)$, d.v.s.

$$P(Y = y_n) > 0 \text{ for alle } n \text{ og } \sum_{n\geq 1} P(Y = y_n) = 1.$$

Da gælder for enhver stokastisk variabel X og Borel funktion $\psi : \mathbf{R}^2 \to \mathbf{R}$, at hvis $E[\,|\psi(X,Y)|] < \infty$, så er

$$E[\psi(X,Y)\,|\,Y] = \tilde{\psi}(Y) \;\; P\text{-n.o.},$$

hvor

$$\tilde{\psi}(\cdot) := \sum_{n\geq 1} a_n \cdot \mathbf{1}_{\{Y=y_n\}}(\cdot), \quad \text{for} \;\; a_n := \frac{1}{P(Y = y_n)} \int_{\{Y=y_n\}} \psi(X, y_n)\,dP \quad n \geq 1.$$

Bm 18 og 19 knytter tæt an til, hvad der normalt kaldes en regulær betinget fordeling af X givet Y. Det er ikke et emne, vi skal gøre meget ud af, men da det spiller en vis rolle i forbindelse med behandlingen af den flerdimensionale normalfordeling, vil jeg ganske kort indføre nogle vigtige begreber og definitioner.

Notation $\{P(A\,|\,\underline{y})\,|\,A \in \mathcal{B}(\mathbf{R}^n),\, \underline{y} \in \mathbf{R}^m\}$ kaldes en *Markov kerne* på $\mathcal{B}(\mathbf{R}^n) \times \mathbf{R}^m$, hvis

a) $A \mapsto P(A\,|\,\underline{y})$ er et Borel sandsynlighedsmål på $\mathbf{R}^n$ for alle $\underline{y} \in \mathbf{R}^m$.

b) $\underline{y} \mapsto P(A\,|\,\underline{y})$ er en Borel funktion på $\mathbf{R}^m$ for alle $A \in \mathcal{B}(\mathbf{R}^n)$.

Lad X og Y betegne h.h.v. en n og en m-dimensional stokastisk vektor. En Markov kerne $\{P_{X\,|\,Y}(A\,|\,\underline{y})\,|\,A \in \mathcal{B}(\mathbf{R}^n),\, \underline{y} \in \mathbf{R}^m\}$ på $\mathcal{B}(\mathbf{R}^n) \times \mathbf{R}^m$ kaldes en *regulær betinget fordeling* for X givet Y, hvis

$$P(X \in A,\, Y \in B) = \int_B P_{X\,|\,Y}(A\,|\,\underline{y})\,P_Y(d\underline{y})$$

for alle $A \in \mathcal{B}(\mathbf{R}^n)$ og $B \in \mathcal{B}(\mathbf{R}^m)$. Sandsynlighedsmålet

$$A \mapsto P_{X\,|\,Y}(A\,|\,\underline{y})$$

kaldes *den betingede fordeling* for X givet $Y = \underline{y}$ og er det absolut kontinuert med tæthed $\underline{x} \mapsto f_{X\,|\,Y}(\underline{x}\,|\,\underline{y})$, kaldes denne en *betinget tæthed* for X givet $Y = \underline{y}$.

Øvelse 20. Udnyt Proposition 2 og sektionsegenskaber ved produktmålelige mængder til at vise at for alle $A \in \mathcal{B}(\mathbf{R}^{n+m}) = \mathcal{B}(\mathbf{R}^n) \times \mathcal{B}(\mathbf{R}^m)$ er

$$P((X,Y) \in A) = \int_{\mathbf{R}^m} P_{X\,|\,Y}(A(\underline{y})\,|\,\underline{y})\, P_Y(d\underline{y}) \quad \square$$

Bemærkning. Øvelse 20 viser, at man ud fra kendskab til en betinget fordeling for X givet Y kan beregne en betinget fordeling for Z givet Y, hvor $Z := \phi(X,Y)$ for en Borel funktion ϕ. Øvelsen viser nemlig, at den betingede fordeling for Z givet $Y = \underline{y}$ er billedmålet af den betingede fordeling for X givet $Y = \underline{y}$ svarende til den målelige afbildning

$$\underline{x} \mapsto \phi(\underline{x}, \underline{y})$$

Betingede fordelinger er teoretisk set et vanskeligt begreb. Men for en- eller flerdimensionale stokastiske vektorer X og Y eksisterer der altid en regulær betinget fordeling for X givet Y, og da $\mathcal{B}(\mathbf{R}^n)$ er separabel, er $P_{X\,|\,Y}(\,\cdot\,|\,\cdot\,)$ entydigt bestemt i en sådan grad, at det har mening at tale om 'den betingede fordeling' for X givet Y. De betingede fordelinger for X givet $Y = \underline{y}$ er nemlig entydigt bestemte for P_Y-n.a. $\underline{y}$. Skønt der således både er eksistens og entydighed, er den eksplicitte beregning ofte vanskelig (se dog nedenstående øvelse), men i anvendelsessituationer er de betingede fordelinger ofte givet ud fra sammenhængen.

Kendskab til en betinget fordeling for X givet Y gør det muligt at generalisere punkterne Bm 14, 18 og 19. For er f en begrænset Borel funktion på $\mathbf{R}^n \times \mathbf{R}^m$, så er

$$E[f(X,Y)\,|\,Y] = \tilde{f}(Y) \;\; P\text{-n.o. hvor } \tilde{f}:\; \underline{y} \mapsto \int_{\mathbf{R}^n} f(\underline{x},\underline{y})\, P_{X\,|\,Y}(d\underline{x}\,|\,\underline{y}).$$

D.v.s. $\tilde{f}$'s værdi i et punkt $\underline{y}$ er middelværdien i den betingede fordeling af $f(X,Y)$ givet $Y = \underline{y}$.

Formlen vises ved først at reducere til produktfunktioner, d.v.s. funktioner

$$(\underline{x},\underline{y}) \mapsto f_1(\underline{x}) \cdot f_2(\underline{y}),$$

hvor f_1 og f_2 er Borel funktioner på $\mathbf{R}^n$ h.h.v. $\mathbf{R}^m$. Dernæst indses ved hjælp af Standardbeviset, at det er nok at se på tilfældet

$$f_1 = \mathbf{1}_A \text{ og } f_2 = \mathbf{1}_B \text{ for } A \in \mathcal{B}(\mathbf{R}^n) \text{ og } B \in \mathcal{B}(\mathbf{R}^m),$$

men dette svarer præcis til ovenstående definitionsligning. Resultatet kan udvides til visse ubegrænsede f.

Øvelse 21. Lad X og Y være givne n og m dimensionale stokastiske vektorer og μ et vilkårligt givet Borel sandsynlighedsmål på $\mathbf{R}^n$. $\square$

I) Hvis (X,Y) er absolut kontinuert m.h.t. λ_{n+m} med tæthed $(\underline{x},\underline{y}) \mapsto f(\underline{x},\underline{y})$, er

$$P_{X|Y}(A\,|\,\underline{y}) := \begin{cases} \int_A f_{X|Y}(\underline{x}|\underline{y})\,\lambda_n(d\underline{x}) & A \in \mathcal{B}(\mathbf{R}^n),\ \underline{y} \in \{f_2 > 0\} \\ \mu(A) & A \in \mathcal{B}(\mathbf{R}^n),\ \underline{y} \notin \{f_2 > 0\} \end{cases}$$

en regulær betinget fordeling for X givet Y. $\underline{y} \mapsto f_2(\underline{y})$ er her en tæthed for Y og

$$f_{X|Y}(\underline{x}|\underline{y}) := f(\underline{x},\underline{y})/f_2(\underline{y}) \cdot \mathbf{1}_{\{f_2>0\}}(\underline{y}) \quad \text{for } \underline{x} \in \mathbf{R}^n,\ \underline{y} \in \mathbf{R}^m.$$

II) Hvis Y er diskret og $Sp(Y) := \{\underline{y} \in \mathbf{R}^m \,|\, P(Y = \underline{y}) > 0\}$, er

$$P_{X|Y}(A\,|\,\underline{y}) := \begin{cases} P(X \in A\,|\,Y = \underline{y}) & A \in \mathcal{B}(\mathbf{R}^n),\ \underline{y} \in Sp(Y) \\ \mu(A) & A \in \mathcal{B}(\mathbf{R}^n),\ \underline{y} \notin Sp(Y) \end{cases}$$

en regulær betinget fordeling for X givet Y.

III) Hvis X og Y er uafhængige, er

$$P_{X|Y}(A\,|\,\underline{y}) := P_X(A) \quad \text{for } A \in \mathcal{B}(\mathbf{R}^n),\ \ \underline{y} \in \mathbf{R}^m$$

en regulær betinget fordeling for X givet Y. $\square$

Punkt III) viser, at hvis X og Y er uafhængige, så afhænger de betingede mål for X givet $Y = y$ ikke af y. Dette ses let at karakterisere uafhængighed, idet flg. udsagn er ækvivalente. μ er her et sandsynlighedsmål på $\mathbf{R}^n$.

1) X og Y er uafhængige

2) $P_{X|Y}(A\,|\,\underline{y}) := \mu(A)$ for $A \in \mathcal{B}(\mathbf{R}^n)$ og $\underline{y} \in \mathbf{R}^m$ er en regulær betinget fordeling for X givet Y.

Øvelse 22. Eftervis ækvivalensen og vis endvidere, at μ i givet fald er fordelingsmålet for X. $\square$

Øvelse 23. Lad (X,Y) være todimensionalt normalt fordelt. Vis at den betingede fordeling for X givet $Y = y$ er en normalfordeling og bestem dens parametre. Vink: Vis at der findes et α, så at $X - \alpha Y$ og Y er uafhængige, og udnyt dernæst at $X = (X - \alpha Y) + \alpha Y$ tillige med bemærkningen efter Øvelse 20. $\square$

Martingaler.
I dette kapitel betragtes modeller for systemer, der udvikler sig i tiden. Tiden modelleres diskret, d.v.s. ved en tidsparametermængde $T \subseteq \mathbf{Z}$, normalt et interval. Til ethvert tidspunkt n i T knytter der sig en variabel X_n og en del σ-algebra $\mathcal{F}_n$ i $\mathcal{F}$. Vi skal tænke på $\mathcal{F}_n$ som den informationsmængde, der er til stede til tid n, og på X_n som en variabel, der beskriver tilstanden til tid n. Flg. krav forekommer derfor naturlige.

a) $\mathcal{F}_n \subseteq \mathcal{F}_m$ hvis $n \leq m$ for tidspunkter n og m i T, d.v.s. informationsmængden vokser med tiden.

b) X_n er målelig m.h.t. $\mathcal{F}_n$, d.v.s. tilstanden til tid n kan observeres på baggrund af den information, der er til stede til tid n.

Med udgangspunkt heri siges en parametriseret familie $(\mathcal{F}_n)_{n\in T}$ af del σ-algebraer i $\mathcal{F}$ at udgøre et *T-filter*, hvis $\mathcal{F}_n \subseteq \mathcal{F}_m$ for $n \leq m$, $n, m \in T$; og b) udtrykkes ofte kort ved at sige, at *processen* $(X_n)_{n\in T}$ er *tilpasset* filtret $(\mathcal{F}_n)_{n\in T}$.

Vi vil kun se på tilfældet, hvor $T = \mathbf{N}_0 := \{0, 1, \ldots\}$, thi herved dækkes også tilfældet $(X_n, \mathcal{F}_n)_{n\geq k}$, idet denne kan opfattes som $(\tilde{X}_n, \tilde{\mathcal{F}}_n)_{n\geq 0}$, hvor

$$\tilde{X}_n = X_{n+k} \ \text{ og } \ \tilde{\mathcal{F}}_n = \mathcal{F}_{n+k} \quad n \geq 0.$$

Tilfældet, hvor T er et endeligt interval $[k, l]$, er også dækket, thi forlænges konstant ud over højre endepunkt kan $(X_n, \mathcal{F}_n)_{n\in[k,l]}$ beskrives ved $(\tilde{X}_n, \tilde{\mathcal{F}}_n)_{n\geq 0}$, hvor

$$\tilde{X}_n = X_{n+k}, \quad \tilde{\mathcal{F}}_n = \mathcal{F}_{n+k} \quad n \leq l-k \text{ og } \tilde{X}_n = X_l, \quad \tilde{\mathcal{F}}_n = \mathcal{F}_l \quad n > l-k.$$

$T = \mathbf{N}_0$ omfatter altså alle situationer, hvor tidsmængden har et endeligt begyndelsespunkt, og der udestår derfor i princippet kun to tilfælde nemlig, $T = \mathbf{Z}$ eller $T = -\mathbf{N}_0$. Her er $T = \mathbf{Z}$ ikke interessant i en martingal sammenhæng og $T = -\mathbf{N}_0$ overlades på grund af manglende tid til et senere kursus.

Til ethvert filter $(\mathcal{F}_n)_{n\geq 0}$ tilknyttes de såkaldte stoptider defineret på flg. vis.

Definition $\tau : \Omega \to \mathbf{N}_0 \cup \{\infty\}$ er en *stoptid* (mere præcist en $(\mathcal{F}_n)_{n\geq 0}$-stoptid), hvis

$$\{\tau > n\} \in \mathcal{F}_n \qquad n \geq 0.$$

Bemærk at målelighedskravet i definitionen ækvivalent kan formuleres som

$$\{\tau \leq n\} \in \mathcal{F}_n \quad n \geq 0 \ \text{ eller } \ \{\tau = n\} \in \mathcal{F}_n \quad n \geq 0.$$

τ siges at være en *endelig* stoptid, hvis $P(\tau < \infty) = 1$, og τ siges at være en *begrænset* stoptid, hvis $\tau \leq M < \infty$, hvor M er et reelt tal.
Til enhver stoptid τ tilordnes σ-algebraen (overvej)

$$\mathcal{F}_\tau := \{F \in \mathcal{F}_\infty \,|\, F \cap \{\tau = n\} \in \mathcal{F}_n \ n \geq 0\,\},$$

hvor $\mathcal{F}_\infty := \sigma(\bigcup_n \mathcal{F}_n)$, d.v.s. den mindste σ-algebra, der indeholder ethvert $\mathcal{F}_n$. $\mathcal{F}_\tau$ er altså en del σ-algebra i $\mathcal{F}_\infty$ og omtales som informationsmængden, der er til stede til tid τ.

Inden vi starter på den egentlige teori gennemgås en række åbenbare konsekvenser af defintionen på en stoptid τ og den tilhørende σ-algebra $\mathcal{F}_\tau$.

Ma 1 τ er $\mathcal{F}_\tau$-målelig, og en $\mathcal{F}_\infty$-målelig stokastisk variabel X er $\mathcal{F}_\tau$-målelig, hvis og kun hvis $X \cdot \mathbf{1}_{\{\tau=n\}}$ er $\mathcal{F}_n$-målelig for alle $n \geq 0$.

Bevis. Den første påstand følger af identiten

$$\{\tau = k\} \cap \{\tau = n\} = \begin{cases} \{\tau = n\} & \text{hvis } n = k \\ \emptyset & \text{hvis } n \neq k. \end{cases}$$

for $n,\, k \in \mathbf{N}_0 \cup \{\infty\}$, og den anden af identiteten

$$\{X \in B\} \cap \{\tau = n\} = \{X \cdot \mathbf{1}_{\{\tau=n\}} \in B\} \cap \{\tau = n\}.$$

for $n \geq 0$ og $B \in \mathcal{B}(\mathbf{R})$, idet denne specielt, hvis $0 \notin B$, giver

$$\{X \in B\} \cap \{\tau = n\} = \{X \cdot \mathbf{1}_{\{\tau=n\}} \in B\}. \quad \diamond$$

De vigtigste stoptider er de såkaldte *First Hitting Times* defineret ved (her og overalt i det følgende defineres $\inf \emptyset := \infty$)

$$\tau_A(\omega) := \inf\{n \geq 0 \,|\, X_n(\omega) \in A\} \qquad \omega \in \Omega,$$

hvor $A \in \mathcal{B}(\mathbf{R})$ og $(X_n)_{n\geq 0}$ er en tilpasset reel proces. Stoptidsegenskaben følger af ligheden

$$\{\tau_A > n\} = \bigcap_{k=0}^{n} \{X_k \in A\}^c = \bigcap_{k=0}^{n} \{X_k \in A^c\} \qquad n \geq 0.$$

Som vist i sektion 7.2 kan dette generaliseres til de såkaldte *Ocurrence Time* $\tau_{\underline{F}}$, hvor $\underline{F} := (F_n)_{n\geq 0}$ er en følge af hændelser, så at $F_n \in \mathcal{F}_n$ for alle $n \geq 0$, defineret ved

$$\tau_{\underline{F}}(\omega) := \inf\{n \geq 0 \,|\, \omega \in F_n\} \qquad \omega \in \Omega,$$

og igen følger stoptidsegenskaben let, idet $\{\tau_{\underline{F}} > n\} = \bigcap_{k=0}^{n} F_k^c$ for $n \geq 0$.

Udnyttes at stoptider kun antager heltallige værdier ses, at enhver stoptid τ kan beskrives som en Ocurrence Time, idet

$$\tau(\omega) = \inf\{n \geq 0 \,|\, \omega \in \{\tau \leq n\}\} \quad \omega \in \Omega.$$

Direkte anvendelse af definitionen viser, at for ethvert $k \in \mathbf{N}_0 \cup \{\infty\}$ er den konstante variabel $\tau(\omega) :\equiv k$ en stoptid, samt at mængden af $(\mathcal{F}_n)_{n\geq 0}$-stoptider er stabil under endelig sum og endelig punktvis max og min dannelse, d.v.s.

Ma 2 τ_1, τ_2 stoptider $\Rightarrow \tau_1 + \tau_2$, $\tau_1 \vee \tau_2$ og $\tau_1 \wedge \tau_2$ stoptider.
Bevis. Følger af lighederne

$$\{\tau_1 \vee \tau_2 \leq n\} = \{\tau_1 \leq n\} \cap \{\tau_2 \leq n\} \quad \text{og} \quad \{\tau_1 \wedge \tau_2 > n\} = \{\tau_1 > n\} \cap \{\tau_2 > n\}$$

samt

$$\{\tau_1 + \tau_2 = n\} = \bigcup_{k=1}^{n} \{\tau_1 = k\} \cap \{\tau_2 = n - k\} \quad \diamond$$

Argumenterne for $\vee$ og $\wedge$ udvider uden videre til tællig mange stoptider, d.v.s.

$$(\tau_i)_{i \geq 1} \text{ stoptider} \Rightarrow \sup_i \tau_i \text{ og } \inf_i \tau_i \text{ stoptider,}$$

og derfor tilsvarende for summen, da $\sum_{i=1}^{\infty} \tau_i = \sup_n \sum_{i=1}^{n} \tau_i$.
Det næste resultat viser, at stoptids σ-algebraerne generaliserer de givne informations σ-algebraer $\mathcal{F}_n$.

Ma 3 $\mathcal{F}_\tau = \mathcal{F}_k$, hvis $\tau \equiv k$ for et $k \in \mathbf{N}_0 \cup \{\infty\}$ og for stoptider τ_1 og τ_2 er

$$\{\tau_1 \leq \tau_2\} \in \mathcal{F}_{\tau_1 \wedge \tau_2} = \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}, \quad \text{d.v.s. specielt} \quad \tau_1 \leq \tau_2 \Rightarrow \mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}.$$

Bevis. Den første påstand overlades til læseren. For $B \in \mathcal{F}_{\tau_1 \wedge \tau_2}$ og $n \geq 0$ er

$$\mathbf{1}_B \cdot \mathbf{1}_{\{\tau_1 = n\}} = \sum_{k=0}^{n} \mathbf{1}_B \cdot \mathbf{1}_{\{\tau_1 \wedge \tau_2 = k\}} \cdot \mathbf{1}_{\{\tau_1 = n\}}$$

$\mathcal{F}_n$-målelig, hvilket ifølge Ma 1 viser, at $B \in \mathcal{F}_{\tau_1}$ og tilsvarende $B \in \mathcal{F}_{\tau_2}$. Hvis omvendt $B \in \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$ viser lighederne for $n \geq 0$

$$\mathbf{1}_B \cdot \mathbf{1}_{\{\tau_1 \wedge \tau_2 = n\}} = \mathbf{1}_B \cdot \mathbf{1}_{\{\tau_1 = n\}} \cdot \mathbf{1}_{\{\tau_2 > n\}} + \mathbf{1}_B \cdot \mathbf{1}_{\{\tau_2 = n\}} \cdot \mathbf{1}_{\{\tau_1 > n\}} + \mathbf{1}_B \cdot \mathbf{1}_{\{\tau_1 = n\}} \cdot \mathbf{1}_{\{\tau_2 = n\}}$$

at $B \in \mathcal{F}_{\tau_1 \wedge \tau_2}$. D.v.s. $\mathcal{F}_{\tau_1 \wedge \tau_2} = \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$. Resten følger tilsvarende af lighederne

$$\mathbf{1}_{\{\tau_1 \leq \tau_2\}} \cdot \mathbf{1}_{\{\tau_1 \wedge \tau_2 = n\}} = \mathbf{1}_{\{\tau_1 \leq \tau_2\}} \cdot \mathbf{1}_{\{\tau_1 = n\}} = \mathbf{1}_{\{n \leq \tau_2\}} \cdot \mathbf{1}_{\{\tau_1 = n\}} = \mathbf{1}_{\{n-1 < \tau_2\}} \cdot \mathbf{1}_{\{\tau_1 = n\}}. \quad \diamond$$

Definition og målelighed m.h.t. $\mathcal{F}_\tau$ af variable af formen X_τ, *tilstanden til tid* τ, for en stoptid τ og en tilpasset proces $(X_n)_{n \geq 0}$ er klar, hvis τ kun antager endelige værdier, idet X_τ da naturligt defineres som

$$\omega \mapsto X_{\tau(\omega)}(\omega) \quad \omega \in \Omega,$$

eller ækvivalent

$$X_\tau = \sum_{n=0}^{\infty} X_n \cdot \mathbf{1}_{\{\tau = n\}}.$$

Der gælder altså at $X_\tau = X_n$ på $\{\tau = n\}$, d.v.s. X_τ er $\mathcal{F}_\tau$-målelig ifølge Ma 1.
Men generelt kan stoptider antage værdien ∞, og vi vil derfor til en given proces $(X_n)_{n \geq 0}$ altid tilordne en variabel X_∞ ifølge opskriften i Lemma 5, d.v.s.

$$X_\infty(\omega) := \lim_n X_n(\omega), \text{ hvis denne eksisterer i } \mathbf{R}, \text{ og } X_\infty(\omega) := 0 \text{ ellers.}$$

Ifølge Lemma 5 er X_∞ målelig m.h.t. $\mathcal{F}_\infty$, hvis $(X_n)_{n\geq 0}$ er tilpasset. Herefter kan vi uden problemer definere X_τ for en vilkårlig stoptid τ ved fastsættelsen

$$X_\tau := \sum_{n=0}^{\infty} X_n \cdot \mathbf{1}_{\{\tau=n\}} + X_\infty \cdot \mathbf{1}_{\{\tau=\infty\}}.$$

Der er tydeligvis tale om en udvidelse af den allerede indførte definition for stoptider med udelukkende endelige værdier. Endvidere gælder flg. alternative beskrivelse. (Sammenlign med sektion 7.4. Bemærk at $\overline{\mathbf{R}}$ er erstattet med $\mathbf{R}$.)

$$X_\tau(\omega) := \begin{cases} \lim_n X_{\tau(\omega)\wedge n}(\omega) & \text{hvis denne eksisterer i } \mathbf{R} \\ 0 & \text{ellers,} \end{cases}$$

d.v.s. X_τ er '∞-variablen' hørende til $(X_{\tau\wedge n})_{n\geq 0}$ defineret i h.h.t. Lemma 5.
Ma 1 og Lemma 5 sikrer derfor, at

Ma 4 X_τ er $\mathcal{F}_\tau$-målelig for enhver tilpasset proces $(X_n)_{n\geq 0}$.

Flg. variant af formel (7.4.6) er vigtig.

Ma 5 For $X \in L^1(P)$ og alle stoptider τ gælder for ethvert $k \in \mathbf{N}_0 \cup \{\infty\}$

$$E[X \mid \mathcal{F}_\tau] \cdot \mathbf{1}_{\{\tau=k\}} = E[X \mid \mathcal{F}_k] \cdot \mathbf{1}_{\{\tau=k\}} \quad P\text{-n.o.}$$

Bevis. Lad k være givet. Da begge sider er integrable og $\mathcal{F}_k$-målelige, er det nok at vise, at de har samme integral over ethvert $B \in \mathcal{F}_k$. Men dette følger af egenskaber ved betingede middelværdier, da $B \cap \{\tau = k\} \in \mathcal{F}_\tau \cap \mathcal{F}_k$ for alle $B \in \mathcal{F}_k$ (overvej) . $\diamond$

7.3 og 7.4 indeholder ikke flg. to vigtige resultater.

Ma 6 For enhver stoptid τ er $\mathcal{F}_\tau = \sigma(\bigcup_n \mathcal{F}_{\tau\wedge n})$.

Bevis. Inklusionen $\supseteq$ er åbenbar, da $\mathcal{F}_{\tau\wedge n} \subseteq \mathcal{F}_\tau$ for alle $n \geq 0$, og ved brug af identiteten $B = \bigcup_{n=0}^{\infty} B \cap \{\tau = n\} \cup B \cap \{\tau = \infty\}$ fås den anden, idet $B \cap \{\tau = n\} \in \mathcal{F}_\tau \cap \mathcal{F}_n = \mathcal{F}_{\tau\wedge n}$ for alle n og alle $B \in \mathcal{F}_\tau$ og

$$\mathcal{F}_\infty = \mathcal{B} := \{B \in \mathcal{F}_\infty \mid B \cap \{\tau = \infty\} \in \sigma(\bigcup_n \mathcal{F}_{\tau\wedge n})\}.$$

$\mathcal{B}$ er nemlig en σ-algebra, som indeholder ethvert $\mathcal{F}_n$, thi for $n \geq 0$ og $B \in \mathcal{F}_n$ er $B \cap \{\tau = \infty\} = \bigcap_{k=n}^{\infty} B \cap \{\tau \geq k\}$ og $B \cap \{\tau \geq k\} \in \mathcal{F}_{\tau\wedge k}$ for alle $k \geq n$. $\diamond$

Ma 7 Lad $(X_n)_{n\geq 0}$ betegne en tilpasset proces og τ en stoptid, så at X_τ er P-integrabel, d.v.s. element i $L^1(P)$. For enhver stoptid σ gælder da

$$E[X_\tau \mid \mathcal{F}_\sigma] = E[X_\tau \mid \mathcal{F}_{\tau\wedge\sigma}] \quad P\text{-n.o.}$$

Bevis. Da $\mathcal{F}_{\tau\wedge\sigma} \subseteq \mathcal{F}_\sigma$ udestår blot at vise, at for et givent $A \in \mathcal{F}_\sigma$ er

$$\int_A E[X_\tau \,|\, \mathcal{F}_{\tau\wedge\sigma}]\, dP = \int_A X_\tau \, dP.$$

Men da $A \cap \{\sigma \leq \tau\} \cap \{\tau \wedge \sigma = n\} = A \cap \{\sigma = n\} \cap \{n \leq \tau\} \in \mathcal{F}_n$ for alle $n \geq 0$, er $A \cap \{\sigma \leq \tau\} \in \mathcal{F}_{\tau\wedge\sigma}$, hvilket sikrer, at integralerne over $A \cap \{\sigma \leq \tau\}$ er ens. Beviset afsluttes ved at observere at $E[X_\tau \,|\, \mathcal{F}_{\tau\wedge\sigma}] \cdot \mathbf{1}_{\{\tau<\sigma\}} = X_\tau \cdot \mathbf{1}_{\{\tau<\sigma\}}$ P-n.o., thi ifølge Ma 5 gælder ligheden

$$\sum_{k=0}^{\infty} E[X_\tau \,|\, \mathcal{F}_k] \cdot \mathbf{1}_{\{\tau\wedge\sigma=k\}} \cdot \mathbf{1}_{\{\tau<\sigma\}} = \sum_{k=0}^{\infty} E[X_\tau \,|\, \mathcal{F}_\tau] \cdot \mathbf{1}_{\{\tau=k\}} \cdot \mathbf{1}_{\{\tau<\sigma\}}. \qquad \diamond$$

Vi skal studere de såkaldte *martingal* processer, dels fordi de dukker op i utrolig mange sammenhænge, og dels fordi man for disse kan vise interessante *konvergensresultater* samt *maksimaluligheder*. Martingalerne defineres i h.h.t. flg. definition, hvor $(\mathcal{F}_n)_{n\geq 0}$ er et filter og $(X_n)_{n\geq 0}$ en følge af reelle stokastiske variable.

Definition $(X_n, \mathcal{F}_n)_{n\geq 0}$ kaldes en *martingal*, hvis $(X_n)_{n\geq 0}$ er tilpasset til $(\mathcal{F}_n)_{n\geq 0}$ og P-integrabel, d.v.s. $X_n \in L^1(P, \mathcal{F}_n)$ for alle $n \geq 0$, samt

$$E[X_{n+1} \,|\, \mathcal{F}_n] = X_n \quad P\text{-n.o.} \quad \text{for alle } n \geq 0.$$

Med notationen $\Delta X_n := X_n - X_{n-1}$ for $n \geq 1$ kan martingalbetingelsen ækvivalent formuleres som

$$E[\Delta X_n \,|\, \mathcal{F}_{n-1}] = 0 \;\; P\text{-n.o.} \quad \text{for alle } n \geq 1.$$

Hvis $=$ erstattes med et $\geq$, taler man om en *submartingal*, og hvis der i stedet gælder $\leq$, taler man om en *supermartingal*. D.v.s. $(X_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal, hvis og kun hvis $(X_n, \mathcal{F}_n)_{n\geq 0}$ både er en sub- og en supermartingal, og $(X_n, \mathcal{F}_n)_{n\geq 0}$ er en supermartingal, hvis og kun hvis $(-X_n, \mathcal{F}_n)_{n\geq 0}$ er en submartingal. Mængden af martingaler m.h.t. $(\mathcal{F}_n)_{n\geq 0}$ udgør derfor et vektorrum, hvorimod mængden af sub- h.h.v. supermartingaler kun er lukket under linearkombinationer med positive konstanter, d.v.s. udgør en såkaldt positiv kegle. D.v.s.

$$(X_n, \mathcal{F}_n)_{n\geq 0} \text{ og } (Y_n, \mathcal{F}_n)_{n\geq 0} \text{ martingaler} \Rightarrow (aX_n + bY_n, \mathcal{F}_n)_{n\geq 0} \text{ martingal}$$

for alle $a, b \in \mathbf{R}$, og hvis udgangspunktet er to submartingaler, så fås igen en submartingal hvis $a, b \in \mathbf{R}_+$. (tilsvarende for supermartingaler)
Endvidere viser Jensen's ulighed for betingede middelværdier flg. resultat. Bemærk at en reel konveks funktion er kontinuert og derfor specielt Borel målelig.

Ma 8 For enhver en reel konveks funktion $\varphi : \mathbf{R} \to \mathbf{R}$ gælder

$$(X_n, \mathcal{F}_n)_{n\geq 0} \text{ martingal og } \varphi(X_n) \in L^1(P)\; n \geq 0 \Rightarrow (\varphi(X_n), \mathcal{F}_n)_{n\geq 0} \text{ submartingal.}$$

Beviset, der beror på at

$$E[\,\varphi(X_{n+1})\,|\,\mathcal{F}_n] \geq \varphi(\,E[X_{n+1}\,|\,\mathcal{F}_n]\,) = \varphi(X_n),$$

afslører, at hvis φ er konveks og voksende, så er konklusionen den samme for enhver submartingal $(X_n, \mathcal{F}_n)_{n\geq 0}$. Læseren opfordres til at formulere tilsvarende udsagn for supermartingaler og konkave funktioner.

Som det fremgår af ovenstående definition, betragtes her kun sub- og supermartingaler bestående af integrable variable, d.v.s. elementer i $L^1(P)$, hvorimod bogen kun kræver, at de ligger i $L(P)$. Man kan derfor her tale om den tilhørende *middelværdifunktion* $n \mapsto E[X_n]$, og det ses let, at denne er voksende for submartingaler, aftagende for supermartingaler og konstant for martingaler. Ligheden $|x| = 2x^+ - x$ giver derfor, at

$$E[\,|X_n|] = 2\cdot E[X_n^+] - E[X_n] \leq 2\cdot E[X_n^+] - E[X_0] \quad n \geq 0,$$

for enhver submartingal $(X_n, \mathcal{F}_n)_{n\geq 0}$, d.v.s.

$$\sup_n E[\,|X_n|] < \infty \Leftrightarrow \sup_n E[X_n^+] < \infty,$$

og dermed tilsvarende for enhver supermartingal $(X_n, \mathcal{F}_n)_{n\geq 0}$

$$\sup_n E[\,|X_n|] < \infty \Leftrightarrow \sup_n E[X_n^-] < \infty.$$

Sektion 7.6 giver eksempler på martingaler, med særlig fokus på de martingaler der fremkommer ud fra uafhængige variable. Følgende er specielt vigtige.

Ma 9 Lad $(X_n)_{n\geq 0}$ betegne en følge af uafhængige integrable stokastiske variable, og lad $(\mathcal{F}_n)_{n\geq 0}$ betegne filtret frembragt af $(X_n)_{n\geq 0}$, d.v.s.

$$\mathcal{F}_n := \sigma(X_0, \ldots, X_n) \quad n \geq 0.$$

Da gælder idet

$$S_n := \sum_{j=0}^{n} X_j \quad \text{og} \quad P_n := \prod_{j=0}^{n} X_j \quad n \geq 0$$

1) $(S_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal, hvis $E[X_n] = 0$ for alle n, og en submartingal, hvis $E[X_n] \geq 0$ for alle n.

Hvis yderligere X_n'erne alle har middelværdi 0 og endelig varians, er processen $(S_n^2 - \sum_{j=0}^n Var(X_j), \mathcal{F}_n)_{n\geq 0}$ også en martingal.

2) $(P_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal, hvis $E[X_n] = 1$ for alle n. Hvis X_n'erne er ikke negative, er $(P_n, \mathcal{F}_n)_{n\geq 0}$ en submartingal, hvis $E[X_n] \geq 1$ for alle n.

En anden vigtig martingal type er de såkaldte Lévy martingaler, d.v.s. processer på formen

$$(E[X\,|\,\mathcal{G}_n], \mathcal{G}_n)_{n\geq 0},$$

hvor $(\mathcal{G}_n)_{n\geq 0}$ er et filter og X et vilkårligt element i $L^1(P)$. Det overlades til læseren at eftervise, at der vitterligt er tale om martingaler. Ifølge Bm 16' er variablene i en Lévy martingal uniformt integrable, og vi skal senere se, at enhver sådan unifomt integrabel martingal er en Lévy martingal.
Lad mig også nævne den såkaldte *Doob dekomposition.* $(X_n, \mathcal{F}_n)_{n\geq 0}$ er her en tilpasset integrabel proces, d.v.s. $X_n \in L^1(P, \mathcal{F}_n)$ for alle $n \geq 0$. For $n \geq 1$ er

$$X_n = X_0 + \sum_{k=1}^{n} \Delta X_k = X_0 + \sum_{k=1}^{n} E[\Delta X_k \,|\, \mathcal{F}_{k-1}] + \sum_{k=1}^{n} (\Delta X_k - E[\Delta X_k \,|\, \mathcal{F}_{k-1}]),$$

d.v.s. dekompositionen $X_n = X_0 + A_n + M_n$ for $n \geq 0$, hvor $A_0 = M_0 \equiv 0$ og

$$A_n = \sum_{k=1}^{n} E[\Delta X_k \,|\, \mathcal{F}_{k-1}] \quad \text{og} \quad M_n = \sum_{k=1}^{n} (\Delta X_k - E[\Delta X_k \,|\, \mathcal{F}_{k-1}])$$

for $n \geq 1$. Processerne $(A_n, \mathcal{F}_n)_{n\geq 0}$ og $(M_n, \mathcal{F}_n)_{n\geq 0}$ er begge tilpassede og integrable, og ved nærmere eftersyn ses, at $(M_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal og $(A_n)_{n\geq 0}$ en såkaldt $(\mathcal{F}_n)-$*predictabel* proces, d.v.s.

$$A_n \text{ er } \mathcal{F}_{n-1}\text{-målelig for alle } n \geq 0 \ (\mathcal{F}_{-1} := \mathcal{F}_0).$$

Bemærk at hvis $(X_n, \mathcal{F}_n)_{n\geq 0}$ er en submartingal og derfor $E[\Delta X_n \,|\, \mathcal{F}_{n-1}] \geq 0$ for alle $n \geq 1$, er $(A_n)_{n\geq 0}$ en *voksende proces*, d.v.s.

$$0 \leq A_n \leq A_{n+1} \quad P\text{-n.o. for } n \geq 0.$$

Overvejelserne kan sammenfattes i flg. udsagn:

Doob's Dekompositionssætning.
Enhver $(\mathcal{F}_n)_{n\geq 0}$-tilpasset integrabel proces $(X_n)_{n\geq 0}$ kan skrives på formen

$$X_n = M_n + A_n, \quad n \geq 1 \quad X_0 = M_0,$$

hvor $(M_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal, og $(A_n)_{n\geq 0}$ er en $(\mathcal{F}_n)$-predictabel proces; og opskrivningen er entydig op til P-nulmængder. D.v.s. hvis

$$X_n = \tilde{M}_n + \tilde{A}_n, \quad n \geq 1 \quad X_0 = \tilde{M}_0,$$

er en anden repræsentation af samme type, så er P-n.o. for alle $n \geq 0$

$$M_n = \tilde{M}_n \ \text{ og } \ A_n = \tilde{A}_n.$$

$(X_n, \mathcal{F}_n)_{n\geq 0}$ er derfor en sub - h.h.v. en supermartingal, hvis og kun hvis $(X_n)_{n\geq 0}$ er $(\mathcal{F}_n)_{n\geq 0}$-tilpasset og på formen

$$(M_n + A_n)_{n\geq 0} \ \text{ h.h.v. } \ (M_n - A_n)_{n\geq 0}$$

hvor $(M_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal og $(A_n)_{n\geq 0}$ en voksende integrabel proces.

Hvis $(X_n, \mathcal{F}_n)_{n\geq 0}$ er en martingal, er

$$E[\Delta X_k \,|\, \mathcal{F}_{k-1}] = 0 \;\text{ og derfor også } E[V_k \cdot \Delta X_k \,|\, \mathcal{F}_{k-1}] = 0 \quad k \geq 1$$

P-n.o., hvis V_k er begrænset og $\mathcal{F}_{k-1}$-målelig for $k \geq 1$. Dette giver umiddelbart anledning til flg. resultat.

Martingal Transforms.
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en martingal og $(V_n)_{n\geq 0}$ en $(\mathcal{F}_n)$-predictabel proces, hvor V_n er begrænset for alle $n \geq 0$. Da er $(V \bullet X_n, \mathcal{F}_n)_{n\geq 0}$ en martingal, hvor

$$V \bullet X_n := V_0 \cdot X_0 + \sum_{k=1}^{n} V_k \cdot \Delta X_k \quad n \geq 0.$$

Hvis $(X_n, \mathcal{F}_n)_{n\geq 0}$ en submartingal, er $(V \bullet X_n, \mathcal{F}_n)_{n\geq 0}$ igen en submartingal, hvis V_n'erne yderligere er ikke-negative. Processer af den her definerede art er i litteraturen kendt under navnet *martingal transforms*.

Martingalerne har som allerede nævnt mange interessante egenskaber, og vi skal i det følgende gennemgå nedenstående fundamentale resultater. Listen omfatter ikke alle de resultater, der er nævnt i Hoffmann's bog, men dog de vigtigste. Ligeledes afviger de her givne beviser ofte fra bogens.

Sætning Ma 1 *Optional Sampling.* (skrabet udgave)
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en martingal. Den standsede *proces $(X_{\tau\wedge n}, \mathcal{F}_n)_{n\geq 0}$ er da en martingal for enhver stoptid τ, og for ethvert par af begrænsede stoptider $0 \leq \sigma \leq \tau$ er X_τ og X_σ elementer i $L^1(P)$ og*

$$E[X_0] = E[X_\sigma] = E[X_\tau] \;\text{ samt }\; E[X_\tau \,|\, \mathcal{F}_\sigma] = X_\sigma \quad P\text{-n.o.}$$

Der gælder helt tilsvarende resultater for sub- og supermartingaler med = erstattet af det relevante ulighedstegn.

Sætning Ma 2 *Maksimaluligheder.*
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en submartingal og lad $\lambda > 0$ være givet. Da gælder for alle $n \geq 1$

$$\lambda \cdot P(\min_{0\leq k\leq n} X_k < -\lambda) \leq E[X_n^+] - E[X_0]$$

$$\lambda \cdot P(\max_{0\leq k\leq n} X_k > \lambda) \leq E[X_n,\, \max_{0\leq k\leq n} X_k > \lambda] \leq E[X_n^+],$$

hvilket ved addition viser, at for alle $n \geq 1$ og $\lambda > 0$ er

$$\lambda \cdot P(\max_{0\leq k\leq n} |X_k| > \lambda) \leq 3 \cdot \max_{0\leq k\leq n} E[\,|X_k|].$$

Den sidste ulighed gælder også for supermartingaler. Lader vi $n \to \infty$ fås derfor

$$\lambda \cdot P(\sup_k |X_k| > \lambda) \leq 3 \cdot \sup_k E[\,|X_k|] \;\text{ for } \lambda > 0,$$

d.v.s. for enhver sub - eller supermartingal $(X_n, \mathcal{F}_n)_{n\geq 0}$ er $P(\sup_k |X_k| < \infty) = 1$, hvis $\sup_k E[\,|X_k|] < \infty$.

Sætning Ma 3 *Opkrydsningsuligheder.*
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en supermartingal. Da gælder for alle $n \geq 1$ og alle reelle tal $r < s$

$$(s-r) \cdot E[U^n_{r,s}] \leq E[(X_n - r)^-] \leq E[X_n^-] + |r|,$$

hvor $U^n_{r,s}$ er antallet af opkrydsninger over $[r, s]$ i tidsintervallet $[0, n]$. Ved brug af Monoton konvergens fås derfor, at

$$(s-r) \cdot E[\sup_n U^n_{r,s}] \leq \sup_n E[X_n^-] + |r|.$$

Det totale antal opkrydsninger over $[r, s]$, d.v.s. variablen

$$U_{r,s} := \sup_n U^n_{r,s}$$

er derfor endelig P-n.o., hvis $\sup_k E[X_k^-]$ eller ækvivalent $\sup_k E[\,|X_k|]$ er endelig.

Sætning Ma 4 *Konvergenssætninger.*
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en sub- eller supermartingal. Da eksisterer $\lim_n X_n(\omega)$ i $\mathbf{R}$ for P-n.a. ω hvis $\sup_k E[\,|X_k|] < \infty$, og grænsefunktionen er P-integrabel. Udtrykt ved hjælp af X_∞ kan dette formuleres som

$$\sup_k E[\,|X_k|] < \infty \;\Rightarrow\; X_n \to X_\infty \;P\text{-n.o. og } X_\infty \in L^1(P).$$

Hvis $\{X_n \,|\, n \geq 1\}$ er uniformt integrabel er der yderligere konvergens i P-middel. Hvis $\{X_n^+ \,|\, n \geq 1\}$ er uniformt integrabel udvider submartingalegenskaben til 'tidspunkt' ∞, idet der gælder

$$X_n \leq E[X_\infty \,|\, \mathcal{F}_n] \quad P\text{-n.o. for alle } n \geq 0,$$

og analogt for supermartingaler. I martingaltilfældet er

$$X_n = E[X_\infty \,|\, \mathcal{F}_n] \quad P\text{-n.o. for alle } n \geq 0,$$

hvis og kun hvis $\{X_n \,|\, n \geq 1\}$ er uniformt integrabel.

Sætning Ma 5 *Optional Sampling.*
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en martingal og lad $0 \leq \sigma \leq \tau$ betegne to stoptider, hvor τ er optional for $(X_n)_{n\geq 0}$. Da er X_τ og X_σ elementer i $L^1(P)$ og

$$E[X_0] = E[X_\sigma] = E[X_\tau] \;\text{ samt }\; E[X_\tau \,|\, \mathcal{F}_\sigma] = X_\sigma \quad P\text{ -n.o.}$$

Der gælder et tilsvarende resultat for sub- og supermartingaler med = erstattet af det relevante ulighedstegn. Optionalitetskravet kan her svækkes lidt, idet det for en submartingal er nok, at τ er optional for processen $(X_n^+)_{n\geq 0}$, og tilsvarende for en supermartingal nok at τ er optional for $(X_n^-)_{n\geq 0}$.

Sætning Ma 1 og Ma 2.

Bevis for **Sætning Ma 1** (submartingal tilfældet). Lad τ betegne en stoptid. Da

$$|X_{\tau\wedge n}| \le \sum_{k=0}^{n} |X_k| \quad \text{og} \quad X_{\tau\wedge n} = \sum_{k=0}^{n-1} \mathbf{1}_{\{\tau=k\}} X_k + \mathbf{1}_{\{\tau\ge n\}} X_n$$

er $X_{\tau\wedge n}$ integrabel og $\mathcal{F}_n$-målelig for alle $n \ge 0$. Ligeledes gælder for ethvert n ifølge regnereglerne for betingede middelværdier, da

$$\mathbf{1}_{\{\tau\le n\}} X_\tau = \mathbf{1}_{\{\tau\le n\}} X_{\tau\wedge n} \text{ og } \{\tau > n\} \text{ er } \mathcal{F}_n\text{-målelige,}$$

at

$$E[X_{\tau\wedge(n+1)} \,|\, \mathcal{F}_n] = E[\,\mathbf{1}_{\{\tau\le n\}} X_\tau + \mathbf{1}_{\{\tau>n\}} X_{n+1} \,|\, \mathcal{F}_n] =$$

$$\mathbf{1}_{\{\tau\le n\}} X_\tau + \mathbf{1}_{\{\tau>n\}} \cdot E[X_{n+1} \,|\, \mathcal{F}_n] \ge \mathbf{1}_{\{\tau\le n\}} X_\tau + \mathbf{1}_{\{\tau>n\}} X_n = X_{\tau\wedge n}.$$

Dette viser den første påstand. Lad dernæst σ og τ betegne to stoptider, så at $\sigma \le \tau \le m$, hvor $m \ge 1$ er et helt tal. Integrabiliteten af X_τ og X_σ følger som ovenfor og ifølge det netop viste, er

$$E[X_\tau \,|\, \mathcal{F}_n] = E[X_{\tau\wedge m} \,|\, \mathcal{F}_n] \ge X_{\tau\wedge n} \quad P\text{-n.o.}$$

for alle $0 \le n \le m$. Men heraf fås ifølge Ma 5, at

$$E[X_\tau \,|\, \mathcal{F}_\sigma] = \sum_{n=0}^{m} E[X_\tau \,|\, \mathcal{F}_n] \cdot \mathbf{1}_{\{\sigma=n\}} \ge \sum_{n=0}^{m} X_{\tau\wedge n} \cdot \mathbf{1}_{\{\sigma=n\}} = X_{\tau\wedge\sigma}.$$

Uligheden $E[X_\sigma] \le E[X_\tau]$ følger herefter umiddelbart ved at tage middelværdi på begge sider, og $E[X_0] \le E[X_\sigma]$ er et specialtilfælde heraf svarende til et passende valg af begrænsede stoptider. $\diamond$

Som en umiddelbar konsekvens ses at for enhver stoptid τ og ethvert $n \ge 0$ er

$$X_{\tau\wedge n} = E[X_n \,|\, \mathcal{F}_{\tau\wedge n}]$$

i martingaltilfældet og

$$X^+_{\tau\wedge n} \le E[X^+_n \,|\, \mathcal{F}_{\tau\wedge n}]$$

i submartingaltilfældet. Egenskaber ved betingede middelværdier giver derfor anledning til flg. korollar

Korollar *For enhver stoptid τ er $(X_{\tau\wedge n}, \mathcal{F}_n)_{n\ge0}$ en uniform integrabel martingal h.h.v. begrænset i $L^1(P)$, hvis dette gælder for $(X_n, \mathcal{F}_n)_{n\ge0}$.*
Tilsvarende er $(X^+_{\tau\wedge n}, \mathcal{F}_n)_{n\ge0}$ en uniform integrabel submartingal h.h.v. begrænset i $L^1(P)$, hvis dette gælder for $(X^+_n, \mathcal{F}_n)_{n\ge0}$.

Som en første anvendelse af den viste skrabede udgave af Optional Sampling vises **Sætning Ma 2**. Lad derfor $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en submartingal og lad $\lambda > 0$ være givet. Definer, idet $\inf \emptyset := \infty$,

$$\tau_\lambda := \inf\{n \geq 0 \,|\, X_n > \lambda\} \quad \text{og} \quad \sigma_\lambda := \inf\{n \geq 0 \,|\, X_n < -\lambda\}.$$

Da $(X_n)_{n\geq 0}$ er tilpasset, er τ_λ og σ_λ stoptider, og for alle n gælder

$$\{\max_{0\leq k\leq n} X_k > \lambda\} = \{\tau_\lambda \leq n\} = \{X_{\tau_\lambda \wedge n} > \lambda\} \cap \{\tau_\lambda \leq n\}$$

og tilsvarende

$$\{\min_{0\leq k\leq n} X_k < -\lambda\} = \{\sigma_\lambda \leq n\} = \{-X_{\sigma_\lambda \wedge n} > \lambda\} \cap \{\sigma_\lambda \leq n\}.$$

Ifølge den ovenfor viste 'skrabede udgave' af Optional Sampling gælder derfor for alle n, da $\{\tau_\lambda \leq n\} \in \mathcal{F}_{\tau_\lambda \wedge n}$ og tilsvarende $\{\sigma_\lambda > n\} = \{\sigma_\lambda \leq n\}^c \in \mathcal{F}_{\sigma_\lambda \wedge n}$, at

$$\lambda \cdot P(\max_{0\leq k\leq n} X_k > \lambda) \leq E[X_{\tau_\lambda \wedge n},\, \tau_\lambda \leq n]$$

$$\leq E[X_n,\, \tau_\lambda \leq n] \leq E[X_n^+, \tau_\lambda \leq n] \leq E[X_n^+].$$

og tilsvarende

$$\lambda \cdot P(\min_{0\leq k\leq n} X_k < -\lambda) \leq E[-X_{\sigma_\lambda \wedge n},\, \sigma_\lambda \leq n]$$

$$= E[X_{\sigma_\lambda \wedge n},\, \sigma_\lambda > n] - E[X_{\sigma_\lambda \wedge n}]$$

$$\leq E[X_n,\, \sigma_\lambda > n] - E[X_{\sigma_\lambda \wedge n}] \leq E[X_n^+] - E[X_0].$$

Da absolutværdien af en martingal er en submartingal, gælder specielt.

Doob's Ulighed.
Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en martingal. Da gælder for alle $\lambda > 0$ og $n \geq 0$

$$\lambda \cdot P(\max_{0\leq k\leq n} |X_k| > \lambda) \leq E[\,|X_n|, \max_{0\leq k\leq n} |X_k| > \lambda\,]$$

og dermed (se nedenstående momentulighed) *for alle $p > 1$*

$$\|\max_{0\leq k\leq n} |X_k|\,\|_p \leq p/(1-p) \cdot \|\, X_n \,\|_p \quad n \geq 0.$$

Momentulighed. Hvis ikke-negative stokastiske variable X og Y opfylder

$$\lambda \cdot P(\,|X| > \lambda) \leq E[\,|Y|, \,|X| > \lambda] \;\text{ for alle } \lambda > 0,$$

er for alle $p > 1$

$$\|X\|_p = E[\,|X|^p]^{1/p} \leq p/(1-p) \cdot E[\,|Y|^p]^{1/p} = p/(1-p) \cdot \|Y\|_p.$$

Sætning Ma 3 og Ma 4.

Bevis for **Sætning Ma 3**. Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ betegne en supermartingal og $r < s$ reelle tal. Da $(X_n)_{n\geq 0}$ er tilpasset, definerer

$$\tau_1 := \inf\{n \geq 0 \,|\, X_n < r\} \quad \text{og} \quad \sigma_1 := \inf\{n \geq \tau_1 \,|\, X_n > s\}$$

$$\tau_k := \inf\{n \geq \sigma_{k-1} \,|\, X_n < r\} \quad \text{og} \quad \sigma_k := \inf\{n \geq \tau_k \,|\, X_n > s\} \quad k > 1$$

stoptider, så at $\tau_1 \leq \sigma_1 \leq \tau_2 \leq \sigma_2 \leq \cdots$. Bemærk at $\sigma_k > k$ samt at $X_{\tau_k} < r$ på $\{\tau_k < \infty\}$ og tilsvarende $X_{\sigma_k} > s$ på $\{\sigma_k < \infty\}$ og dermed da $\tau_k \leq \sigma_k$

$$X_{\sigma_k} - X_{\tau_k} > s - r \quad \text{på} \;\; \{\sigma_k < \infty\}.$$

Definer herudfra for $n \geq 1$ antallet af *opkrydsninger* $U^n_{r,s}$ over intervallet $[r, s]$ i tidsrummet $\{0, 1, \ldots n\}$, d.v.s.

$$U^n_{r,s} := \sup\{k \,|\, \sigma_k \leq n\} = \sum_{k=1}^{n} \mathbf{1}_{\{\sigma_k \leq n\}}.$$

For ethvert n, $k \geq 1$ er

$$\mathbf{1}_{\{\tau_k \leq n\}} = \mathbf{1}_{\{\sigma_k \leq n\}} + \mathbf{1}_{\{\tau_k \leq n < \sigma_k\}},$$

d.v.s.

$$\sum_{k=1}^{n}(X_{\sigma_k \wedge n} - X_{\tau_k \wedge n}) \cdot \mathbf{1}_{\{\tau_k \leq n\}} = \sum_{k=1}^{n}(X_{\sigma_k} - X_{\tau_k}) \cdot \mathbf{1}_{\{\sigma_k \leq n\}}$$

$$+ \sum_{k=1}^{n}(X_n - X_{\tau_k}) \cdot \mathbf{1}_{\{\tau_k \leq n < \sigma_k\}} \geq (s - r) \cdot U^n_{r,s} + \sum_{k=1}^{n}(X_n - r) \cdot \mathbf{1}_{\{\tau_k \leq n < \sigma_k\}}$$

$$\geq (s - r) \cdot U^n_{r,s} - (X_n - r)^- \cdot \sum_{k=1}^{n} \mathbf{1}_{\{\tau_k \leq n < \sigma_k\}} \geq (s - r) \cdot U^n_{r,s} - (X_n - r)^-,$$

og dermed

$$(s - r) \cdot U^n_{r,s} \leq (X_n - r)^- + \sum_{k=1}^{n}(X_{\sigma_k \wedge n} - X_{\tau_k \wedge n}) \cdot \mathbf{1}_{\{\tau_k \leq n\}}$$

Men for alle $1 \leq k \leq n$ gælder ifølge den 'skrabede udgave' af Optional Sampling

$$E[(X_{\sigma_k \wedge n} - X_{\tau_k \wedge n}) \cdot \mathbf{1}_{\{\tau_k \leq n\}}] = E[\, E[X_{\sigma_k \wedge n} - X_{\tau_k \wedge n} \,|\, \mathcal{F}_{\tau_k \wedge n}] \cdot \mathbf{1}_{\{\tau_k \leq n\}}] \leq 0,$$

og dermed

$$(s - r) \cdot E[U^n_{r,s}] \leq E[(X_n - r)^-] \leq E[X_n^-] + |r|. \qquad \diamond$$

Korollar *Lad* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *betegne en sub- eller en supermartingal, som er begrænset i* $L^1(P)$, *d.v.s.* $\sup_n E[\,|X_n|] < \infty$. *Da er*

$$P(-\infty < \liminf_n X_n = \limsup_n X_n < \infty) = 1,$$

d.v.s. $\lim_n X_n(\omega)$ *eksisterer i* $\mathbf{R}$ *for* P*-n.a.* ω.

Bevis. Den simple sammenhæng mellem sub - og supermartingaler viser, at vi uden tab af generalitet kan antage, at $(X_n, \mathcal{F}_n)_{n\geq 0}$ er en supermartingal. Da

$$\sup_n E[\,|X_n|] < \infty \Rightarrow P(\sup_n |X_n| < \infty) = 1$$

ifølge maksimalulighederne, udestår kun at vise, at lighedstegnet holder P-n.o. Men gælder dette ikke, eksisterer der, da $\mathbf{R}$ er separabel, reelle tal $r < s$, så at

$$0 < P(\liminf_n X_n < r < s < \limsup_n X_n) \leq P(\sup_n U^n_{r,s} = \infty),$$

hvilket strider mod Sætning Ma 3, da $E[\sup_n U^n_{r,s}] = \sup_n E[U^n_{r,s}] < \infty$. $\diamond$

Bemærkning. Da $E[\,|X_n|] = E[X_n] \leq E[X_0]$ for enhver ikke-negativ supermartingal $(X_n, \mathcal{F}_n)_{n\geq 0}$ er en sådan altid konvergent P-n.o.

Med X_∞ defineret i henhold til den vedtagne konvention, kan det viste alternativt formuleres som.

Martingal Konvergenssætningen.
For enhver sub- eller supermartingal $(X_n, \mathcal{F}_n)_{n\geq 0}$ *gælder*

$$\sup_k E[\,|X_k|] < \infty \;\Rightarrow\; X_n \to X_\infty \ P\text{-n.o. og } X_\infty \in L^1(P).$$

Hvis $\{X_n \,|\, n \geq 0\}$ *er uniformt integrabel, er der yderligere konvergens i* $L^1(P)$.

Integrabiliteten følger af Fatou's Lemma, idet

$$E[\,|X_\infty|] \leq \liminf_n E[\,|X_n|] \leq \sup_n E[\,|X_n|] < \infty,$$

og konvergensen i $L^1(P)$ følger dernæst af Sætning 6. Det sidste punkt kan præciseres yderligere.

Korollar *For enhver martingal* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *gælder*

$$\{X_n \,|\, n \geq 0\} \text{ uniformt integrabel} \;\Rightarrow\; X_n = E[X_\infty \,|\, \mathcal{F}_n] \quad P\text{-n.o. } n \geq 0,$$

og hvis $(X_n, \mathcal{F}_n)_{n\geq 0}$ *er en submartingal gælder*

$$\{X_n^+ \,|\, n \geq 0\} \text{ uniformt integrabel} \;\Rightarrow\; X_n \leq E[X_\infty \,|\, \mathcal{F}_n] \quad P\text{-n.o. } n \geq 0,$$

Bevis. Da betinget middelværdi er en kontraktion i $L^1(P)$, har vi

$$X_k \to X_\infty \text{ i } L^1(P) \;\Rightarrow\; E[X_k \,|\, \mathcal{F}_n] \to_{k\to\infty} E[X_\infty \,|\, \mathcal{F}_n] \text{ i } L^1(P)$$

for alle n, hvoraf martingaltilfældet umiddelbart følger, da

$$X_n = E[X_k \,|\, \mathcal{F}_n] \quad P\text{-n.o. for } k \geq n.$$

Betragt dernæst submartingaltilfældet. Antagelsen sikrer at $\sup_k E[\,|X_k|] < \infty$ og dermed $X_n \to X_\infty$ P-n.o. For ethvert $m \geq 0$ konvergerer derfor

$$X_n \vee (-m) \to X_\infty \vee (-m) \;\; P\text{-n.o.},$$

og også i $L^1(P)$, idet $\{X_n \vee (-m) \,|\, n \geq 0\}$ er uniformt integrabel, da

$$|\, X_n \vee (-m)| \leq X_n^+ + m \quad \text{for alle } n,\, m \geq 0.$$

Ved fornyet brug af, at betingede middelværdier er kontraktioner i $L^1(P)$, fås derfor

$$E[X_k \vee (-m) \,|\, \mathcal{F}_n] \to_{k\to\infty} E[X_\infty \vee (-m) \,|\, \mathcal{F}_n] \;\text{ i } L^1(P)$$

for alle $n \geq 0$, og dermed $X_n \leq E[X_\infty \vee (-m) \,|\, \mathcal{F}_n]$ P-n.o. for alle $m \geq 0$, da

$$X_n \leq X_n \vee (-m) \leq E[X_k \vee (-m) \,|\, \mathcal{F}_n] \;\; P\text{-n.o.}$$

for alle $n \leq k$. Det ønskede resultat følger nu ved grænseovergang, idet

$$E[X_\infty \,|\, \mathcal{F}_n] = \inf_{m\geq 0} E[X_\infty \vee (-m) \,|\, \mathcal{F}_n] \;\; P\text{-n.o.}$$

for ethvert $n \geq 0$ ifølge egenskaber ved betingede middelværdier. ◇

Konvergenssætningen giver anledning til et par interessante korollarer.

Lévy's Sætning
For ethvert $X \in L^1(P)$ og ethvert filter $(\mathcal{G}_n)_{n\geq 0}$ konvergerer

$$E[X \,|\, \mathcal{G}_n] \to E[X \,|\, \mathcal{G}_\infty] \quad P\text{-n.o. og i } L^1(P),$$

specielt gælder for enhver $(\mathcal{G}_n)$-stoptid τ, at hvis $X_n = E[X \,|\, \mathcal{G}_n]$ for $n \geq 0$, så er

$$X_\tau = E[X \,|\, \mathcal{G}_\tau] \quad P\text{-n.o.}$$

Bevis. Lad $X \in L^1(P)$ og $(\mathcal{G}_n)_{n\geq 0}$ være givet. $(E[X \,|\, \mathcal{G}_n], \mathcal{G}_n)_{n\geq 0}$ udgør da en uniform integrabel martingal, og ifølge martingalkonvergenssætningen findes der derfor et element $\tilde{X} \in L^1(P)$, som er $\mathcal{G}_\infty$-målelig, så at

$$E[X \,|\, \mathcal{G}_n] \to \tilde{X} \;\; P\text{-n.o. og i } L^1(P) \;\text{ og }\; E[X \,|\, \mathcal{G}_n] = E[\tilde{X} \,|\, \mathcal{G}_n] \;\; P\text{-n.o. } n \geq 0.$$

For alle $n \geq 0$ og alle $B \in \mathcal{G}_n$ gælder dermed

$$\int_B X\,dP = \int_B E[X\,|\,\mathcal{G}_n]\,dP = \int_B E[\tilde{X}\,|\,\mathcal{G}_n]\,dP = \int_B \tilde{X}\,dP$$

hvilket, da $\bigcup_n \mathcal{G}_n$ er en algebra, som frembringer $\mathcal{G}_\infty$, ifølge Proposition 5 Lemma 2 viser, at

$$\tilde{X} = E[X\,|\,\mathcal{G}_\infty] \quad P\text{-n.o.}$$

og dermed korollarets første del. Da

$$X_{\tau\wedge n} = E[X_n\,|\,\mathcal{G}_{\tau\wedge n}] = E[X\,|\,\mathcal{G}_{\tau\wedge n}] \quad P\text{-n.o.}$$

ifølge den skrabede udgave af Optional Sampling og $\sigma(\cup_n \mathcal{G}_{\tau\wedge n}) = \mathcal{G}_\tau$ ifølge Ma 6, sikrer første del, at

$$X_{\tau\wedge n} \to E[X\,|\,\mathcal{G}_\tau] \quad P\text{-n.o.}$$

Men ifølge martingal konvergenssætningen gælder også $X_{\tau\wedge n} \to X_\tau$ P-n.o., hvilket tilsammen viser den ønskede lighed. $\diamond$

Bemærkning. Da enhver betinget middelværdi er en kontraktion i $L^1(P)$ kan punkt 1 i Levý's Sætning generaliseres til

$$X_n \to X \text{ i } L^1(P) \;\Rightarrow\; E[X_n\,|\,\mathcal{G}_n] \to E[X\,|\,\mathcal{G}_\infty] \text{ i } L^1(P).$$

L^p-konvergens. $(p > 1)$
For enhver martingal $(M_n, \mathcal{F}_n)_{n\geq 0}$ og ethvert $p > 1$ gælder:

$$\sup_n E[\,|M_n|^p] < \infty \;\Leftrightarrow\; \lim_n M_n \text{ eksisterer } P\text{-n.o. og i } L^p(P).$$

Bevis. Implikationen $\Leftarrow$ er en umiddelbar konsekvens af L^p-konvergens. Men hvis $\sup_n E[\,|M_n|^p] < \infty$ er $\{M_n\,|\,n \geq 0\}$ uniformt integrabel, og ifølge martingal konvergenssætningen findes der derfor et $M \in L^1(P, \mathcal{F}_\infty)$, så at

$$M_n = E[M\,|\,\mathcal{F}_n] \;\text{ og }\; M_n \to M \quad P\text{-n.o.}$$

Ifølge Fatou's Lemma gælder derfor

$$E[\,|M|^p] \leq \liminf_n E[\,|M_n|^p] \leq \sup_n E[\,|M_n|^p] < \infty,$$

d.v.s. $M \in L^p(P)$, og da

$$|M_n|^p \leq E[\,|M|^p\,|\,\mathcal{F}_n] \quad P\text{-n.o. for alle } n \geq 0$$

ifølge Jensen's ulighed for betingede middelværdier, har vi alt i alt, at $\lim_n M_n$ eksisterer P-n.o. og $\{|M_n|^p\,|\,n \geq 0\}$ er uniformt integrabel. $L^p(P)$-konvergensen følger derfor af korollaret til Sætning 6. $\diamond$

Korollaret omhandler, som det fremgår, kun tilfældet $p > 1$ og gælder generelt ikke for $p = 1$, men hvis den betragtede martingal er afsnitsfølgen hørende til en sum af uafhængige centrerede variable, kan man ved hjælp af korollaret til Ottavianis ulighed udvide til $p = 1$. Der gælder nemlig flg. resultat

Lad $(X_i)_{i\geq 1}$ betegne en følge af uafhængige stokastiske variable med tilhørende afsnitssummer $S_n := \sum_{i=1}^n X_i$ for $n \geq 1$. Da gælder

$$\sup_n E[\,|S_n|] < \infty \;\Rightarrow\; \lim_n (S_n - E[S_n]) \;\text{ eksisterer } P\text{-n.o. og i } L^1(P).$$

D.v.s. hvis $(S_n)_{n\geq 1}$ er begrænset i $L^1(P)$, så eksisterer $\lim_n S_n$ P-n.o. og i $L^1(P)$, hvis alle X_i'erne har middelværdi 0, eller hvis $(S_n)_{n\geq 1}$ er konvergent i fordeling.

Bevis. Da $|E[S_n]| \leq E[\,|S_n|]$ for alle n udgør $(E[S_n])_{n\geq 1}$ en begrænset talfølge, og dermed ifølge trekantuligheden er $\sup_n E[\,|S_n - E[S_n]|\,] < \infty$. Martingal konvergenssætningen brugt på martingalen

$$(S_n - E[S_n])_n = (\sum_{i=1}^n (X_i - E[X_i]))_n$$

giver derfor, at

$$\lim_n (S_n - E[S_n]) \;\text{ eksisterer } P\text{ -n.o.}$$

Men ifølge Lebesgue's Sætning er der også konvergens i $L^1(P)$, da ovenstående omtalte korollar giver vurderingen

$$E[\sup_n |S_n - E[S_n]\,|\,] \leq 6 \cdot \sup_n E[\,|S_n - E[S_n]\,|\,] < \infty,$$

Bemærkningen er klar, hvad angår tilfældet med middelværdi 0. I forbindelse med det andet benyttes at en reel talfølge $(a_n)_{n\geq 1}$ er konvergent i $\mathbf{R}$, hvis der findes en følge af stokastiske variable $(Y_n)_{n\geq 1}$, så at

$$(Y_n)_{n\geq 1} \text{ og } (Y_n - a_n)_{n\geq 1} \text{ begge konvergerer i fordeling.}$$

Thi anvendes dette på $(S_n)_{n\geq 1}$ og $(E[S_n])_n$ ses, at $\lim_n E[S_n]$ eksisterer i $\mathbf{R}$ og dermed ved addition, at

$$\lim_n S_n \;\text{ eksisterer } P\text{-n.o. og i } L^1(P).$$

Optionalitet.

Notation. Hvis $(\mathcal{F}_n)_{n\geq 0}$ er et filter i $\mathcal{F}$ og $(Y_n)_{n\geq 0}$ en reel tilpasset stokastisk proces, siges en stoptid τ at være *optional* for $(Y_n)_{n\geq 0}$, hvis $\{Y_{\tau\wedge n}\,|\,n\geq 0\}$ er uniformt integrabel.

Optionalitet er altså et krav om, at processen $(Y_n)_{n\geq 0}$ ikke må blive 'for stor' før tidspunkt τ. Dette giver anledning til flg. simple kriterier for optionalitet.

a) τ er optional for $(Y_n)_{n\geq 0}$, hvis $|Y_{\tau\wedge n}|\leq Y$ P-n.o. for alle $n\geq 0$ for et $Y\in L^1(P)$, eller ækvivalent hvis $\sup_n |Y_{\tau\wedge n}|\in L^1(P)$.

b) τ er optional for $(Y_n)_{n\geq 0}$, hvis der findes et $\alpha>1$, så at $\sup_n E[\,|Y_{\tau\wedge n}|^\alpha]<\infty$.

Hvis Y_n'erne er integrable, d.v.s. elementer i $L^1(P)$, er enhver begrænset stoptid τ optional for $(Y_n)_{n\geq 0}$, thi hvis $\tau\leq M$ for et positivt helt tal M er

$$|Y_{\tau\wedge n}|\leq\sum_{k=0}^{M}|Y_k|\quad\text{for alle } n\geq 0,$$

og da højresiden er integrabel, fås optionaliteten umiddelbart af a).

Ud over kriterierne a) og b) er det også værd at huske, at Bm 16 sommetider kan benyttes til at vise optionalitet, samt at

c) τ er optional for $(Y_n)_{n\geq 0}$, hvis $\lim Y_{\tau\wedge n}$ eksisterer i $L^1(P)$.

Opskrivningerne

$$Y_{\tau\wedge\sigma\wedge n}=\mathbf{1}_{\{\sigma>\tau\}}Y_{\tau\wedge n}+\mathbf{1}_{\{\sigma\leq\tau\}}Y_{\sigma\wedge n},\quad Y_{(\tau\vee\sigma)\wedge n}=\mathbf{1}_{\{\sigma<\tau\}}Y_{\tau\wedge n}+\mathbf{1}_{\{\sigma\geq\tau\}}Y_{\sigma\wedge n}$$

og dermed

$$|Y_{\tau\wedge\sigma\wedge n}|\leq|Y_{\tau\wedge n}|+|Y_{\sigma\wedge n}|\quad\text{og}\quad|Y_{(\tau\vee\sigma)\wedge n}|\leq|Y_{\tau\wedge n}|+|Y_{\sigma\wedge n}|,$$

viser, at hvis τ og σ er optionale stoptider for $(Y_n)_{n\geq 0}$, så er både $\tau\wedge\sigma$ og $\tau\vee\sigma$ optional for $(Y_n)_{n\geq 0}$.

Pr. definition af Y_τ er $|Y_\tau|\leq\liminf_n|Y_{\tau\wedge n}|$ P-n.o., og ifølge Fatou's Lemma gælder derfor, at

$$E[\,|Y_\tau|]\leq\liminf_n E[\,|Y_{\tau\wedge n}|]\leq\sup_n E[\,|Y_{\tau\wedge n}|],$$

d.v.s. τ kan kun være optional, hvis $E[\,|Y_\tau|]<\infty$. Men betingelsen er langt fra tilstrækkelig, og vi vil nu se undersøge på, hvad der yderligere skal til. Ligheden

$$Y_{\tau\wedge n}=\mathbf{1}_{\{\tau>n\}}Y_n+\mathbf{1}_{\{\tau\leq n\}}Y_\tau$$

viser, da $\{\tau>n\}$ og $\{\tau\leq n\}$ er disjunkte, at τ er optional, hvis og kun hvis

$$\{\mathbf{1}_{\{\tau\leq n\}}Y_\tau\mid n\geq 0\}\ \text{ og }\ \{\mathbf{1}_{\{\tau>n\}}Y_n\,|\,n\geq 0\}$$

begge er uniformt integrable. Her er der ingen problemer med den første, thi da $\mathbf{1}_{\{\tau\le n\}}|Y_\tau| \uparrow \mathbf{1}_{\{\tau<\infty\}}|Y_\tau|$ P-n.o. ses at

$$\{\mathbf{1}_{\{\tau\le n\}}Y_\tau \mid n \ge 0\} \text{ uniformt integrabel } \Leftrightarrow \int_{\{\tau<\infty\}} |Y_\tau|\, dP < \infty.$$

Den anden er derimod mere kompliceret. Men opskrivningen

$$\mathbf{1}_{\{\tau>n\}}Y_n = \mathbf{1}_{\{n<\tau<\infty\}}Y_n + \mathbf{1}_{\{\tau=\infty\}}Y_n$$

viser som før, da $\mathbf{1}_{\{n<\tau<\infty\}}Y_n \to 0$ P-n.o., at

$$\{\mathbf{1}_{\{\tau>n\}}Y_n \mid n \ge 0\} \text{ uniformt integrabel hvis og kun hvis}$$

$$\{\mathbf{1}_{\{\tau=\infty\}}Y_n \mid n \ge 0\} \text{ uniformt integrabel og } \lim_n \int_{\{n<\tau<\infty\}} |Y_n|\, dP = 0.$$

Alt i alt har vi vist flg. optionalitetskriterium.

En endelig stoptid τ er optional for $(Y_n)_{n\ge0}$ hvis og kun hvis

$$\int_{\{\tau<\infty\}} |Y_\tau|\, dP < \infty \quad \text{og} \quad \lim_n \int_{\{n<\tau<\infty\}} |Y_n|\, dP = 0,$$

Bemærk at da τ er endelig, er

$$\int_{\{\tau<\infty\}} |Y_\tau|\, dP = E[\,|Y_\tau|\,] \quad \text{og} \quad \int_{\{n<\tau<\infty\}} |Y_n|\, dP = E[\,|Y_n|; \tau > n] \quad n \ge 0$$

Som allerede nævnt er en stoptid τ optional, hvis processen ikke bliver for stor før tid τ. Det er derfor nærliggende at tro, at optionalitet af τ medfører optionalitet af enhver stoptid σ, som er mindre end τ, d.v.s. opfylder $\sigma \le \tau$. Dette gælder dog ikke altid, men ligheden

$$Y_{\sigma\wedge n} = \mathbf{1}_{\{\sigma\le n\}}Y_\sigma + \mathbf{1}_{\{\sigma>n\}}Y_n = \mathbf{1}_{\{\sigma\le n\}}Y_\sigma + \mathbf{1}_{\{\sigma>n\}}Y_{\tau\wedge n}$$

viser, at det er sandt, hvis $\{\mathbf{1}_{\{\sigma\le n\}}Y_\sigma \,|\, n \ge 0\}$ er uniformt integrabel, hvilket, som ovenfor vist, er ensbetydende med, at

$$\int_{\{\sigma<\infty\}} |Y_\sigma|\, dP < \infty.$$

Men som før bemærket, er denne egenskab også nødvendig, d.v.s. hvis en stoptid τ er optional for $(Y_n)_{n\ge0}$, så gælder for enhver stoptid $\sigma \le \tau$, at

$$\sigma \text{ er optional for } (Y_n)_{n\ge0} \Leftrightarrow \int_{\{\sigma<\infty\}} |Y_\sigma|\, dP < \infty.$$

Sætning Ma 5.

Bevis for **Sætning Ma 5**. Lad $(X_n, \mathcal{F}_n)_{n\geq 0}$ være en submartingal og $0 \leq \sigma \leq \tau$ stoptider, hvor τ antages optional for $(X_n^+)_{n\geq 0}$, d.v.s. $\{X^+_{\tau\wedge n} \,|\, n \geq 0\}$ er uniformt integrabel.
Da $\sigma \wedge n \leq \tau \wedge n$ viser den 'skrabede udgave' af Optional Sampling at

$$X_{\sigma\wedge n} \leq E[X_{\tau\wedge n} \,|\, \mathcal{F}_{\sigma\wedge n}] \text{ og dermed } X^+_{\sigma\wedge n} \leq E[X^+_{\tau\wedge n} \,|\, \mathcal{F}_{\sigma\wedge n}]$$

for $n \geq 0$, og ifølge egenskaber ved betingede middelværdier er $\{X^+_{\sigma\wedge n} \,|\, n \geq 0\}$ derfor uniformt integrabel, d.v.s. σ er også optional for $(X_n^+)_{n\geq 0}$.
Benyttes martingal konvergenssætningen på de to submartingaler $(X_{\sigma\wedge n}, \mathcal{F}_n)_{n\geq 0}$ og $(X_{\tau\wedge n}, \mathcal{F}_n)_{n\geq 0}$ fås derfor, at X_σ og X_τ er elementer i $L^1(P)$ og

$$X_{\sigma\wedge n} \rightarrow_{n\to\infty} X_\sigma \text{ og } X_{\tau\wedge n} \rightarrow_{n\to\infty} X_\tau \quad P\text{-n.o.}$$

Korollaret til martingal konvergenssætningen viser endvidere, at

$$X_{\tau\wedge n} \leq E[X_\tau \,|\, \mathcal{F}_n] \quad P\text{-n.o.}$$

og dermed ifølge den skrabede udgave af Optional Sampling

$$X_{\sigma\wedge n} \leq E[X_{\tau\wedge n} \,|\, \mathcal{F}_{\sigma\wedge n}] \leq E[\, E[X_\tau \,|\, \mathcal{F}_n] \,|\, \mathcal{F}_{\sigma\wedge n}] = E[X_\tau \,|\, \mathcal{F}_{\sigma\wedge n}].$$

Som allerede nævnt konvergerer venstresiden her P-n.o. mod X_σ, og da

$$E[X_\tau \,|\, \mathcal{F}_{\sigma\wedge n}] \rightarrow_{n\to\infty} E[X_\tau \,|\, \mathcal{F}_\sigma] \quad P\text{-n.o.},$$

ifølge Lévy's Sætning ses at

$$X_\sigma \leq E[X_\tau \,|\, \mathcal{F}_\sigma] \quad P\text{-n.o.}$$

og dermed også uligheden $E[X_\sigma] \leq E[X_\tau]$. ◇

I martingaltilfældet gælder tilsvarende formler blot med lighedstegn overalt.

Som en konsekvens af Ma 7 og det viste resultat har vi flg. korollar.

Korollar. *Lad* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *være en martingal og* τ *en stoptid, som er optional for* $(X_n)_{n\geq 0}$. *Da er*

$$X_{\sigma\wedge\tau} = E[X_\tau \,|\, \mathcal{F}_\sigma] \quad P\text{-n.o.}$$

for enhver stoptid σ, *d.v.s.* $\{X_\sigma \,|\, \sigma \text{ stoptid}, \ \sigma \leq \tau\}$ *er uniformt integrabel.*

Hvis $(X_n, \mathcal{F}_n)_{n\geq 0}$ *er en submartingal og* τ *optional for* $(X_n^+)_{n\geq 0}$, *gælder tilsvarende*

$$X_{\sigma\wedge\tau} \leq E[X_\tau \,|\, \mathcal{F}_\sigma] \quad P\text{-n.o.}$$

for enhver stoptid σ *og* $\{X_\sigma^+ \,|\, \sigma \text{ stoptid}, \ \sigma \leq \tau\}$ *er derfor uniformt integrabel.*

Appendiks A. Konvekse funktioner.

I det følgende betegner $J \subseteq \mathbf{R}$ et interval med endepunkter v_J og h_J, hvor $-\infty \le v_J < h_J \le \infty$. $J^\circ :=]v_j, h_J[$.

Notation. $l: J \to \mathbf{R}$ siges at være *affin*, hvis

$$l(x) = ax + b \quad x \in J \quad \text{hvor } a, b \in \mathbf{R}.$$

$\varphi: J \to \mathbf{R}$ siges at være *konveks*, hvis

$$\varphi(\alpha x + (1-\alpha)y) \le \alpha\varphi(x) + (1-\alpha)\varphi(y) \quad x, y \in J, \ \alpha \in [0,1].$$

Affine funktioner er konvekse, og Middelværdisætningen viser, at hvis J er åben og φ differentiabel i J, er φ konveks, hvis $x \mapsto \varphi'(x)$ er voksende, f.eks. hvis φ'' eksisterer og er ikke-negativ. Desuden er ethvert overalt endeligt punktvis supremum af konvekse funktioner igen konvekst.

Lad i det følgende φ betegne en given konveks funktion. Konveksitetsligningen viser, at $\{\varphi \le a\}$ for ethvert $a \in \mathbf{R}$ er en konveks mængde, d.v.s. et delinterval, i J.

$$\{\varphi > a\} = J \setminus \{\varphi \le a\}$$

er derfor enten et delinterval i J eller en foreningsmængde af to sådanne. Heraf følger at

$$\varphi(x+) := \lim_{y \downarrow x} \varphi(y) \ \text{ og } \ \varphi(x-) := \lim_{y \uparrow x} \varphi(y) \ \text{ eksisterer i } \overline{\mathbf{R}} \text{ for } x \in J. \tag{1}$$

Endvidere eksisterer $\varphi(v_J+)$ *og* $\varphi(h_J-)$.

Bevis. Vi ser kun på $\varphi(\cdot\,-)$, da den anden går analogt. Lad $x \in J \cup \{h_J\}$ være givet. For ethvert $a \in \mathbf{R}$ har vi

$$\limsup_{y \uparrow x} \varphi(y) > a \ \Rightarrow \ \exists \epsilon > 0 \ ]x - \epsilon, x[\subseteq \{\varphi > a\} \ \Rightarrow \ \liminf_{y \uparrow x} \varphi(y) \ge a.$$

Dette viser, at

$$\limsup_{y \uparrow x} \varphi(y) = \liminf_{y \uparrow x} \varphi(y)$$

og dermed, at $\varphi(x-)$ eksisterer i $\overline{\mathbf{R}}$. ◇

$$\varphi(x) \ge \varphi(x+) \vee \varphi(x-) \quad x \in J. \tag{2}$$

Bevis. Vi viser kun $\varphi(\cdot) \ge \varphi(\cdot\,+)$. Lad $x < y$ være punkter i J.

$$\varphi(x+) = \limsup_{z \downarrow x} \varphi(z) = \limsup_{\alpha \uparrow 1} \varphi(\alpha x + (1-\alpha)y)$$

$$\le \limsup_{\alpha \uparrow 1} \alpha\varphi(x) + \limsup_{\alpha \uparrow 1} (1-\alpha)\varphi(y) = \varphi(x). \quad ◇$$

Kontinuitet. *φ er kontinuert i J°.*

Bevis. Lad $x \in J^\circ$ være givet. Da J° er åben og φ konveks, er

$$\varphi(x) \leq \frac{1}{2}(\varphi(x-h)+\varphi(x+h)) \quad \text{for } h \text{ lille,}$$

d.v.s.

$$\varphi(x) \leq \frac{1}{2}\limsup_{h\downarrow 0}(\varphi(x-h)+\varphi(x+h)) = \frac{1}{2}(\varphi(x-)+\varphi(x+)).$$

Hvilket sammen med (2) viser, at $\varphi(x)=\varphi(x-)=\varphi(x+)$. $\diamond$

$$\frac{\varphi(y)-\varphi(x_0)}{y-x_0} \leq \frac{\varphi(x)-\varphi(x_0)}{x-x_0} \qquad y<x_0<x \text{ i } J.$$

Bevis. Lad $y<x_0<x$ i J være givet.

$$\frac{\varphi(y)-\varphi(x_0)}{y-x_0} \leq \frac{\varphi(x)-\varphi(x_0)}{x-x_0} \Leftrightarrow (x-x_0)(\varphi(y)-\varphi(x_0)) \geq (x_0-y)(\varphi(x_0)-\varphi(x))$$

$$\Leftrightarrow \varphi(x_0) \leq \frac{x-x_0}{x-y}\,\varphi(y) + (1-\frac{x-x_0}{x-y})\,\varphi(x)$$

hvilket er opfyldt ifølge konveksiteten, da

$$x_0 = \frac{x-x_0}{x-y}\cdot y + (1-\frac{x-x_0}{x-y})\cdot x. \quad \diamond$$

Heraf ses at for ethvert $x_0 \in J^\circ$ er tallene

$$d^+ := \inf_{x>x_0}\frac{\varphi(x)-\varphi(x_0)}{x-x_0} \quad \text{og} \quad d^- := \sup_{x<x_0}\frac{\varphi(x)-\varphi(x_0)}{x-x_0}$$

begge endelige og $d^- \leq d^+$. Pr. definition af d^- og d^+ har vi

$$\varphi(x) \geq d^+\cdot(x-x_0)+\varphi(x_0) \quad x>x_0 \quad \text{og} \quad \varphi(x) \geq d^-\cdot(x-x_0)+\varphi(x_0) \quad x<x_0$$

og derfor

$$\varphi(x) \geq d^-(x-x_0)+\varphi(x_0) \quad x \in J.$$

Herudfra kan vi nu uddrage to vigtige konsekvenser.
For ethvert $x_0 \in J^\circ$ eksisterer der derfor en affin 'støttefunktion', d.v.s. en affin funktion l_{x_0}, så at

$$l_{x_0} \leq \varphi \quad \text{og} \quad l_{x_0}(x_0)=\varphi(x_0). \tag{3}$$

Lad $(q_i)_{i\geq 1}$ betegne en nummerering af $\mathbf{Q}\cap J^\circ$ og lad for $i \geq 1$ l_i betegne en affin støttefunktion hørende til q_i. Funktionen

$$x \mapsto \sup_i l_i(x) \quad x \in J$$

er derfor konveks, da den er domineret af φ og derfor overalt endelig. Men da den pr. definition stemmer overens med φ på mængden $\mathbf{Q} \cap J^\circ$ følger, da konvekse funktioner er kontinuerte i J°, at

$$\varphi \geq \sup_i l_i \quad \text{og} \quad \varphi(x) = \sup_i l_i(x) \quad x \in J^\circ. \tag{4}$$

Punkt (3) benyttes i beviset for Jensen's ulighed og punkt (4) i beviset for Jensen's ulighed for betingede middelværdier.

Appendiks B.
I gennemgangen af integralteorien har vi udelukkende fokuseret på reelle integranter. Men i forbindelse med emnet karakteristiske funktioner får vi brug for at kunne tale om integralet af en kompleks funktion. Udvidelsen, som naturligvis sker via integration af h.h.v. realdel og imaginærdel, kræver lidt ny notation. $(E, \mathcal{E}, \mu)$ betegner her et vilkårligt målrum.
Definer

$$\mathrm{M}(\mathcal{E}, \mathbf{C}) := \{f : E \to \mathbf{C} \,|\, \Re f, \Im f \in \mathrm{M}(\mathcal{E})\}$$

$$L^1(\mu, \mathbf{C}) := \{f \in \mathrm{M}(\mathcal{E}, \mathbf{C}) \,|\, \Re f, \Im f \in L^1(\mu)\} = \{f \in \mathrm{M}(\mathcal{E}, \mathbf{C}) \,|\, |f| \in L^1(\mu)\}.$$

Det sidste lighedstegn beror på uligheden $|\Re f| \vee |\Im f| \leq |f| \leq |\Re f| + |\Im f|$.
Egenskaber ved $\mathrm{M}(\mathcal{E})$ og $L^1(\mu)$ sikrer, at funktionsrummene $\mathrm{M}(\mathcal{E}, \mathbf{C})$ og $L^1(\mu, \mathbf{C})$ er komplekse vektorrum, samt at $\mathrm{M}(\mathcal{E}, \mathbf{C})$ er stabil under punktvis konvergens. Bemærk at $\mathrm{M}(\mathcal{E})$ og $L^1(\mu)$ er underrum i h.h.v. $\mathrm{M}(\mathcal{E}, \mathbf{C})$ og $L^1(\mu, \mathbf{C})$.
Afbildningen

$$L^1(\mu, \mathbf{C}) \ni f \to \int f \, d\mu := \int \Re f \, d\mu + i \int \Im f \, d\mu \in \mathbf{C}.$$

kaldes stadigt integralet m.h.t. μ, og udnyttes at integralet er lineært på $L^1(\mu)$ ses, at denne afbildning, der udvider integralet på $L^1(\mu)$, er kompleks lineær samt opfylder

$$\overline{\int f \, d\mu} = \int \overline{f} \, d\mu \qquad f \in L^1(\mu, \mathbf{C}).$$

Endvidere er

$$|\int f \, d\mu| \leq \int |f| \, d\mu \quad f \in L^1(\mu, \mathbf{C}),$$

da der for ethvert $f \in L^1(\mu, \mathbf{C})$ findes et $\lambda \in \mathbf{C}$ med længde 1, så at

$$|\int f \, d\mu| = \lambda \int f \, d\mu$$

og dermed, da $|\Re(\lambda f)| \leq |\lambda f| = |f|$,

$$|\int f \, d\mu| = \int \lambda f \, d\mu = \Re(\int \lambda f \, d\mu) = \int \Re(\lambda f) \, d\mu \leq \int |f| \, d\mu.$$

Sammen med Lebesgue's Sætning viser dette flg. konvergensresultat, hvor f og $(f_n)_{n \geq 1}$ er elementer i $L^1(\mu, \mathbf{C})$.

$$f_n \to f \;\; \mu\text{-n.o. og } \; |f_n| \leq g \;\; \mu\text{-n.o., hvor } \; g \in \mathrm{M}(\mathcal{E})_+ \text{ og } \int g \, d\mu < \infty \; \Rightarrow$$

$$\lim_n \int |f_n - f| \, d\mu = 0 \;\; \text{og dermed} \;\; \lim_n \int f_n \, d\mu = \int f \, d\mu.$$

Ikke overraskende gælder der også versioner af Fubini's Sætning for funktioner med komplekse værdier. Men i stedet for at søge efter den mest generelle nøjes vi her med flg. udgave, som vi gang på gang skal benytte i forbindelse med gennemgangen af karakteristiske funktioner.

Lad $(E, \mathcal{E}, \mu)$ og $(F, \mathcal{F}, \nu)$ betegne to endelige målrum og lad $f \in \mathrm{bM}(\mathcal{E} \otimes \mathcal{F}, \mathbf{C})$ være givet (b'et indikerer som sædvanligt, at $|f| \leq M < \infty$ for en positiv konstant M). Da er $f \in L^1(\mu \otimes \nu, \mathbf{C})$ og

$$x \mapsto \int_F f(x,y)\, \nu(dy) \in L^1(\mu, \mathbf{C}) \quad \text{og} \quad y \mapsto \int_E f(x,y)\, \mu(dx) \in L^1(\nu, \mathbf{C})$$

samt

$$\int_{E \times F} f \, d\mu \otimes \nu = \int_E \{ \int_F f(x,y)\, \nu(dy) \}\, \mu(dx) = \int_E \{ \int_F f(x,y)\, \mu(dx) \}\, \nu(dy).$$

Appendiks C.
I forbindelse med gennemgangen af konvergens i fordeling udnyttede vi flg. egenskab ved den reelle akse.

Enhver åben begrænset delmængde af den reelle akse er en højst tællelig disjunkt forening af åbne intervaller.

'Bevis'. Lad $G \subseteq \mathbf{R}$ betegne en begrænset ikke tom åben delmængde og lad $K > 0$ være valgt så at $G \subseteq [-K, K\,]$. Definer for ethvert $x \in G$

$$x_h := \inf\{y \,|\, y > x,\ y \notin G\} \quad \text{og} \quad x_v := \sup\{y \,|\, y < x,\ y \notin G\}.$$

Da G er åben og begrænset, og $\mathbf{Q}$ er tæt i $\mathbf{R}$ ses let at flg. betingelser er opfyldte for alle $x \in G$:

a) $-K \leq x_v < x < x_h \leq K$. b) $(x_v, x_h) \subseteq G$. c) $\mathbf{Q} \cap (x_v, x_h) \neq \emptyset$.

samt

d) $(x_v, x_h) \cap (\tilde{x}_v, \tilde{x}_h) \neq \emptyset \Rightarrow (x_v, x_h) = (\tilde{x}_v, \tilde{x}_h)$ for vilkårlige $x,\ \tilde{x} \in G$.

Specielt findes der altså for alle $x \in G$ et $\tilde{x} \in \mathbf{Q} \cap G$, så at

$$(x_v, x_h) = (\tilde{x}_v, \tilde{x}_h).$$

Lader vi derfor $(x(n))_{n \geq 1}$ betegne en nummerering af $\mathbf{Q} \cap G$ har vi

$$G = \bigcup_{x \in G} (x_v, x_h) = \bigcup_{n=1}^{\infty} (x(n)_v, x(n)_h),$$

og intervallerne er enten sammenfaldende eller disjunkte. Ved udtynding og overgang til en eventuel delfølge fås derfor, at G kan skrives som en højst tællelig disjunkt forening af åbne intervaller. $\diamond$

Det er på sin plads at bemærke, at resultatet kun gælder i dimension 1. Et tilsvarende resultat i højere dimensioner er flg.

Enhver åben mængde $G \subseteq \mathbf{R}^n$ kan skrives som en højst tællelig disjunkt foreningsmængde af 'halvåbne' kasser, d.v.s. mængder på formen

$$\prod_{i=1}^{n} \,]\, a_i, b_i \,]$$

hvor $-\infty < a_i < b_i < \infty$ for $i = 1, \ldots, n$. Dette resultat udnyttes i beviset for Transformationssætningen, d.v.s. Sætning 13.

Appendiks D.
Fra Mat 11 er det velkendt, at for reelle tal $(a_n)_{n\geq 1}$ og a gælder implikationen

$$a_n \to_{n\to\infty} a \Rightarrow (1+\frac{a_n}{n})^n \to_{n\to\infty} e^a.$$

Men i forbindelse med den centrale grænseværdisætning har vi brugt, at resultatet også gælder for komplekse tal. Et argument herfor går som følger.

Lad $(a_n)_{n\geq 1}$ og a betegne komplekse tal, så at $a_n \to_{n\to\infty} a$. Ifølge definitionen på konvergens af komplekse tal har vi derfor

$$|\,a_n| \to_{n\to\infty} |\,a|,\ \ \Re a_n \to_{n\to\infty} \Re a \ \text{ og } \ \Im a_n \to_{n\to\infty} \Im a.$$

Da $a_n/n \to 0$ gælder yderligere fra et vist trin at regne, at

$$1+\frac{a_n}{n} = |\,1+\frac{a_n}{n}\,| \cdot e^{i\arctan\theta_n} \qquad \text{hvor } \theta_n = \frac{\Im a_n/n}{1+\Re a_n/n},$$

og dermed

$$(1+\frac{a_n}{n})^n = |\,1+\frac{a_n}{n}\,|^n \cdot e^{in\arctan\theta_n}.$$

$x \mapsto \arctan x$ betegner her hoveddeterminationen af $\tan^{-1}$. Da denne er differentiabel i 0 med differentialkvotient 1 og $\theta_n \to_{n\to\infty} 0$ fås derfor, at

$$n\arctan\theta_n = n\cdot\theta_n \cdot \frac{\arctan\theta_n}{\theta_n} = \frac{\Im a_n}{1+\Re a_n/n}\cdot\frac{\arctan\theta_n}{\theta_n} \to_{n\to\infty} \Im a.$$

Tilsvarende fås ved brug af ovenstående reelle udgave, at

$$|\,1+\frac{a_n}{n}\,|^n = ((1+\Re a_n/n)^2 + (\Im a_n/n)^2)^{n/2} = \left(\left(1+\frac{|a_n|^2}{n^2}+\frac{2\Re a_n}{n}\right)^n\right)^{1/2} =$$

$$\left(\left(1+\frac{|a_n|^2/n+2\Re a_n}{n}\right)^n\right)^{1/2} \to_{n\to\infty} \sqrt{e^{2\Re a}} = e^{\Re a}.$$

D.v.s. alt i alt

$$\lim_n (1+\frac{a_n}{n})^n = \lim_n |\,1+\frac{a_n}{n}\,|^n \cdot \lim_n e^{in\arctan\theta_n} = e^{\Re a}\cdot e^{i\Im a} = e^a.$$

Appendiks E. Metriske rum.
Som det er velkendt definerer

$$(x, y) \mapsto |x - y| \quad x,\, y \in \mathbf{R}$$

og

$$(\underline{x}, \underline{y}) \mapsto \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad \underline{x},\, \underline{y} \in \mathbf{R}^n$$

et afstandsbegreb i $\mathbf{R}$ h.h.v. $\mathbf{R}^n$, som er ikke-negativ, symmetrisk og opfylder trekantsuligheden. Disse såkaldte *Euklidiske metrikker* danner baggrund for flg. notation.

Lad S betegne en ikke-tom mængde. En funktion $d : S \times S \to \mathbf{R}_+$ kaldes en *metrik* på S, hvis

1) $d(x, y) = d(y, x) \quad x,\, y \in S$

2) $d(x, y) \leq d(x, z) + d(z, y) \quad x,\, y,\, z \in S$

3) $d(x, y) = 0 \Leftrightarrow x = y \quad x,\, y \in S.$

Hvis 3) erstattes af det svagere krav

$$d(x, x) = 0 \quad x \in S,$$

taler man om en *pseudometrik*.

(S, d) kaldes i givet fald et metrisk h.h.v. et pseudometrisk rum, og værdien $d(x, y)$ omtales som d-afstanden mellem punkterne x og y.

Tilstedeværelsen af en metrik d på S gør det muligt direkte at oversætte en række velkendte afstandsrelaterede begreber fra $\mathbf{R}^n$ så som punktkonvergens, åbne kugler, åbne og lukkede mængder samt kontinuitet. Generalisationerne foregår i h.h.t. flg. definitioner. Læg mærke til at begreberne afhænger eksplicit af den valgte metrik d.

Punktkonvergens

$$x_n \to x \Leftrightarrow d(x_n, x) \to 0.$$

hvor $(x_n)_{n \geq 1}$ og x er punkter i S.

Åbne kugler

$$b(x, r) := \{y \in S \mid d(x, y) < r\}$$

hvor $x \in S$ og $r > 0$ kaldes den *åbne kugle* med centrum x og radius r.

Åbne mængder

$$U \subseteq S \text{ er åben} \Leftrightarrow \forall\, x \in S\ \exists\, r > 0 : \quad b(x, r) \subseteq U.$$

Lukkede mængder

$$F \subseteq S \text{ er lukket } \Leftrightarrow F^c \text{er åben.}$$

Kontinuitet

$$f : S \to \mathbf{R} \text{ er kontinuert } \Leftrightarrow x_n \to x \Rightarrow f(x_n) \to f(x) \text{ for alle } (x_n)_{n\geq 1} \text{ og } x \text{ i } S$$

$$\text{(overvej)} \Leftrightarrow \forall\, x \in S\; \forall\, \epsilon > 0\; \exists\, \delta > 0 :\; d(x,y) < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Med udgangspunkt i velkendte egenskaber ved Euklidiske metrikker indføres yderligere to begreber.

Separabilitet
Et metrisk rum (S, d) siges at være *separabelt*, hvis der findes en følge $(x_n)_{n\geq 1}$ af punkter i S, så at

$$\forall\, x \in S\; \forall\, r > 0\; \exists\, n \geq 1 :\; x_n \in b(x, r).$$

Man udtrykker dette ved at sige, at $(x_n)_{n\geq 1}$ er *tæt* i S.

Fuldstændighed
Et metrisk rum (S, d) siges at være *fuldstændigt*, hvis enhver *Cauchyfølge* er konvergent. En følge $(x_n)_{n\geq 1}$ af punkter i S, siges her at være en Cauchyfølge, hvis

$$\forall\, \epsilon > 0\; \exists\, n \geq 1 :\; m, k \geq n \Rightarrow d(x_m, x_k) < \epsilon.$$

Det er ofte vigtigt at arbejde med forskellige metrikker på en mængde S. I denne forbindelse er flg. begreb åbenlyst vigtigt.

Ækvivalens
To metrikker d_1 og d_2 på S siges at være *ækvivalente* hvis

$$d_1(x_n, x) \to 0 \Leftrightarrow d_2(x_n, x) \to 0$$

for vilkårlige punkter $(x_n)_{n\geq 1}$ og x i S. D.v.s. ækvivalente metrikker har samme konvergente følger.

Da man forholdsvis let kan vise, at ækvivalensbetingelsen kan omformuleres til

$$\forall\, x \in S\; \forall\, r > 0\; \exists\, r_1, r_2 > 0 :\; b_{d_1}(x, r_1) \subseteq b_{d_2}(x, r) \text{ og } b_{d_2}(x, r_2) \subseteq b_{d_1}(x, r),$$

ses, at egenskaberne åbenhed, lukkethed, kontinuitet samt separabilitet bevares ved overgang mellem ækvivalente metrikker. Dette gælder derimod ikke nødvendigvis for fuldstændighed, som flg. eksempel viser. (tænk over dette)

$$S := (0, 1] \text{ samt } d_1(x, y) := |x - y| \text{ og } d_2(x, y) := |1/x - 1/y| \quad x, y \in S.$$

d_1 og d_2 er da ækvivalente metrikker på S, og det metriske rum (S, d_2) er fuldstændigt, men ikke (S, d_1). F.eks. er $(1/n)_{n\geq 1}$ en d_1-Cauchyfølge i S, som ikke konvergerer i (S, d_1).

INDHOLDSFORTEGNELSE

Forelæsningsnoterne skal ses i sammenhæng med Hoffmann's bog, og der er flg. forbindelse mellem bogen og noternes forskellige afsnit.

Målelige rum henter stof fra sektionerne 1.1, 1.2, 1.5, 1.6, 1.11, 1.12 og 1.43.
Bemærkning. Sektion 1.6 indeholder blandt andet Dynkin's Lemma.

Borel σ-algebraer henter stof fra sektionerne 1.8, 1.9, 1.10 og 1.43.
Bemærkning. Sektion 1.9 indeholder blandt andet identiteten $\mathcal{B}(\mathbf{R}^n) = \mathcal{B}(\mathbf{R})^n$, samt at Borel σ-algebraen i et seperabelt metrisk rum er frembragt af mængdesystemet bestående af alle kugler.

Målelige funktioner henter stof fra sektionerne 1.38, 1.39, 1.40, 1.41, 1.42, 1.43, 1.45, 1.46 og 1.48.

Mål henter stof fra sektionerne 1.3, 1.4, 1.7, 1.13, 1.14, 1.15, 1.19, 1.20 og 1.44.

Mål på metriske rum henter stof fra sektionerne 1.34 og 1.37.

Integralet, konstruktion og egenskaber henter stof fra sektionerne 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.15, 3.17 og 3.26.
Bemærkning. Integralet af en ikke-negativ funktion defineres straks i sektion 3.1, hvorefter det vises, at det har de ønskede egenskaber.

Uligheder henter stof fra sektionerne 3.9, 3.10, 3.11 og 3.12.

Produktmål, eksistens og egenskaber henter stof fra sektionerne 3.20, 3.21 og 3.30. Bemærk at noterne bruger Dynkin's Lemma i stedet for sektion 1.49.

Konvergensformer henter stof fra sektionerne 3.13, 3.22, 3.23, 3.24 og 3.25.

L^p**-rum** henter stof fra sektionerne 3.22 og 3.25.

Radon-Nikodym's Sætning henter stof fra sektionerne 3.18 og 3.19. Beviset i sektion 3.18 er helt anderledes end det i noterne.

Konstruktion af mål henter stof fra sektionerne $1.21, 1.22, 1.23, 1.24, 1.25, 1.26$ og 1.27.
Bemærkning. Noterne løser eksistensproblemet ved brug af ydre mål. Den duale konstruktion via indre mål berøres overhovedet ikke.

Lebesgue-Stieltjes mål på R henter stof fra sektionerne 1.32 og 3.27.
Bemærkning. Beviset for eksistens af Lebesgue-Stieltjes mål på **R** er ændret lidt, da noterne, som allerede nævnt, kun behandler målkonstruktion via ydre mål.

Afsnittene **Lebesgue målet på** $\mathbf{R}^n$ og **Transformation af tætheder** har ingen direkte pendant i bogen, men den formulerede transformationssætning er et specialtilfælde af en generel sætning, som findes i bind II.

Sandsynlighedsfelter og stokastiske funktioner henter stof fra sektionerne $2.1, 2.7, 2.8, 2.14, 4.1$ og 4.2.

Uafhængighed henter stof fra sektionerne $2.3, 2.4$ og 2.5.

De store tals love I og II henter stof fra sektionerne $4.7, 4.8, 4.9, 4.11, 4.12, 4.13, 4.31, 4.32, 4.33, 4.34$ og 4.36.

Fordelingskonvergens henter stof fra sektionerne $5.1\, 5.2, 5.3, 5.4$ og 5.9.

Kontinuitetetssætningen for karakteristiske funktioner henter stof fra sektionerne $5.5, 5.6, 5.8$ og 5.10.

Den Centrale Grænseværdisætning henter stof fra sektionerne $5.16, 5.21, 5.22$ og 5.23.

Betingede middelværdier henter stof fra sektionerne $6.1, 6.4, 6.7, 6.8, 6.9, 6.10$ og 6.11.

Martingaler henter stof fra sektionerne $7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 7.10$ $7.12, 7.13, 7.14, 7.16$ og 7.17.

Afsnittene **Momentproblemet** og **Den flerdimensionale normalfordeling** skal ses samen med sektion 4.17 og sektion 4.22.