Konvergens i fordeling

Husk: $\{X_n\}$ av stokastiske variabler konvergerer i fordeling X dersom

$$\lim_{n\to\infty} F_n(x) = F(x)$$

i de punktene x hvor F er kontinuerlig.

Definisjon: En delmengde $D \subseteq \mathbb{R}$ er tett dersom ethvert åpent intervall (a,b) inneholder elementer fra D.

Eksempel: $\mathbb{Q}$

Setning: Anta at $\lim_{n\to\infty} F_n(x) = F(x)$ i en tett mengde av punkter. Da er $\lim_{n\to\infty} F_n(x) = F(x)$ i alle kontinuitetspunkter til F.

Bevis: Anta at x er et kontinuitetspunkt til F og at $\varepsilon > 0$ er et gitt tall. Siden F er kont i x, finnes det punkter a og b i den tette mengden slik $a<x<b$ og at
$F(b) < F(x)+\varepsilon$ og $F(a) > F(x)-\varepsilon$. Da

$$\left.\begin{array}{l} \limsup_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(b) = F(b) < F(x)+\varepsilon \\ \quad\rotatebox{90}{$\le$} \\ \liminf_{n\to\infty} F_n(x) \ge \liminf_{n\to\infty} F_n(a) = F(a) > F(x)-\varepsilon \end{array}\right\} \text{ for alle } \varepsilon$$

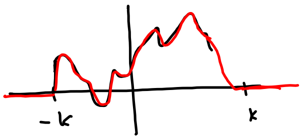
Siden dette gjelder for alle ε, er $\limsup F_n(x) = \liminf F_n(x) = F(x)$, og dermed $\lim F_n(x) = F(x)$.

Teorem: $\{X_n\}$ konvergerer mot X i fordeling hvis og bare hvis

(*) $$E[f(X_n)] \to E[f(X)]$$

for alle begrensede og kontinuerlige funksjoner f.

Bevis: Anta at $X_n \to X$ i fordeling. Skal først vise (*) dersom f har kompakt bærer, dvs at det finnes en k slik at $f(x) = 0$ når $|x| \geq k$.



Siden f er uniformt kontinuerlig finnes det for gitt $\varepsilon > 0$ en funksjon

$$\psi = \sum_{i=1}^{n} a_i \mathbb{1}_{(x_{i-1}, x_i]}$$

slik at $|\psi(x) - f(x)| < \frac{\varepsilon}{3}$ for alle x, og alle x_i er kontinuitetspunkter for F.

$$|E[f(X) - f(X_n)]| \leq \underbrace{|E[f(X) - \psi(X)]|}_{< \frac{\varepsilon}{3}} + |E[\psi(X) - \psi(X_n)]| + \underbrace{|E[\psi(X_n) - f(X_n)]|}_{\varepsilon/3}$$

$$\leq \frac{2}{3}\varepsilon + |E[\psi(X) - \psi(X_n)]|$$

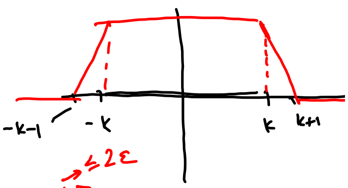
$$\leq \frac{2}{3}\varepsilon + \overbrace{\left|\sum_{i=1}^{n} a_i [F(x_i) - F(x_{i-1})] - \sum_{i=1}^{n} a_i [F_n(x_i) - F_n(x_{i-1})]\right|}^{< \frac{\varepsilon}{3}} < \varepsilon.$$

Siden $\varepsilon > 0$ er vilkårlig, betyr dette at $\lim E[f(X_n)] = E[f(X)]$.

Vi må nå utvide resultatet til å gjelde for alle begrensede, kontinuerlige f. Velg M slik at $|f(x)| \leq M$ for alle x. Gitt $\varepsilon > 0$, velg k så stor at $F(-k) = P[X \leq -k] < \varepsilon$ og $F(k) = P[X \leq k] \geq 1 - \varepsilon$, og slik at k og $-k$ er kontinuitetspunkter for F.

La $g(x) = \begin{cases} 1 & \text{når } |x| \leq k \\ 0 & \text{når } |x| > k+1 \end{cases}$ være kontinuerlig.

Da er $\underbrace{f(x)g(x)}_{\text{kont.}}$ en funksjon med kompakt bærer, og $f(X)g(X) = f(X)$ når $|X| \leq k$.



Dermed har vi

$$|E[f(X_n)] - E[f(X)]| = \overbrace{|E[f(X_n) - f(X_n)g(X_n)]|}^{M P[|X_n| > k] \leq 2\varepsilon}$$

$$+ \underbrace{|E[f(X_n)g(X_n) - f(X)g(X)]|}_{\to 0 \text{ nå det er vist.}} + \underbrace{|E[f(X)g(X) - f(X)]|}_{\leq M \underbrace{P[|X| \geq k]}_{\leq 2\varepsilon}}$$

$$\leq 2\varepsilon M + \to 0 + 2\varepsilon M \leq 5\varepsilon M.$$

Siden $\varepsilon > 0$ er vilkårlig, betyr dette at $E[f(X_n)] \to E[f(X)]$.

Dermed er halve setningen bevist; det gjenstår å vise at $E[f(X_n)] \to E[f(X)]$ medfører konvergens i fordeling.

Oppgave 5.11: $X_1, X_2, \dots$ uavhengige stokastiske variabler, $E[X_n]=0$ og endelig varians. $S_n = X_1 + X_2 + \dots + X_n$.

a) $\varepsilon > 0$ gitt. La

$$\Lambda = \{\omega : \max_{1\le k\le n} |S_k(\omega)| \ge \varepsilon\}$$

$$\Lambda_k = \{\omega : \max_{1\le j\le k-1} |S_j(\omega)| < \varepsilon \text{ og } |S_k(\omega)| \ge \varepsilon\}$$

Vis at Λ_k-ene er disjunkte og at $\bigcup_{k=1}^{n} \Lambda_k = \Lambda$

b) Vis at $S_n - S_k$ er uavhengig av S_k:

$$\underbrace{S_k = X_1 + X_2 + \dots + X_k}_{\sigma(X_1, \dots, X_k)\text{-målbar}}, \quad \underbrace{S_n - S_k = X_{k+1} + X_{k+2} + \dots + X_n}_{\sigma(X_{k+1}, \dots, X_n)\text{-målbar.}}$$

c) Vis at $\int_{\Lambda_k} S_n^2 \, dP \ge \int_{\Lambda_k} S_k^2 \, dP$. Hint: $S_n = S_k + (S_n - S_k)$

$$\int_{\Lambda_k} S_n^2 \, dP = E[\mathbb{1}_{\Lambda_k} S_n^2] = E[\mathbb{1}_{\Lambda_k}(S_k + (S_n - S_k))^2]$$

$$= E[\mathbb{1}_{\Lambda_k}(S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2)]$$

$$= \int_{\Lambda_k} S_k^2 \, dP + 2E[\mathbb{1}_{\Lambda_k} S_k (S_n - S_k)] + \underbrace{E[\mathbb{1}_{\Lambda_k}(S_n - S_k)^2]}_{\ge 0}$$

uavhengige

$$\ge \int_{\Lambda_k} S_k^2 \, dP + 2E[\mathbb{1}_{\Lambda_k} S_k] \underbrace{E[S_n - S_k]}_{E[X_{k+1} + X_{k+2} + \dots + X_n] = 0} + 0 = \int_{\Lambda_k} S_k^2 \, dP.$$

d) Vis at $E[S_n^2] \ge \sum_{k=1}^{n} \int_{\Lambda_k} S_k^2 \, dP \ge \varepsilon^2 P(\Lambda)$

Vi har

$$E[S_n^2] \ge E[\mathbb{1}_\Lambda S_n^2] = E\left[\sum_{k=1}^{n} \mathbb{1}_{\Lambda_k} S_n^2\right]$$

$$\overset{(c)}{\ge} \sum_{k=1}^{n} E[\mathbb{1}_{\Lambda_k} \underbrace{S_k^2}_{\varepsilon^2}] \ge \sum_{k=1}^{n} \varepsilon^2 P(\Lambda_k) = \varepsilon^2 P(\Lambda)$$

e) Vis Kolmogorovs ulikhet:

$$P\left[\max_{1\le k\le n} |S_k| \ge \varepsilon\right] \le \frac{1}{\varepsilon^2} E[S_n^2].$$

Vi har

$$P\Big[\overbrace{\max_{1\le k\le n} |S_k| \ge \varepsilon}^{\Lambda}\Big] = P(\Lambda) \overset{d)}{\le} \frac{1}{\varepsilon^2} E[S_n^2]$$

5.12: $X_1, X_2, \ldots$ uavhengige stok. variable, $E[X_i]=0$, endelig varians.

Vis at hvis $\sum_{n=1}^{\infty} E[X_n^2] < \infty$, så konvergerer $S_n = X_1 + X_2 + \cdots + X_n$ nesten overalt.

Cauchy-følger: $\{x_n\}$ er en Cauchy-følge dersom det for enhver $\varepsilon > 0$ finnes en $N \in \mathbb{N}$ slik at $|x_n - x_N| < \varepsilon$ for alle $n > N$. En følge er konvergent hvis og bare hvis den er en Cauchy-følge.

Dersom $\{S_n\}$ ikke konvergerer n.o.a., så finnes det en $\varepsilon > 0$ en mengde Δ m.d. $P(\Delta) > 0$ slik at ... og alle $n \in \mathbb{N}$, så finnes det en $m > n$ slik at $|S_m - S_n| \geq \varepsilon$.

Dette betyr

$$\Delta \subseteq \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{\omega : \max_{n \leq k \leq m} |S_k - S_n| \geq \varepsilon\}$$

Kolm. ulikhet
av mål.

$$0 < P(\Delta) \leq \lim_{n\to\infty} \lim_{m\to\infty} P\{\omega : \max_{n\leq k\leq m} |S_k - S_n| \geq \varepsilon\} \overset{\lim\lim}{\leq} \frac{1}{\varepsilon^2} E[|S_m - S_n|^2]$$

$$\overset{\lim\lim}{=} \frac{1}{\varepsilon^2} E\Big[\Big(\sum_{j=n}^{m} X_j\Big)^2\Big] \overset{\lim\lim}{=} \frac{1}{\varepsilon^2} \sum_{j=n}^{m} E[X_j^2]$$

$$\overset{\lim\lim}{\leq} \underbrace{\frac{1}{\varepsilon^2} \sum_{j=n}^{\infty} E[X_j^2]}_{\sum_{j=1}^{\infty} E[X_j^2] < \infty} \to 0$$

Selvmotsigelse, så S_n konvergerer nesten overalt.