Borel-Cantellis sats och stora talens lag

Gunnar Englund
Matematisk statistik
KTH

Vt 2005

1 Inledning

Borel-Cantellis sats är en intressant och användbar sats framför allt för att bevisa stora talens lag i stark form.
Vi betraktar en följd händelser $A_1, A_2, A_3, \ldots$ och är intresserade av frågan om oändligt många av dessa inträffar eller om möjligen bara ett ändligt antal av dem inträffar. Vi bildar

$$F_n = \bigcup_{k=n}^{\infty} A_k \text{ och } G_n = \bigcap_{k=n}^{\infty} A_k.$$

Att G_n inträffar innebär att alla A_k för $k \geq n$ inträffar. Om det finns något sådant n innebär det alltså att från och med detta n inträffar alla A_k för $k \geq n$. Med

$$H = \bigcup_{n=1}^{\infty} G_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

innebär detta att om H inträffar finns ett n så att alla A_k med $k \geq n$ inträffar. Ibland betecknas H med $\liminf A_k$.

Att F_n inträffar innebär att det finns något A_k för $k \geq n$ som inträffar. Om F_n inträffar för alla n innebär detta att oändligt många av A_k:na inträffar. Vi bildar därför

$$E = \bigcap_{n=1}^{\infty} F_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Om E inträffar, så inträffar alltså oändligt många av A_k:na. Ibland skrives detta lite behändigt som $E = \{A_n \text{ i.o.}\}$ där i.o. står för "infinitely often", dvs oändligt många gånger. E betecknas ibland med $\limsup A_k$.

Man noterar att F_n avtar i n och att alltså

$$P(E) = P(\bigcap_{n=1}^{\infty} F_n) = \lim_{n\to\infty} P(F_n) = \lim_{n\to\infty} P(\bigcup_{k=n}^{\infty} A_k).$$

Vi har dock enligt Booles sats att

$$P(\bigcup_{k=n}^{\infty} A_k) \leq \sum_{k=n}^{\infty} P(A_k)$$

och denna serie $\to 0$ då $n \to \infty$ om serien $\sum_1^\infty P(A_k)$ konvergerar. Detta innebär att vi visat följande sats som kallas Borel-Cantellis sats.

Sats 1 *Borel-Cantellis sats*
Om $\sum_{n=1}^{\infty} P(A_n) < \infty$ så gäller att $P(E) = P(A_n \text{ i.o}) = 0$, dvs att med sannolikhet 1 inträffar bara ändligt många A_n. $\square$

Man kan notera att det i satsen inte krävs någon form av oberoende utan satsen gäller helt generellt.

Det finns en omvändning av Borel-Cantellis sats om man antar att händelserna $A_1, A_2, \ldots$ är oberoende.

Sats 2 *Omvändning till Borel-Cantellis sats*
Om $A_1, A_2, \ldots$ är oberoende och

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

så gäller att $P(E) = P(A_n \text{ i.o}) = 1$, dvs att det med sannolikhet 1 inträffar oändligt många A_n. $\square$

Bevis: Vi har

$$P(\bigcap_{k=n}^{\infty} A_k^*) = \prod_{k=n}^{\infty} P(A_k^*) = \prod_{k=n}^{\infty} (1 - P(A_k)).$$

Eftersom $1 - x \leq e^{-x}$ erhålls $1 - P(A_k) \leq e^{-P(A_k)}$ och vi får

$$P(\bigcap_{k=n}^{\infty} A_k^*) \leq \exp(-\sum_{k=n}^{\infty} P(A_k)).$$

Om nu $\sum_{n=1}^{\infty} P(A_n) = \infty$ så gäller att serien i exponenten divergerar och man får

$$P(\bigcap_{k=n}^{\infty} A_k^*) = 0.$$

Alltså gäller även

$$P(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^*) = 0.$$

som innebär att

$$P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = 1 - P(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^*) = 1 - 0 = 1$$

dvs att oändligt många A_k:n inträffar med sannolikhet 1. $\square$

2 Några exempel på tillämpningar

Exempel 2.1 Låt $X_1, X_2, X_3 \dots$ vara oberoende likafördelade med kontinuerlig fördelning. Vi låter

$$U_n = \begin{cases} 1 & \text{om } X_n > X_j \text{ för } j = 1, 2, \dots n-1 \\ 0 & \text{annars.} \end{cases}$$

Detta betyder att $U_n = 1$ om X_n utgör ett ”rekord”, dvs är det hitills största värdet. Vi låter $A_n = \{U_n = 1\}$.
Man inser att $P(U_n = 1) = 1/n$ eftersom sannolikheten att det största av n värden skall komma i omgång n är $1/n$ av symmetriskäl. Vidare är $A_1, A_2, \dots$ oberoende. Vi har ju

$$P(A_m \cap A_{m+1} \cap \dots A_{m+k}) = P(A_m | A_{m+1} \cap \dots A_{m+k}) P(A_{m+1} \cap \dots A_{m+k})$$

och A_m och $A_{m+1} \cap \dots A_{m+k}$ är naturligtvis oberoende eftersom A_m bara berör storleksförhållandena bland de första m av X-variablerna.

Vi får då eftersom

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

att $P(A_n \text{ i.o.}) = 1$ dvs oändligt många A_n inträffar. Vi får oändligt många rekord – ett resultat som kanske (?) är självklart. Vidare får vi

$$E\Big(\sum_{n=1}^{\infty} U_n U_{n+1}\Big) = \sum_{n=1}^{\infty} E(U_n U_{n+1}) = \sum_{n=1}^{\infty} P(A_n) P(A_{n+1}) = \sum_{n=1}^{\infty} \frac{1}{n(n+1)} < \infty$$

och alltså är $\sum_1^{\infty} U_n U_{n+1}$ ändlig med sannolikhet 1. Slutsatsen är att det bara inträffar ett ändligt antal ”dubbelrekord”, dvs rekord två gånger i rad – ett resultat som på intet sätt är trivialt.

Exempel 2.2 Låt $X_1, X_2, \dots$ vara oberoende och likafördelade. Då gäller att

$$\begin{aligned} E(|X_1|) = \int_0^{\infty} P(|X_1| > x)dx &= \sum_{n=0}^{\infty} \int_n^{n+1} P(|X_1| > x)dx \\ &\leq \sum_{n=0}^{\infty} P(|X_1| > n) \leq 1 + \sum_{n=1}^{\infty} P(|X_n| > n), \end{aligned}$$

men också

$$E(|X_1|) = \sum_{n=0}^{\infty} \int_n^{n+1} P(|X_1| > x)dx \geq \sum_{n=0}^{\infty} P(|X_1| > n+1) = \sum_{n=1}^{\infty} P(|X_n| > n).$$

Om nu $E(|X_1|) < \infty$ ser vi att $\sum_{n=1}^{\infty} P(|X_n| > n) < \infty$ som enligt Borel-Cantellis sats medför $P(|X_n| > n \text{ i.o.}) = 0$. Å andra sidan, om $E(|X_1|) = \infty$ och alltså $\sum_{n=0}^{\infty} P(|X_n| > n) = \infty$, så ger omvändningen till Borel-Cantellis sats att $P(|X_n| > n \text{ i.o.}) = 1$. Om $E(|X_k|)$ är ändligt (respektive oändligt) kommer $|X_n|$ att bli större än n oändligt många gånger med sannolikhet 0 (respektive 1).

3 Bevis av stora talens lag i stark form

Vi låter $X_1, X_2, \ldots$ vara oberoende likafördelade med $E(X_i) = m$ och $V(X_i) = \sigma^2 < \infty$ och definierar $S_n = X_1 + X_2 + \cdots + X_n$. Vi är intresserade att visa att med sannolikhet 1 gäller att $S_n/n \to m$ då $n \to \infty$. Detta betyder alltså att vi vill visa att

$$P(\lim_{n\to\infty} \frac{S_n}{n} = m) = 1,$$

dvs att det finns en mängd Ω_0 med $P(\Omega_0) = 1$ där för varje $\omega \in \Omega_0$ gäller att

$$\lim_{n\to\infty} |\frac{S_n}{n} - m| = 0.$$

Vi behöver alltså visa att för varje $\omega \in \Omega_0$ och för varje $\varepsilon > 0$ finns ett $N(\omega, \varepsilon)$ så att om $n \geq N(\omega, \varepsilon)$ gäller att $|S_n/n - m| \leq \varepsilon$. Det räcker att visa att

$$\lim_{N\to\infty} P(|\frac{S_n}{n} - m| > \varepsilon \text{ något } n \geq N) = 0.$$

Notera skillnaden mot stora talens lag i svag form, som säger att för alla $\varepsilon > 0$

$$P(|\frac{S_n}{n} - m| > \varepsilon) \to 0 \text{ då } n \to \infty.$$

I stora talens lag i stark form måste $|S_n/n - m|$ vara litet för alla tillräckligt stora n för alla $\omega \in \Omega_0$ där $P(\Omega_0) = 1$. Vid slantsingling kan vi koda krona och klave som 0 respektive 1, och kan identifiera ett ω med ett på måfå valt tal på intervallet $[0, 1]$ där binärbråksutvecklingen ger sekvensen. Vad stora talens lag då säger är att vi med sannolikhet 1 får ett tal sådant att andelen 1:or i sekvensen konvergerar mot 1/2. Det kan finnas ”undantags”-ω - t ex är vid slantsingling sekvensen $000\ldots$ möjlig, men sådana undantagssekvenser har sammanlagt sannolikhet 0.
Utan inskränkning kan vi anta att $E(X_i) = m = 0$ eftersom vi annars kan betrakta $X_i - m$. Vi har $V(S_n) = n\sigma^2$. Enligt Tjebyshovs olikhet gäller

$$P(|S_n| > n\varepsilon) \leq \frac{V(S_n)}{(n\varepsilon)^2} = \frac{n\sigma^2}{(n\varepsilon)^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Tyvärr divergerar den harmoniska serien $\sum_1^\infty 1/n$ så vi kan inte använda Borel-Cantellis sats direkt. Dock gäller $\sum_1^\infty 1/n^2 < \infty$ och detta betyder att vi kan använda satsen för n^2, $n = 1, 2, \ldots$. Vi har

$$P(|S_{n^2}| > n^2\varepsilon) \leq \frac{\sigma^2}{n^2\varepsilon^2}.$$

Alltså gäller enligt Borel-Cantellis sats att $P(|\frac{S_{n^2}}{n^2}| > \varepsilon \text{ i.o.}) = 0$ som visar att (med sannolikhet 1) $S_n^2/n^2 \to 0$. Vi har alltså lyckats visa att för delsekvensen $n^2,\ n = 1, 2, \ldots$ så har vi konvergens med sannolikhet 1. Återstår att reda ut vad som kan hända mellan dessa n^2. Vi definierar därför

$$D_n = \max_{n^2 \leq k < (n+1)^2} |S_k - S_{n^2}|,$$

dvs den största avvikelsen från S_{n^2} som kan inträffa mellan n^2 och $(n+1)^2$. Vi får

$$D_n^2 = \max_{n^2 \le k < (n+1)^2} (S_k - S_{n^2})^2 \le \sum_{k=n^2}^{(n+1)^2-1} (S_k - S_{n^2})^2,$$

där vi gjort den grova uppskattningen att $\max(|x|, |y|) \le (|x| + |y|)$. Detta ger

$$E(D_n^2) \le \sum_{k=n^2}^{(n+1)^2-1} E((S_k - S_{n^2})^2).$$

Men $E((S_k - S_{n^2})^2) = (k - n^2)\sigma^2 \le 2n\sigma^2$ då $n^2 \le k < (n+1)^2$ och det är $2n$ termer i summan och det ger

$$E(D_n^2) \le (2n)(2n)\sigma^2 = 4n^2\sigma^2.$$

Med Tjebyshovs olikhet ger detta

$$P(D_n > n^2\varepsilon) \le \frac{4n^2\sigma^2}{(n^2\varepsilon)^2} = \frac{4\sigma^2}{n^2\varepsilon^2}.$$

Alltså gäller $D_n/n^2 \to 0$ med sannolikhet 1. Till slut ger detta för k mellan n^2 och $(n+1)^2$

$$|\frac{S_k}{k}| \le \frac{|S_{n^2}| + D_n}{k} \le \frac{|S_{n^2}| + D_n}{n^2} \to 0.$$

Detta innebär att vi lyckats visa att $S_n/n \to 0$ med sannolikhet 1. Vi har gjort detta under tilläggsvillkoret att $V(X_i) = \sigma^2$, men med stor möda kan man visa att detta tilläggsvillkor ej är nödvändigt.