Erling Arnold Tønseth Svela

Den Cramér-stokastiske modellen for primtala

Bacheloroppgåve i matematiske fag
Veileder: Kristian Seip
Mai 2020

Bacheloroppgåve

NTNU
Noregs teknisk-naturvitskaplege universitet
Fakultet for informasjonsteknologi og elektroteknikk
Institutt for matematiske fag

NTNU
Kunnskap for en bedre verden

Erling Arnold Tønseth Svela

Den Cramér-stokastiske modellen for primtala

Bacheloroppgåve i matematiske fag
Veileder: Kristian Seip
Mai 2020

Noregs teknisk-naturvitskaplege universitet
Fakultet for informasjonsteknologi og elektroteknikk
Institutt for matematiske fag

NTNU
Kunnskap for ei betre verd

Samandrag

Målet med denne oppgåva er å introdusera Cramér-modellen, og å drøfta om den modellerer primtala på ein rimeleg måte. Motivasjonen bak modellen vert forklart og modellen vert definert. Somme av krava modellar av primtala burde ha vert drøfta, og det vert vist at Cramér-modellen oppfyller desse krava. Vidare nytter vi Cramér-modellen til å utleia tre formodingar som betrar allereie eksisterande resultat. Til slutt vert problem og avvik i Cramér-modellen diskutert, og det vert drøfta om det er mogleg å gjera modellen betre.

Innhald

Forord

Denne oppgåva er skriven under rettleiing av Kristian Seip. Eg takkar han for å ha føreslått ei interessant oppgåve, glimrande rettleiing under skriveprosessen, samt å ha planta den første interessa mi for analyse då eg byrja å studere for tre år sidan.
Ein platonisk takk går til Andrew Granville, som skreiv den strålande oversiktsartikkelen "Harald Cramér and the distribution of prime numbers" som eg har hatt god bruk for i arbeidet mitt. Vidare ønsker eg å takka studievenane mine for lærerike faglege diskusjonar gjennom heile studieløpet. I lys av verdssituasjonen i skrivande stund ønsker eg til slutt også å takka familien min for å ha vist stor forståing når eg har isolert meg sjølv på rommet for å gjera matematikk.

Erling A.T. Svela, april 2020.

1 Innleiing

Harald Cramér vert rekna som den første til å nytta element frå sannsynsteorien i analytisk talteori. I 1920 gjekk han statistisk til verks for å vise at, for ein vilkårleg $\epsilon > 0$, har vi

$$p_{n+1} - p_n = O(p_n^\epsilon)$$

for alle primtal $p_n \leq x$, med unntak av maksimalt $x^{1-3\epsilon/2}$ av desse. Denne tilnærminga endra delar av tenkemåten i fagfeltet. Det var naudsynt både å forstå den gjennomsnittlege eller "vanlege"oppførselen til primtala, og oppførselen i ekstreme tilfelle. Denne typen statistisk tilnærming vart seinare teken opp og betra, mellom andre av Selberg.
I denne oppgåva skal vi, med utgangspunkt i primtalssatsen, laga ein stokastisk modell for primtala. Denne modellen vil vera mykje enklare å arbeida med enn dei faktiske primtala, og vi skal nytta dette til å koma fram til tre formodingar om primtala. Rimelegheita til desse formodingane er avhengig av om den stokastiske modellen vår er ei rimeleg etterlikning av primtala, og dette vil bli diskutert.

1.1 Notasjon

Gjennom heile oppgåva vil p_n vera det n-te primtalet, medan $\mathcal{P}$ er mengda av alle primtal. Vi seier at ei følgje ξ_n konvergerer mot ξ med sannsyn 1 dersom $\text{Prob}(\xi_n \to \xi) = 1$. Vi skriv dette som $\xi_n \to \xi$ *a.s.* For to funksjonar f og g vil følgjande notasjon bli nytta for å skildra storleiksordenen til f i forhold til g:

- $f(x) = O(g(x))$ dersom det eksisterer $C > 0$ og $\tilde{x}$ slik at

$$|f(x)| \leq Cg(x) \qquad \text{for alle } x \geq \tilde{x},$$

- $f(x) \sim g(x)$ dersom

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1,$$

- $f(x) = o(g(x))$ dersom

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0,$$

- $f(x) = \Omega(g(x))$ dersom

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \infty.$$

2 Nokre innleiande resultat om primtal

For å kunne studere fordelinga til primtala vil det vera nyttig for oss å introdusera funksjonen $\pi(x)$.

Definisjon 1 (Teljefunksjonen for primtal)**.** Funksjonen $\pi(x) : \mathbb{R} \to \mathbb{N}$ er definert ved

$$\pi(x) = \{\text{Antal primtal mindre enn eller lik } x\} = \sum_{p \leq x} 1. \tag{1}$$

På grunn av den tilsynelatande sporadiske fordelinga av primtal, kan det sjå ut som at å finna noko mønster er ei umogleg oppgåve. Dette er ei sanning med modifikasjonar. Mellom anna veit vi mykje om dei asymptotiske eigenskapane til $\pi(x)$.

Teorem 1 (Primtalssatsen)**.**

$$\pi(x) \sim \frac{x}{\log(x)}.$$

Vi veit altså at når x er veldig stor, kan $\pi(x)$ tilnærmast ved hjelp av $\frac{x}{\log(x)}$. Dette er ei tilstrekkeleg tilnærming, men $\pi(x)$ kan tilnærmast enno betre av den logaritmiske integralfunksjonen.

Definisjon 2. For x i intervallet $[2,\infty)$ definerer vi funksjonen

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log(t)}.$$

Vi kallar $\mathrm{Li}(x)$ for den logaritmiske integralfunksjonen.

Det er klart at $\mathrm{Li}(x)$ har dei same asymptotiske eigenskapane som $\frac{x}{\log(x)}$.

Lemma 1.

$$\mathrm{Li}(x) \sim \pi(x).$$

Prov. På grunn av primtalssatsen er det nok å visa at $\mathrm{Li}(x) \sim \frac{x}{\log(x)}$. Delvis integrasjon med $u = \frac{1}{\log(t)}$ og $v' = 1$ gjev

$$\int_2^x \frac{dt}{\log(t)} = \left[\frac{t}{\log(t)}\right]_2^x + \int_2^x \frac{dt}{\log(t)^2} = \frac{x}{\log(x)} - \frac{2}{\log(2)} + \int_2^x \frac{dt}{\log(t)^2}.$$

Det siste leddet i uttrykket kan estimerast.

$$\int_2^x \frac{dt}{\log(t)^2} = \int_2^{\sqrt{x}} \frac{dt}{\log(t)^2} + \int_{\sqrt{x}}^x \frac{dt}{\log(t)^2} \leq \frac{\sqrt{x}-2}{\log(2)^2} + \frac{x-\sqrt{x}}{\log(\sqrt{x})^2}.$$

Vi viser så at estimatet har orden $o\left(\frac{x}{\log(x)}\right)$.

$$\lim_{x\to\infty} \frac{\frac{\sqrt{x}-2}{\log(2)^2} + \frac{x-\sqrt{x}}{\log(\sqrt{x})^2}}{\frac{x}{\log(x)}} = \lim_{x\to\infty} \left(\frac{(\sqrt{x}-2)\log(x)}{x\log(2)^2} + \frac{(x-\sqrt{x})\log(x)}{x\log(\sqrt{x})^2}\right)$$

$$\lim_{x\to\infty}\left(\frac{\sqrt{x}\log(x)}{x\log(2)^2} - \frac{2\log(x)}{x\log(2)^2} + \frac{x\log(x)}{\frac{x}{4}\log(x)^2} - \frac{\sqrt{x}\log(x)}{\frac{x}{4}\log(x)^2}\right)$$

$$= \lim_{x\to\infty}\left(\frac{\log(x)}{\sqrt{x}\log(2)^2} - \frac{2\log(x)}{x\log(2)^2} + \frac{4}{\log(x)} - \frac{4}{\sqrt{x}\log(x)}\right)$$

$$= 0 \implies \int_2^x \frac{dt}{\log(t)^2} = o\left(\frac{x}{\log(x)}\right).$$

Då får vi, sidan det midtarste leddet er konstant, og dermed også har orden $o\left(\frac{x}{\log(x)}\right)$,

$$\lim_{x\to\infty} \frac{\mathrm{Li}(x)}{\frac{x}{\log(x)}} = \lim_{x\to\infty} \frac{\frac{x}{\log(x)} + o\left(\frac{x}{\log(x)}\right)}{\frac{x}{\log(x)}}$$

$$= \lim_{x\to\infty} \frac{\frac{x}{\log(x)}}{\frac{x}{\log(x)}} = 1 \implies \mathrm{Li}(x) \sim \frac{x}{\log(x)}.$$

Dette impliserer at $\pi(x) \sim \mathrm{Li}(x)$. □

Grunnen til at vi ønsker å nytte $\mathrm{Li}(x)$ i staden for $\frac{x}{\log(x)}$ er fordi feilleddet mellom $\pi(x)$ og $\mathrm{Li}(x)$ vert minimalt, som forklart i følgjande teorem.

Teorem 2. *Anta Riemann-hypotesen (formoding 4 i appendikset). Då har vi*

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x}\log(x)).$$

Prov. Vi lét $\psi(x)$ vera definert som i definisjon 10 i appendikset. Dersom formoding 4 gjeld, får vi

$$\psi(x) = x - \sum_{\zeta(\rho)=0} \frac{x^\rho}{\rho} = x - \sum_{\zeta(\frac{1}{2}+i\gamma)=0} \frac{x^{\frac{1}{2}+i\gamma}}{\frac{1}{2}+i\gamma}.$$

Vi estimerer summen over nullpunkta.

$$\left| \sum_{\zeta(\frac{1}{2}+i\gamma)=0} \frac{x^{\frac{1}{2}+i\gamma}}{\frac{1}{2}+i\gamma} \right| \leq \sum_{\zeta(\frac{1}{2}+i\gamma)=0} \left| \frac{x^{\frac{1}{2}+i\gamma}}{\frac{1}{2}+i\gamma} \right| = \sum_{\zeta(\frac{1}{2}+i\gamma)=0} \frac{\sqrt{x}}{\sqrt{\frac{1}{4}+\gamma^2}}$$

$$\leq C\sqrt{x} \sum_{0\leq\gamma\leq\sqrt{x}} \frac{1}{\gamma},$$

der C er ein konstant. Ved lemma 7 har vi vidare

$$\leq C\sqrt{x} * O(\log\left(\sqrt{x}\right)^2) = C\sqrt{x} * O(\log(x)^2) = O(\sqrt{x}\log(x)^2)$$

$$\implies \psi(x) = x + O(\sqrt{x}\log(x)^2). \tag{2}$$

Vi definerer så funksjonen $\pi_1(x)$ ved

$$\pi_1(x) = \sum_{p^m\leq x} \frac{\log(p)}{m\log(p)}.$$

Som vist i [5] gjeld relasjonane

$$\pi_1(x) = \pi(x) + \frac{1}{2}\pi(x^{\frac{1}{2}}) + \frac{1}{3}\pi(x^{\frac{1}{3}}) + ...$$

og

$$\pi_1(x) = \int_2^x \frac{\psi(t)dt}{t\log(t)^2} + \frac{\psi(x)}{\log(x)}.$$

Dersom vi nytter likning (2) på den andre relasjonen får vi

$$\pi_1(x) = \int_2^x \left(\frac{t}{t\log(t)^2} + \frac{O(\sqrt{t}\log(t)^2)}{t\log(t)^2} \right) dt + \frac{x}{\log(x)} + \frac{O(\sqrt{x}\log(x)^2)}{\log(x)}$$

$$\int_2^x \frac{dt}{\log(t)^2} + O(\sqrt{x}) + \frac{x}{\log(x)} + O(\sqrt{x}\log(x))$$

$$= \frac{x}{\log(x)} + \int_2^x \frac{dt}{\log(t)^2} + O(\sqrt{x}\log(x)).$$

Frå provet for lemma 1 kjenner vi att dei to første ledda som $\mathrm{Li}(x) + O(1)$, så vi får at

$$\pi_1(x) = \mathrm{Li}(x) + O(\sqrt{x}\log(x)).$$

Vi veit at $\pi(x) \leq x$ for alle x, så vi får

$$\frac{1}{2}\pi(x^{\frac{1}{2}}) + \frac{1}{3}\pi(x^{\frac{1}{3}}) + ... = O(x^{\frac{1}{2}}) = O(\sqrt{x}),$$

som betyr at

$$\begin{aligned}\pi(x) = \pi_1(x) + O(\sqrt{x}) &= \mathrm{Li}(x) + O(\sqrt{x}\log(x)) + O(\sqrt{x}) \\ &= \mathrm{Li}(x) + O(\sqrt{x}\log(x)).\end{aligned}$$

□

Sjølv om vi veit at $\pi(x) \sim \mathrm{Li}(x)$, er vi langt i frå å kunne alt om verken $\pi(x)$ eller om primtala. Vidare i oppgåva skal vi bruke desse innleiande resultata til å utvikla ein probabilistisk modell som etterlikner primtala. Denne modellen vil forhåpentleg vis hjelpa oss å forstå primtala betre.

3 Den Cramér-stokastiske modellen

Det følgjande korollaret er ein direkte konsekvens av primtalssatsen.

Korollar 1. *Gitt eit heiltal x er tettleiken av primtal rundt x asymptotisk lik $\frac{1}{\log(x)}$.*

Prov. Betrakt $\pi(x+1) - \pi(x)$. Ved primtalssatsen får vi

$$\pi(x+1) - \pi(x) \sim \frac{x+1}{\log(x+1)} - \frac{x}{\log(x)} \sim \frac{x+1}{\log(x)} - \frac{x}{\log(x)} = \frac{1}{\log(x)},$$

der den andre relasjonen følgjer av at vi lét x gå mot uendeleg. □

Ein kan tolka dette som eit utsegn om sannsyn. Gitt eit naturleg tal x er sannsynet for at x er eit primtal $\frac{1}{\log(x)}$. Dette gjev grunnlaget for den Cramér-stokastiske modellen.

Definisjon 3 (Den Cramér-stokastiske modellen)**.** For $n = 3, 4, ...$ definerer vi den diskrete stokastiske variabelen z_n ved

$$\mathrm{Prob}(z_n = 1) = \frac{1}{\log(n)}, \qquad \mathrm{Prob}(z_n = 0) = 1 - \frac{1}{\log(n)} \tag{3}$$

der alle z_n er parvis uavhengige. Dette er den Cramér-stokastiske modellen for primtala.
Definer vidare følgja (P_ν) ved

$$P_{\nu+1} = m \qquad \text{dersom} \sum_{n=3}^{m} z_n = \nu \text{ og } \sum_{n=3}^{m-1} z_n = \nu - 1 \tag{4}$$

og funksjonen $\Pi(x)$ ved

$$\Pi(x) = \sum_{n=3}^{x} z_n. \tag{5}$$

Følgja (P_ν) vert kalla følgja av probabilistiske primtal, og $\Pi(x)$ vert kalla den probabilistiske teljefunksjonen for primtal.

Merk. Av tekniske årsaker har vi valt å starta indekseringa på $n = 3$. Dette betyr at 2 ikkje er ein del av modellen vår, og dermed strengt tatt ikkje har moglegheit for å vera eit primtal. Sidan vi i hovudsak er opptatt av store primtal er ikkje dette noko stort problem, men ein kan lett løysa dette ved å introdusera ein ny uavhengig variabel, z_2, og setja den konstant lik 1.

Valet av notasjon er ikkje tilfeldig. Her speler variablane $(z_n)_{n\geq 3}$ rolla til dei naturlege tala, medan følgja $(P_\nu)_{\nu\geq 2}$ av probabilistiske primtal svarer til følgja (p_n) av primtal. Dette stemmer overeins med primtalssatsen, sidan gitt eit naturleg tal n er sannsynet for at det er eit primtal, altså $\text{Prob}(z_n = 1)$, lik $\frac{1}{\log(n)}$. Tilsvarande svarer $\Pi(x)$ til funksjonen $\pi(x)$ for dei naturlege tala, som tel primtala opp til x.

Sidan (P_ν) og $\Pi(x)$ er summar av dei stokastiske variablane (z_n) vil også (P_ν) og $\Pi(x)$ vera stokastiske variablar. Det er ein ekvivalent definisjon av $\Pi(x)$ som speler meir direkte på samsvaret mellom $\pi(x)$ og $\Pi(x)$:

$$\Pi(x) = \{P_n \mid P_n \leq x\} = \sum_{P_n \leq x} 1. \tag{6}$$

Gitt ei binær følgje $(e_n)_{n=3}^{\infty}$ av utfall i variablane $(z_n)_{n=3}^{\infty}$ vil (P_ν) vera ei monoton følgje. Utfallsrommet for modellen vår vil derfor vera mengda $C = \{(a_n)_{n=3}^{\infty} \mid a_i \leq a_{i+i}\ \forall i \in \mathbb{N}\}$ av monotone følgjer. Då vil følgja (p_n) av primtala vera eit element i C, og det same vil mange andre følgjer som i større eller mindre grad etterlikner følgja av primtal. Denne modellen kan sjølvsagt ikkje prova resultat om primtala, men dersom vi antar at primtala er modellerte på ein tilstrekkeleg måte kan vi koma fram til fleire interessante formodingar om dei. Historisk sett har stokastiske modellar, mellom anna Cramér-modellen, vore nytta til å presentera formodingar om primtala, som seinare har vorte prova med hardare analytiske metodar.

Vi ser med ei gong at Cramér-modellen etterlikner dei mest grunnleggjande eigenskapane til primtala.

Teorem 3. *Vi lét A vera hendinga*

$$\{\textit{Det er uendeleg mange probabilistiske primtal}\}.$$

Då har vi

$$\text{Prob}(A) = 1.$$

Prov. Dersom det er endeleg mange probabilistiske primtal eksisterer det ein ñ slik at $z_n = 0$ for $n \geq \tilde{n}$. Kall denne hendinga for E. Sidan variablane (z_n) er uavhengige får vi

$$\begin{aligned}\mathrm{Prob}(E) &= \prod_{n=3}^{\tilde{n}} \mathrm{Prob}(z_n \in \{0,1\}) * \prod_{n=\tilde{n}+1}^{\infty} \mathrm{Prob}(z_n = 0) \\ &= 1 * \prod_{n=\tilde{n}+1}^{\infty} 1 - \frac{1}{\log(n)} = \prod_{n=1}^{\infty} 1 - \frac{1}{\log(\tilde{n}+n)} \\ &\sim \prod_{n=1}^{\infty} 1 - \frac{1}{\log(\tilde{n})} = \prod_{n=1}^{\infty} k = \lim_{n\to\infty} k^n = 0\end{aligned}$$

Sidan $k = 1 - \frac{1}{\log(\tilde{n})} < 1$.
$\mathrm{Prob}(E) = 0$, så då er sannsynet for at det er endeleg mange probabilistiske primtal 0. Då må sannsynet for den komplementære hendinga, nemleg A, vera 1. $\square$

Ved primtalssatsen er $\pi(x)$ og $\mathrm{Li}(x)$ asymptotisk like. Vi har eit tilsvarande resultat for $\Pi(x)$:

Teorem 4.

$$\mathrm{E}(\Pi(x)) \sim \mathrm{Li}(x)$$

Prov. z_m er binomisk fordelt med $n = 1$ og $p = \frac{1}{\log(m)}$, så $\mathrm{E}(z_m) = np = 1 * \frac{1}{\log(m)} = \frac{1}{\log(m)}$. Sidan forventingsverdi er additivt får vi

$$\mathrm{E}(\Pi(x)) = \mathrm{E}\left(\sum_{n=3}^{x} z_n\right) = \sum_{n=3}^{x} \mathrm{E}(z_n) = \sum_{n=3}^{x} \frac{1}{\log(n)},$$

som er asymptotisk lik $\mathrm{Li}(x)$ når vi lét $x \to \infty$. $\square$

Vidare i oppgåva skal vi ta for oss ein del interessante resultat som Cramérmodellen føreslår, samt å sjå på avvik mellom primtala og dei probabilistiske primtala.

4 Storleiksordenen til $|\Pi(x) - \mathrm{Li}(x)|$

Vi veit at $\pi(x) = \mathrm{Li}(x) + O(f(x))$, men her er det mykje rom for variasjon. Det er vist at

$$f(x) = \Omega\left(\frac{\sqrt{x}}{\log(x)}\right) \tag{7}$$

og samstundes, dersom ein antar Riemann-hypotesen

$$f(x) = O(\sqrt{x} * \log(x)). \tag{8}$$

Dette er framleis meir variasjon enn det ein skulle ønske. I denne delen skal vi gje eit forslag til ei strengare avgrensing av $|\pi(x) - \mathrm{Li}(x)|$ ved å nytta Cramérmodellen.
Målet vårt er å visa følgjande resultat om $\Pi(x)$:

Teorem 5.

$$\limsup_{x\to\infty} \frac{|\Pi(x) - \mathrm{Li}(x)|}{\sqrt{2x} * \sqrt{\frac{\log\log(x)}{\log(x)}}} = 1 \ a.s$$

Vi skal visa dette ved å nytta følgjande resultat frå sannsynsteorien:

Teorem 6 (Den generelle lova om den itererte logaritmen)**.** *Vi lét $(x_n)_{n\in\mathbb{N}}$ vera ei følgje av uavhengige stokastiske variablar slik at*

$$\mathrm{E}(x_n) = 0, \qquad \mathrm{E}(x_n^2) = \sigma_n^2, \qquad \mathrm{E}|x_n|^3 = \mu_n,$$

for alle $n \in \mathbb{N}$. Definer

$$S_x = \sum_{n\le x} x_n, \qquad B_x^2 = \sum_{n\le x} \sigma_n^2, \qquad B_x = \sqrt{B_x^2}, \qquad M_x = \sum_{n\le x} \mu_n.$$

Dersom $\lim_{x\to\infty} B_x = \infty$ og $\frac{M_x}{B_x^3} \le \frac{c}{\log(B_x)}$ for ein $c \le \infty$ har vi

$$\limsup_{x\to\infty} \frac{|S_x|}{B_x * \sqrt{2\log\log(B_x)}} = 1 \ a.s$$

Prov. Sjå Borovkov [2]. □

Teorem 5 kan no visast ved å nytta den generelle lova om den itererte logaritmen.

Prov av teorem 5. Vi lét (x_n) vera følgja $(z_n - \frac{1}{\log(n)})$ der (z_n) er dei vanlige variablane i Cramér-modellen. Det følgjer klart at $\mathrm{E}(x_n) = 0$, $\mathrm{E}(x_n^2) = \mathrm{Var}(z_n) = \frac{1}{\log(n)} - \frac{1}{\log(n)^2}$ og $\mathrm{E}(x_n^3) = \frac{1}{\log(n)} - \frac{3}{\log(n)^2} + \frac{3}{\log(n)^3}$. Vi får

$$S_x = \sum_{n\le x} z_n - \frac{1}{\log(n)},$$

$$B_x = \sqrt{\sum_{n\le x} \frac{1}{\log(n)} - \frac{1}{\log(n)^2}},$$

$$M_x = \sum_{n\le x} \frac{1}{\log(n)} - \frac{3}{\log(n)^2} + \frac{3}{\log(n)^3}.$$

Då er det klart at $\lim_{x\to\infty} B_x^2 = \infty$, så $\lim_{x\to\infty} B_x = \infty$. I tillegg får vi

$$\frac{M_x}{B_x^3} = \frac{\sum_{n\le x} \frac{1}{\log(n)} - \frac{3}{\log(n)^2} + \frac{3}{\log(n)^3}}{\left(\sqrt{\sum_{n\le x} \frac{1}{\log(n)} - \frac{1}{\log(n)^2}}\right)^3}$$

$$\sim \frac{\sum_{n\le x} \frac{1}{\log(n)}}{\left(\sqrt{\sum_{n\le x} \frac{1}{\log(n)}}\right)^3} = \frac{1}{\sqrt{\sum_{n\le x} \frac{1}{\log(n)}}}$$

$$\leq \frac{1}{\log\left(\sqrt{\sum_{n\leq x}\frac{1}{\log(n)}}\right)} = \frac{1}{\log(B_x)}.$$

Så alle krava for den generelle lova om den itererte logaritmen er oppfylte, og vi får

$$1 = \limsup_{x\to\infty} \frac{\left|\sum_{n\leq x} z_n - \frac{1}{\log(n)}\right|}{\sqrt{\sum_{n\leq x}\frac{1}{\log(n)} - \frac{1}{\log(n)^2}} * \sqrt{2\log\log\left(\sqrt{\sum_{n\leq x}\frac{1}{\log(n)} - \frac{1}{\log(n)^2}}\right)}}$$

$$= \limsup_{x\to\infty} \frac{|\Pi(x) - \mathrm{Li}(x)|}{\sqrt{\sum_{n\leq x}\frac{1}{\log(n)}} * \sqrt{2\log\log(x)}}$$

$$= \limsup_{x\to\infty} \frac{|\Pi(x) - \mathrm{Li}(x)|}{\sqrt{\frac{x}{\log(x)}} * \sqrt{2\log\log(x)}} = \limsup_{x\to\infty} \frac{|\Pi(x) - \mathrm{Li}(x)|}{\sqrt{2x}\sqrt{\frac{\log\log(x)}{\log(x)}}} \; a.s$$

□

I lys av teorem 5 er det naturleg å tenka at den tilsvarande relasjonen gjeld for $\pi(x)$:

Formoding 1.

$$\limsup_{x\to\infty} \frac{|\pi(x) - \mathrm{Li}(x)|}{\sqrt{2x} * \sqrt{\frac{\log\log(x)}{\log(x)}}} = 1 \; a.s$$

Så teorem 5 føreslår at feilleddet mellom $\pi(x)$ og $\mathrm{Li}(x)$ er av storleiksorden $\sqrt{x}\sqrt{\frac{\log\log(x)}{\log(x)}}$. Dette stemmer overeins med (7) og (8), men føreslår samstundes ei forbetring av (8), nemleg

$$|\pi(x) - \mathrm{Li}(x)| = O\left(\sqrt{x}\sqrt{\frac{\log\log(x)}{\log(x)}}\right) \tag{9}$$

Som gjev oss

$$\pi(x) = \mathrm{Li}(x) + O\left(\sqrt{x}\sqrt{\frac{\log\log(x)}{\log(x)}}\right)$$

Så det Cramér-modellen føreslår er at $\mathrm{Li}(x)$ er ei enno betre tilnærming av $\pi(x)$ enn det ein til no har trudd. Kanskje enno meir interessant er det at vi får ei betre avgrensing av feilleddet til $|\pi(x) - \mathrm{Li}(x)|$ enn det vi hadde gjort sjølv om ein antok Riemann-hypotesen! Dersom formoding 1 stemmer, ville vi vore i stand til å avgjera eigenskapane til $\pi(x)$ med mykje større nøyaktigheit. Uheldigvis er vi, på dette tidspunktet, ikkje i nærleiken av verken å visa eller å motvisa formoding 1.

5 Primtal i korte intervall

I denne delen skal vi sjå på intervall av typen $[x, x + \lambda \log x]$, der $\lambda \in \mathbb{R}_{>0}$, og på kor mange primtal ein kan finna i slike intervall. Cramér-modellen gjev eit forslag til fordelinga til primtal i slike intervall.
Vi skal ta i bruk det følgjande klassiske resultatet frå sannsynsteorien:

Teorem 7. *Vi lét $(p_n)_{n\in\mathbb{N}}$ vera ei følgje av reelle tal i intervallet $[0,1]$ der $lim_{n\to\infty} n * p_n = \lambda$ og $\lambda \in \mathbb{R}$. Då har vi*

$$\lim_{n\to\infty} \binom{n}{k} p^k * (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Vi kjenner att uttrykka i følgja som punktsannsynet til ei binomisk fordeling med parametrar n og p_n, og tilsvarande ser vi at grense-uttrykket er punktsannsynet til ei poissonfordeling med parameter $\lambda = \lim_{n\to\infty} np_n$. Teorem 7 kan derfor tolkast på denne måten: Ei følgje av binomisk fordelte variablar, der følgja av forventingsverdiar np_n konvergerer mot λ, vil konvergera mot ein poissonfordelt variabel med parameter λ.

Prov. Sidan $\lim_{n\to\infty} n * p_n = \lambda$ får vi

$$\lim_{n\to\infty} \binom{n}{k} p^k * (1-p)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n(n-1)(n-2)...(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k * \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n^k + O(n^{k-1})}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \frac{\lambda^k}{k!} * \left(1 - \frac{\lambda}{n}\right)^n * \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$= \frac{\lambda^k}{k!} e^{-\lambda}.$$

□

Sidan z_n er binomisk fordelt for alle n, er $\Pi(x)$ binomisk fordelt. Ein kan derfor, ved hjelp av teorem 7 visa følgjande overraskande resultat om $\Pi(x)$:

Teorem 8. *Gitt ein $\lambda > 0 \in \mathbb{R}$ og $k \geq 0 \in \mathbb{Z}$ har vi, med sannsyn 1*

$$\#\{x \leq X \mid x \in \mathbb{Z},\ \Pi(x + \lambda \log(x)) - \Pi(x) = k\} \sim e^{-\lambda} \frac{\lambda^k}{k!} X \tag{10}$$

når $X \to \infty$

Prov. Vi nytter det førre resultatet. Sidan $\Pi(x)$ er ein sum av binomiske variablar får vi

$$\Pi(x + \lambda \log(x)) - \Pi(x) = \sum_{n=3}^{x+\lambda \log(x)} z_n - \sum_{n=3}^{x} z_n = \sum_{n=x+1}^{x+\lambda \log(x)} z_n = \sum_{n=1}^{\lambda \log(x)} z_{x+n}.$$

Vi ser så på forventingsverdien til $\Pi(x+\lambda\log(x))-\Pi(x)$

$$\mathrm{E}(\Pi(x+\lambda\log(x))-\Pi(x)) = \mathrm{E}\left(\sum_{n=1}^{\lambda\log(x)} z_{x+n}\right) = \sum_{n=1}^{\lambda\log(x)} \frac{1}{\log(x+n)}.$$

n er såpass mykje mindre enn x at $\frac{1}{\log(x+n)} \sim \frac{1}{\log(x)}$, så vi får

$$\mathrm{E}(\Pi(x+\lambda\log(x))-\Pi(x)) \sim \sum_{n=1}^{\lambda\log(x)} \frac{1}{\log(x)} = \frac{\lambda\log(x)}{\log(x)} = \lambda.$$

Så forventingsverdien er konstant og uavhengig av x. Dersom vi no lét x gå mot uendeleg vil $\Pi(x+\lambda\log(x))-\Pi(x)$ vera poissonfordelt med parameter λ. Då får vi

$$\mathrm{Prob}(\Pi(x+\lambda\log(x))-\Pi(x)=k) = \frac{\lambda^k}{k!}e^{-\lambda},$$

og resultatet følgjer. □

Teorem 8 føreslår ein asymptotisk storleik for talet på korte delintervall i $[0,X]$ som inneheld k primtal. Ein kan spørja seg om dette er rimeleg å anta. Vidare i denne delen blir det ført eit argument for at relasjonen

$$\#\{x \le X \mid \pi(x+\lambda\log(x))-\pi(x)=k\} \sim \frac{\lambda^k}{k!}e^{-\lambda}X \tag{11}$$

burde gjelda når $X\to\infty$.
Ein kan ikkje visa at likning (11) gjeld utan antakinger. Vi er nøydde til å anta den såkalla Hardy-Littlewood-formodinga om k-tuplar av primtal først.

Definisjon 4. Vi lét mengda $H_k = \{h_1, h_2, ..., h_k\}$ bestå av k distinkte naturlege tal. Vi lét så $v_{H_k}(p)$ vera antal distinkte heiltal modulo p som har eit element frå H_k i seg. Dersom $v_{H_k}(p) < p$ for alle primtal p kallar vi H_k for ei tillateleg mengde. Vi definerer funksjonen $\pi(x;H_k)$ ved

$$\pi(x;H_k) = \#\{n \le x \mid n+h_1, n+h_2, ..., n+h_k \text{ er primtal}\}. \tag{12}$$

Vi definerer $C(H_k)$ ved

$$C(H_k) = \prod_{p \text{ primtal}} \left(1-\frac{1}{p}\right)^{-k}\left(1-\frac{v_{H_k}(p)}{p}\right), \tag{13}$$

der $v_{H_k}(p)$ er antal distinkte restklassar modulo p blant elementa i H_k. Merk at dersom H_k er tillateleg er $C(H_k)\neq 0$.

Formoding 2 (Hardy-Littlewood-formodinga om k-tuplar av primtal)**.** *For eit heiltal $k\ge 2$ og ei tillateleg mengde H_k har vi*

$$\pi(x;H_k) = C(H_k)\frac{x}{\log(x)^k}(1+o_k(1))$$

uniformt for $H_k \subset [1,h]$, der $h\sim\lambda\log(x)$ når $x\to\infty$ og $\lambda>0$ er ein konstant.

Konstanten $C(H_k)$ vil dukka opp i argumentet vårt, og det er derfor nyttig å ha ei avgrensing på $C(H_k)$.

Lemma 2. *Vi lét $H_k = \{h_1, h_2, ..., h_k\}$. Gitt ein konstant $k \geq 2$ og $h \to \infty$, har vi*

$$\sum_{1 \leq h_1, h_2, \ldots, h_k \leq h} C(H_k) = h^k + O(h^{k-\frac{1}{2}-\epsilon}).$$

Prov. Dette er ei grov skisse av provet for dette resultatet. Sjå Gallagher [6] for eit fullstendig prov.
Den p-ande faktoren i $C(H_k)$ er

$$1 - \frac{p^k - v_{H_k}(p) * p^{k-1} - (p-1)^k}{(p-1)^k} = 1 + a(p, v_{H_k}(p)),$$

der

$$1 + a(p, v) = \begin{cases} O(\frac{1}{p-1}^2) & v = r \\ O(\frac{1}{p-1}) & v < r \end{cases} \tag{14}$$

og følgjeleg konvergerer produktet $C(H_k)$. Dersom vi no definerer $a_H(q)$ for kvadratfrie q ved

$$a_H(q) = \prod_{p|q} a(p, H_k(p))$$

får vi ei absolutt konvergent rekkjeutvikling av $C(H_K)$:

$$C(H_k) = \sum_{q \text{ kvadratfri}} a_H(q)$$

Gitt ein $x \leq h$, og ved å nytta likning 14 kan det visast at

$$\sum_{q \geq x} |a_H(q)| = O((xh)^\epsilon / x).$$

Vi får då

$$\begin{aligned}
\sum_{1 \leq h_1, h_2, \ldots, h_k \leq h} C(H_k) &= \sum_{1 \leq h_1, h_2, \ldots, h_k \leq h} \left(\sum_{q \leq x} a_H(q) + \sum_{q \geq x} a_H(q) \right) \\
&= \sum_{1 \leq h_1, h_2, \ldots, h_k \leq h} \sum_{q \leq x} a_H(q) + \sum_{1 \leq h_1, h_2, \ldots, h_k \leq h} \sum_{q \geq x} a_H(q) \\
&= \sum_{q \leq x} \sum_{1 \leq h_1, h_2, \ldots, h_k \leq h} a_H(q) + O(h^k((xh)^\epsilon / x))
\end{aligned}$$

Som det, ved å nytta det kinesiske restteoremet og eit gitterpunkt-argument, kan visast at er

$$h^k + O(h^{k-\frac{1}{2}-\epsilon})$$

□

Dersom vi reknar med at formoding 2 gjeld, kan vi vise følgjande teorem:

Teorem 9. *Anta formoding 2. Når $X \to \infty$ har vi*

$$\#\{x \leq X \mid \pi(x+h) - \pi(x) = k\} \sim \frac{\lambda^k}{k!} e^{-\lambda} X \qquad \text{der } h \sim \lambda \log(X). \tag{15}$$

Vi skal visa dette ved å sjå på momenta til dei to storleikane. For å behandla momentet til poissonfordelingar introduserer vi det følgjande kombinatoriske uttrykket.

Definisjon 5. Stirling-talet av type 2 for k og r, $\left\{ {k \atop r} \right\}$ er definert ved

$$\left\{ {k \atop k} \right\} = 1, \qquad \left\{ {k \atop 1} \right\} = 1,$$

$$\left\{ {k \atop r} \right\} = \frac{1}{r!} \sum_{i=0}^{r} (-1)^i \binom{r}{i} (r-i)^k.$$

Talet $\left\{ {k \atop r} \right\}$ er talet på måtar å partisjonera ei mengde av k element inn i r ikkje-tomme delmengder.

Det følgjande korollaret er ein direkte konsekvens av definisjonen.

Korollar 2. *For $k > 0$ gjeld rekurensrelasjonen*

$$\left\{ {k+1 \atop r} \right\} = r \left\{ {k \atop r} \right\} + \left\{ {k \atop r-1} \right\}.$$

Vi kan nytte stirling-talet til å skildre poisson-moment av vilkårlege ordenar, som vist i det neste lemmaet.

Lemma 3. *Det k-ande momentet til ei poissonfordeling med parameter λ er*

$$m_k(\lambda) = \sum_{r=1}^{k} \left\{ {k \atop r} \right\} \lambda^r.$$

Prov. Den momentgenererande funksjonen til ei poissonfordeling med parameter λ er $f(t) = e^{\lambda(e^t - 1)}$. Det k-ande momentet til fordelinga er då $\frac{d^k}{(dt)^k} f(t)\Big|_{t=0}$. Vi skal ved induksjon på k visa at

$$\frac{d^k}{(dt)^k} f(t) = \sum_{r=1}^{k} \left\{ {k \atop r} \right\} \lambda^r e^{rt} e^{\lambda(e^t - 1)}. \tag{16}$$

Dersom vi deriverer f, får vi

$$\frac{d}{dt} f(t) = e^{\lambda\left(e^t - 1\right)} * \left(e^t \lambda\right) = \sum_{r=1}^{1} \left\{ {1 \atop i} \right\} \lambda^i e^{i*t} e^{\lambda\left(e^t - 1\right)}.$$

Så relasjonen gjeld for $k = 1$. Anta så at relasjonen gjeld for $k = n$, altså at

$$\frac{d^n}{(dt)^n} f(t) = \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \lambda^r e^{rt} e^{\lambda(e^t - 1)}.$$

Vi får då

$$\frac{d^{n+1}}{(dt)^{n+1}} f(t) = \frac{d}{dt} \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \lambda^r e^{rt} e^{\lambda(e^t-1)} = \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \lambda^r e^{rt+\lambda(e^t-1)} \left(r + \lambda e^t\right)$$

$$= \sum_{r=1}^{n} \left\{ {n \atop r} \right\} r \lambda^r e^{rt+\lambda(e^t-1)} + \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \lambda^{r+1} e^{rt+1+\lambda(e^t-1)}.$$

Dersom vi no grupperer ledda etter eksponenten til λ-faktoren får vi

$$= \left\{ {n \atop 1} \right\} \lambda e^t e^{\lambda(e^t-1)} + \sum_{r=2}^{n} \left(\left\{ {n \atop r} \right\} r + \left\{ {n \atop r-1} \right\} \right) \lambda^r e^{rt} e^{\lambda(e^t-1)} + \left\{ {n \atop n} \right\} \lambda^{n+1} e^{(n+1)t} e^{\lambda(e^t-1)}$$

som, ved korollar 2, er

$$= \left\{ {n+1 \atop 1} \right\} \lambda e^t e^{\lambda(e^t-1)} + \sum_{r=2}^{n} \left\{ {n+1 \atop r} \right\} \lambda^r e^{rt} e^{\lambda(e^t-1)} + \left\{ {n+1 \atop n+1} \right\} \lambda^{n+1} e^{(n+1)t} e^{\lambda(e^t-1)}$$

$$= \sum_{r=1}^{n+1} \left\{ {n+1 \atop r} \right\} \lambda^r e^{rt} e^{\lambda(e^t-1)}.$$

Så likning (16) gjeld for alle $k \in \mathbb{N}$. Då får vi, for alle $k \in \mathbb{N}$, at det k-ande momentet til fordelinga er

$$\sum_{r=1}^{k} \left\{ {k \atop r} \right\} \lambda^r e^{rt} e^{\lambda(e^t-1)} \Bigg|_{t=0} = \sum_{r=1}^{k} \left\{ {k \atop r} \right\} \lambda^r * 1 = \sum_{r=1}^{k} \left\{ {k \atop r} \right\} \lambda^r = m_k(\lambda).$$

$\square$

Vi er no klare for å visa teorem 9.

Prov av teorem 9. To storleikar er like dersom momenta deira er like. Det k-ande momentet til $\pi(x+h) - \pi(x)$ er

$$M_k(X) = \sum_{x \le X} (\pi(x+h) - \pi(x))^k = \sum_{x \le X} \sum_{x \le p_1, p_2, \ldots, p_n \le X} 1.$$

Dersom vi no samlar ledda med same antal distinkte primtal blant $p_1, p_2, \ldots, p_n$ får vi

$$M_k(X) = \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \sum_{1 \le h_1, h_2, \ldots h_r \le h} \pi(x; H_k). \tag{17}$$

Ved formoding 2 får vi

$$M_k(X) \sim \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \sum_{1 \le h_1, h_2, \dots h_r \le h} C(H_k) \frac{X}{\log(X)^r}$$

og ved lemma 2 får vi så

$$\sim \sum_{r=1}^{n} \left\{ {n \atop r} \right\} h^r \frac{X}{\log(X)^r} \sim \sum_{r=1}^{n} \left\{ {n \atop r} \right\} X \left(\frac{h}{\log(X)} \right)^r \sim X \sum_{r=1}^{n} \left\{ {n \atop r} \right\} \lambda^r.$$

Det siste leddet er det k-ande momentet til ei poissonfordeling med parameter λ, skalert med X. Sidan momenta er asymptotisk like, må også storleikane vera asymptotisk like. Resultatet følgjer. □

Utan å vita om formoding 2 stemmer kan vi ikkje trygt seia om teorem 9 er sant eller ikkje. Det vi derimot kan seie, er at sidan vi har god grunn til å tru at 2 stemmer, har vi også god grunn til å tru at teorem 9 stemmer. Teorem 8 gjeld med sannsyn 1 i Cramér-modellen, og vi ser då at Cramér-modellen har dei eigenskapane som vi, med god grunn, trur gjeld for primtala.

6 Den maksimale verdien til $p_{n+1} - p_n$

I denne delen skal vi studera maksimale primtalsgap, altså $\max_{p_n} p_{n+1} - p_n$, der p_n er eit primtal. Vi veit at denne storleiken er avgrensa nedanifrå av $\log(p_n)$. Dersom ein reknar med at Riemann-hypotesen stemmer får ein dette resultatet.

Teorem 10. *Anta formoding 4. Då har vi*

$$p_{n+1} - p_n = O(\sqrt{p_n} \log(p_n)). \tag{18}$$

Prov. Sjå Cramér [3]. □

Så, gitt Riemann-hypotesen er storleiken $p_{n+1} - p_n$ avgrensa ovanfrå av $\sqrt{p_n} \log(p_n)$. Sidan Riemann-hypotesen framleis ikkje er prova, er det ikkje mogleg å seia sikkert om denne relasjonen stemmer. Mykje innsats har vore lagt inn på å forsøka å prova (18) utan Riemann-hypotesen. Det beste resultatet som er oppnådd til no vart vist av Baker, Harman og Pintz i 2000, og er svært nært til resultatet vi har ved Riemann-hypotesen.

Teorem 11.

$$p_{n+1} - p_n = O(p_n^{0.525}).$$

Bevis. Sjå Baker, Harman og Pintz [1]. □

Så utan å anta Riemann-hypotesen veit vi at storleiksordenen til $p_{n+1} - p_n$ ligg mellom $\log(p_n)$ og $p_n^{0.525}$. Dette er eit greitt estimat, men vi ser at det enno er mykje rom for variasjon mellom endepunkta.

Vi skal no nytta Cramér-modellen til å utleia ei formoding om ei veldig mykje betre øvre grense for $p_{n+1} - p_n$. Mesteparten av denne prosessen handlar om å prove følgjande resultat for dei probabilistiske primtala

Teorem 12.
$$\limsup_{n\to\infty} \frac{P_{n+1} - P_n}{\log(P_n)^2} = 1 \ a.s$$

Før vi set i gong med provet treng vi litt sannsynsteori.

Definisjon 6. Vi lét $(E_n)_{n\in\mathbb{N}}$ vera ei følgje av utfall i eit sannsynsrom. Vi definerer følgjande mengder av utfall:

$$\limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n}^{\infty} E_n, \tag{19}$$

$$\liminf_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} \bigcap_{k\geq n}^{\infty} E_n. \tag{20}$$

Lemma 4 (Det første Borel-Cantelli-lemmaet)**.** *Vi lét* $E_1, E_2, ..$ *vera ei følgje av hendingar i eit sannsynsrom. Dersom*

$$\sum_{n=1}^{\infty} \text{Prob}(E_n) < \infty,$$

så har vi

$$\text{Prob}\left(\limsup_{n\to\infty} E_n\right) = 0.$$

Prov. Betrakt følgja $(\bigcup_{n=N}^{\infty} E_n)_{N=1}^{\infty}$. Dette er ei monotont minkande følgje, så vi får, ved monoton konvergens av sannsyn

$$\text{Prob}\left(\limsup_{n\to\infty} E_n\right) = \lim_{N\to\infty} \text{Prob}\left(\bigcup_{n=N}^{\infty} E_n\right)$$

$$\leq \lim_{N\to\infty} \sum_{n=N}^{\infty} \text{Prob}(E_n).$$

Sidan $\sum_{n=1}^{\infty} \text{Prob}(E_n) < \infty$ får vi

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} \text{Prob}(E_n) = 0.$$

□

Lemma 5 (Det andre Borel-Cantelli-lemmaet)**.** *Vi lét* $E_1, E_2, ..$ *vera ei følgje av uavhengige hendingar i eit sannsynsrom. Dersom*

$$\sum_{n=1}^{\infty} \text{Prob}(E_n) = \infty,$$

så har vi

$$\text{Prob}\left(\limsup_{n\to\infty} E_n\right) = 1.$$

Prov. Dette er ekvivalent med å visa at $1 - \text{Prob}\left(\limsup_{n\to\infty} E_n\right) = 0$. Vi får

$$1 - \text{Prob}\left(\limsup_{n\to\infty} E_n\right) = \text{Prob}\left(\limsup_{n\to\infty} E_n\right)^c$$

$$= \text{Prob}\left(\left(\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}^{\infty} E_n\right)^c\right).$$

Ved enkle mengdeteoretiske reglar er uttrykket over det same som

$$= \text{Prob}\left(\bigcup_{n=1}^{\infty}\bigcap_{k\geq n}^{\infty} (E_n)^c\right) = \lim_{N\to\infty} \text{Prob}\left(\bigcap_{n=N}^{\infty} (E_n)^c\right).$$

Det er no nok å visa at $\text{Prob}\left(\bigcap_{n=N}^{\infty}(E_n)^c\right) = 0 \ \forall n \in \mathbb{N}$. Sidan hendingane er uavhengige får vi

$$\text{Prob}\left(\bigcap_{n=N}^{\infty} (E_n)^c\right) = \prod_{n=N}^{\infty} \text{Prob}(E_n^c) = \prod_{n=N}^{\infty} (1 - \text{Prob}(E_n))$$

$$\leq \prod_{n=N}^{\infty} \exp(-\text{Prob}(E_n)) = \exp\left(-\sum_{n=N}^{\infty} \text{Prob}(E_n)\right) = 0,$$

sidan rekkja divergerer. □

Vi kan no prove teorem 12.

Prov av teorem 12. Vi definerer tre hendingar:

- A=$\{P_{n+1} - P_n > c\log(P_n)^2$ for uendeleg mange $n\}$
- E_m=$\{z_{m+v} = 0$, når v oppfyller $1 \leq v \leq c\log(m)^2\}$
- B=$\{E_m$ skjer for uendeleg mange $m\}$

Det er klart at dersom E_m skjer, så er relasjonen $P_{m+1} - P_m > c(\log(P_n))^2$ oppfylt. Det betyr at dersom A skal skje, så må E_m skje for uendeleg mange m. Derfor har vi då $\text{Prob}(A) = \text{Prob}(B)$. Så, for å finna $\text{Prob}(A)$ er det derfor nok å finna $\text{Prob}(B)$. Hendinga B er at hendingane E_m skjer for uendeleg mange m. Dette betyr at $B = \bigcap_{m=1}^{\infty}\bigcup_{k\geq m}^{\infty} E_k = \limsup_{m\to\infty} E_m$. Vi har derfor lyst å nytte Borel-Cantelli-lemmaa. Vi starter med å rekne ut $\text{Prob}(E_m)$ for ein gitt m. Vi får, sidan variablane z_n er uavhengige,

$$\text{Prob}(E_m) = \text{Prob}(z_{m+v} = 0 \text{ for } 1 \leq v \leq c(\log(m))^2)$$

$$= \prod_{v=1}^{c\log(m)^2} \text{Prob}(z_{m+v} = 0) = \prod_{v=1}^{c\log(m)^2} \left(1 - \frac{1}{\log(m+v)}\right). \tag{21}$$

Her er det ønskeleg å kunne avgrensa denne storleiken med ein litt enklare funksjon. Vi veit at for logaritmen er $\log(a+b) \sim \log(a)$ dersom a er veldig mykje større enn b. Vi vel oss derfor ein $\tilde{m}$ slik at storleiken til v er ubetydeleg samanlikna med m. Då får vi

$$\mathrm{Prob}(E_m) = \prod_{v=1}^{c\log(m)^2} \left(1 - \frac{1}{\log(m+v)}\right) \sim \prod_{v=1}^{c\log(m)^2} \left(1 - \frac{1}{\log(m)}\right).$$

Vidare har vi ved den førsteordens taylorutviklinga til $\exp(-x)$

$$\prod_{v=1}^{c\log(m)^2} \left(1 - \frac{1}{\log(m)}\right) \leq \prod_{v=1}^{c\log(m)^2} \exp\left(-\frac{1}{\log(m)}\right) = \exp\left(-\sum_{v=1}^{c\log(m)^2} \frac{1}{\log(m)}\right).$$

Summen i eksponenten er uavhengig av v, så vi får

$$\exp\left(-\sum_{v=1}^{c\log(m)^2} \frac{1}{\log(m)}\right) = \exp\left(-\frac{c\log(m)^2}{\log(m)}\right) = \exp\left(-c\log(m)\right)$$

$$= \exp\left(\log\left(\frac{1}{m^c}\right)\right) = \frac{1}{m^c}.$$

Derfor gjeld følgjande relasjon:

$$\frac{K}{m^c} < \mathrm{Prob}(E_m) < \frac{L}{m^c} \tag{22}$$

for konstantar K og L.
Så dersom vi ser på $\sum_m \mathrm{Prob}(E_m)$ får vi

$$\sum_m \mathrm{Prob}(E_m) = \sum_{m=1}^{\tilde{m}-1} \mathrm{Prob}(E_m) + \sum_{\tilde{m}}^{\infty} \mathrm{Prob}(E_m)$$

$$< \sum_{m=1}^{\tilde{m}-1} \mathrm{Prob}(E_m) + L * \sum_{\tilde{m}}^{\infty} \frac{1}{m^c} < \infty$$

dersom $c > 1$, og

$$\sum_m \mathrm{Prob}(E_m) = \sum_{m=1}^{\tilde{m}-1} \mathrm{Prob}(E_m) + \sum_{\tilde{m}}^{\infty} \mathrm{Prob}(E_m)$$

$$> \sum_{m=1}^{\tilde{m}-1} \mathrm{Prob}(E_m) + K * \sum_{\tilde{m}}^{\infty} \frac{1}{m^c} > \infty$$

dersom $c < 1$. Ved Borel-Cantelli betyr dette at sannsynet for B, og dermed også for A, er 0 dersom $c > 1$, og 1 dersom $c < 1$. Med omsyn på ulikskapen som definerte hendinga A får vi

$$(1-\delta) * (\log(P_n))^2 \leq P_{n+1} - P_n \leq (1+\delta) * (\log(P_n))^2 \tag{23}$$

for uendeleg mange n og alle $\delta > 0$. Dette betyr at

$$\lim_{n\to\infty} \frac{P_{n+1} - P_n}{\log(P_n)^2} \leq \lim_{n\to\infty} \frac{(1+\delta)\log(P_n)^2}{\log(P_n)^2} = 1 + \delta.$$

Men, samstundes har vi

$$\lim_{n\to\infty} \frac{P_{n+1} - P_n}{\log(P_n)^2} \geq \lim_{n\to\infty} \frac{(1-\delta)\log(P_n)^2}{\log(P_n)^2} = 1 - \delta.$$

Så når vi lét δ gå mot 0 må vi få at

$$\lim_{n\to\infty} \frac{P_{n+1} - P_n}{\log(P_n)} = 1$$

og resultatet er vist. $\square$

Sidan $\limsup_{n\to\infty} \frac{P_{n+1}-P_n}{\log(P_n)^2} = 1$ gjeld for følgja (P_n) av variablar i Cramér-modellen er det rimeleg å tru at den same relasjonen gjeld for følgja (p_n) av primtal.

Formoding 3 (Cramér-formodinga)**.** *For følgja (p_n) av primtal gjeld*

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{\log(p_n)^2} = 1 \ a.s \tag{24}$$

Altså har vi

$$\max_{p_n \leq x}(p_{n+1} - p_n) \sim \log(x)^2. \tag{25}$$

Som vi ser, føreslår Cramér-modellen ei ganske mykje strengare øvre grense for storleiken på $p_{n+1} - p_n$, og som i dei tidlegare avsnitta er denne grensa betre enn den vi får når vi antar Riemann-hypotesen.
Fram til no har vi sett på formodingar som Cramér-modellen føreslår som forbetrar resultata om primtala vi hadde frå før. Dette er interessante resultat, men spørsmålet om kor rimelege desse resultata er direkte avhengig av kor rimeleg modellen vår. Dette betyr at rimelegheita til mellom anna Cramér-formodinga er avhengig av kor bra Cramér-modellen etterlikner primtala. Som vi no skal sjå, finn ein ikkje alle eigenskapane til primtala i følgja (P_n) i Cramér-modellen.

7 Problem i Cramér-modellen

Følgja av primtal er deterministisk, medan følgjene i Cramér-modellen er probabilistiske. Dette fører til eit par openbare avvik mellom primtala og følgja (P_ν). Mellom anna har vi følgjande teorem.

Teorem 13. *Det er, asymptotisk sett, like mange odde og jamne probabilistiske primtal. Med andre ord*

$$\liminf_{\nu\to\infty}(P_{\nu+1} - P_\nu) = 1 \ a.s$$

Prov. Dette er det same som å visa at vi, med sannsyn 1, kan finna uendeleg mange etterfølgjande probabilistiske primtal. Vi kallar hendinga {Det er uendeleg mange etterfølgjande probabilistiske primtal} for A. For å visa at A har sannsyn 1 definerer vi hendingane E_m:

- $E_m = \{z_{m+i}=1, \text{ for } i \in [1, \log(m)^2]\}$

Sidan variablane (z_n) er uavhengige, blir sannsynet for E_m

$$\mathrm{Prob}(E_m) = \prod_{i=1}^{\log(m)^2} \frac{1}{\log(m+i)}.$$

Med eit tilsvarande triks som i provet for teorem 12 eksisterer det $\tilde{m}$, K og L slik at

$$\frac{K}{m} \leq \mathrm{Prob}(E_m) \leq \frac{L}{m},$$

når $\tilde{m} \leq m$. Då har vi at

$$\sum_{m=1}^{\infty} \mathrm{Prob}(E_m) = \sum_{m=1}^{\tilde{m}} \mathrm{Prob}(E_m) + \sum_{m=\tilde{m}+1}^{\infty} \mathrm{Prob}(E_m)$$

$$\geq \sum_{m=1}^{\tilde{m}} \mathrm{Prob}(E_m) + K * \sum_{m=\tilde{m}+1}^{\infty} \frac{1}{m} = \infty.$$

Så, ved 2.Borel-Cantelli er

$$\mathrm{Prob}(A) = \mathrm{Prob}(\limsup_{m\to\infty} E_m) = 1.$$

□

Det er sjølvsagt heilt absurd å anta at dette stemmer for følgja av primtal, og det eksisterer fleire slike "openbare"avvik. Til dømes kan ein også visa, slik som i [10], at i Cramér-modellen er det, asymptotisk sett, $\frac{x}{\log(x)^2}$ tvillingprimtal mindre enn x, som motseier formoding 2. Felles for dei "openbare"avvika er at dei omhandlar lokale problem. Dette er stort sett problem som ein uansett ikkje er interesserte i å nytta Cramér-modellen på. Det er først og fremst når ein ser på primtal i intervall av ei viss lengde at det er nyttig å bruka Cramér-modellen.

Det vart lenge trudd at Cramér-modellen etterlikna primtala på ein korrekt måte i globale og semi-globale problem. Det følgjande resultatet til Maier viser at dette heller ikkje er tilfellet

Teorem 14. *Vi lét $\lambda > 0$. Då har vi*

$$\begin{aligned} \limsup_{n\to\infty} \frac{\pi(x+\log(x)^{\lambda}) - \pi(x)}{\log(x)^{\lambda-1}} &> 1, \\ \liminf_{n\to\infty} \frac{\pi(x+\log(x)^{\lambda}) - \pi(x)}{\log(x)^{\lambda-1}} &< 1. \end{aligned} \tag{26}$$

Prov. Sjå Maier [9]. □

Dette står i kontrast til det ein forventar frå Cramér-modellen.

Teorem 15. *For alle $\lambda > 2$ har vi*

$$\Pi(x + \log(x)^\lambda) - \Pi(x) \sim \frac{\log(x)^\lambda}{\log(x)} = \log(x)^{\lambda-1}.$$

Vi treng eit lite lemma for å prova dette.

Lemma 6. *Gitt ein $y \in \mathbb{R}$, anta at y=O(f(x)) og at $\frac{y}{\log(x)^2} \to \infty$. Då har vi*

$$\Pi(x + y) - \Pi(x) = \int_x^{x+y} \frac{dt}{\log(t)} + O(\sqrt{y}),$$

og dermed

$$\Pi(x + y) - \Pi(x) = (1 + o(1))\frac{y}{\log(x)}$$

med sannsyn 1.

Prov. Vi lét y vera gitt som i lemmaet.

$$\Pi(x + y) - \Pi(x) =$$

$$\Pi(x + y) - \mathrm{Li}(x + y) - \Pi(x) + \mathrm{Li}(x) + \mathrm{Li}(x + y) - \mathrm{Li}(x).$$

Ved teorem 5 og additivitet av integralet får vi at dette er asymptotisk det same som

$$\sqrt{2(x + y)} * \sqrt{\frac{\log\log(x + y)}{\log(x + y)}} - \sqrt{2x} * \sqrt{\frac{\log\log(x)}{\log(x)}} + \int_x^{x+y} \frac{dt}{\log(t)}. \tag{27}$$

Dersom vi no ser bort i frå integralet, har vi at

$$\sqrt{2(x + y)} * \sqrt{\frac{\log\log(x + y)}{\log(x + y)}} - \sqrt{2x} * \sqrt{\frac{\log\log(x)}{\log(x)}}$$

$$\sim \sqrt{\frac{\log\log(x)}{\log(x)}} * \left(\sqrt{2(x + y)} - \sqrt{2x}\right) \leq \sqrt{2(x + y)} - \sqrt{2x}$$

$$= \frac{2x + 2y - 2x}{\sqrt{2(x + y)} + \sqrt{2x}} = \sqrt{2}\frac{y}{\sqrt{x + y} + \sqrt{x}} \leq \sqrt{2}\frac{y}{\sqrt{x + y}}$$

$$\leq \sqrt{2}\frac{y}{\sqrt{y}} = \sqrt{2}\sqrt{y}$$

$$\implies \sqrt{2(x + y)} * \sqrt{\frac{\log\log(x + y)}{\log(x + y)}} - \sqrt{2x} * \sqrt{\frac{\log\log(x)}{\log(x)}} = O(\sqrt{y}).$$

Så (27) vil bli

$$\int_x^{x+y} \frac{dt}{\log(t)} + O(\sqrt{y})$$

og første del av teoremet er då prova.
Del 2 av teoremet følgjer frå den første. Sidan

$$\max_{t\in[x,x+y]} \frac{1}{\log(t)} = \frac{1}{\log(x)},$$

$$\min_{t\in[x,x+y]} \frac{1}{\log(t)} = \frac{1}{\log(x+y)}$$

får vi, ved monotonitet av integralet,

$$\frac{y}{\log(x+y)} \leq \int_x^{x+y} \frac{dt}{\log(t)} \leq \frac{y}{\log(x)}.$$

Dersom vi no lét x gå mot uendeleg får vi

$$\lim_{x\to\infty} \frac{\int_x^{x+y} \frac{dt}{\log(t)}}{\frac{y}{\log(x)}} \leq \lim_{x\to\infty} \frac{\frac{y}{\log(x)}}{\frac{y}{\log(x)}} = 1,$$

men samstundes

$$\lim_{x\to\infty} \frac{\int_x^{x+y} \frac{dt}{\log(t)}}{\frac{y}{\log(x)}} \geq \lim_{x\to\infty} \frac{\frac{y}{\log(x+y)}}{\frac{y}{\log(x)}} = \lim_{x\to\infty} \frac{\frac{y}{\log(x)}}{\frac{y}{\log(x)}} = 1.$$

Så vi konkluderer med at

$$\lim_{x\to\infty} \frac{\int_x^{x+y} \frac{dt}{\log(t)}}{\frac{y}{\log(x)}} = 1$$

$$\implies \int_x^{x+y} \frac{dt}{\log(t)} \sim \frac{y}{\log(x)} \sim (1+o(1))\frac{y}{\log(x)}.$$

□

Når vi har prova lemma 6 er det grei skuring å prova teorem 15.

Prov av teorem 15. Vi lét $\lambda > 2$ og set $y = \log(x)^\lambda$. Det er då klart at

$$\lim_{x\to\infty} \frac{y}{\log(x)^2} = \lim_{x\to\infty} \log(x)^{\lambda-2} \geq \lim_{x\to\infty} \log(x) = \infty.$$

Frå lemma 6 får vi då

$$\Pi(x+\log(x)^\lambda) - \Pi(x) = (1+o(1))\frac{\log(x)^\lambda}{\log(x)} = (1+o(1))\log(x)^{\lambda-1} \sim \log(x)^{\lambda-1}.$$

□

Det eksisterer også globale avvik mellom Cramér-modellen og primtala. For å visa dette skal vi, av tekniske årsaker, betrakta dei vekta variablane

$$\xi_n = \log(n) * z_n \qquad (n \geq 3)$$

slik at vi får $\mathrm{E}(\xi_n) = 1$. Dersom Cramér-modellen etterlikner primtala, vil det vera naturleg å tru at feilleddet mellom $\pi(x)$ og $\frac{x}{\log(x)}$, $\Delta'(x)$, definert ved

$$\Delta'(x) = \sum_{2<p\leq x} \log(p) - \sum_{2<n\leq x} 1, \qquad x \in \mathbb{Z}$$

kan tilnærmast bra ved hjelp av standardavviket til funksjonen $\Pi(x)$. Dersom vi følgjer denne tankegangen får vi

$$\mathrm{Var}(\Pi(x)) = \mathrm{Var}\left(\sum_{3\leq n\leq x} \xi_n\right) = \sum_{3\leq n\leq x} \mathrm{Var}(\xi_n)$$

$$= \sum_{3\leq n\leq x} \mathrm{Var}(\log(n) z_n) = \sum_{3\leq n\leq x} \log(n)^2 \mathrm{Var}(z_n)$$

$$= \sum_{3\leq n\leq x} \log(n)^2 * \left(\frac{1}{\log(n)} - \frac{1}{\log(n)^2}\right) = \sum_{3\leq n\leq x} (\log(n) - 1) \sim x\log(x)$$

$$\Longrightarrow \Delta'(x) \approx \sqrt{\mathrm{Var}(\Pi(x))} \sim \sqrt{x\log(x)}.$$

Så ein vil forventa at storleiksordenen til $\Delta'(x)$ er, i gjennomsnitt, $\sqrt{x\log(x)}$. Dette stemmer overeins med resultata vi har for $\Delta'(x)$ (vist i [7] og [8], henhaldsvis), nemleg

$$\Delta'(x) = O(\sqrt{x}\log(x)^2) \qquad \text{gitt formoding 4.} \tag{28}$$

$$\Delta'(x) = \Omega(\sqrt{x}\log\log\log(x)). \tag{29}$$

Derimot er det i strid med følgjande teorem

Teorem 16. *Anta formoding 4. Då gjeld følgjande relasjon for alle $Y \in \mathbb{R}$*

$$\frac{1}{Y}\int_2^Y \Delta'(x)^2 dx = O(Y)$$

Prov. Dette teoremet er vist i [4], ironisk nok av Cramér sjølv. □

Teorem 16 impliserer at $\Delta'(x)^2$, i gjennomsnitt, er av storleiksorden $O(x)$, som igjen impliserer at $|\Delta'(x)|$ har gjennomsnittleg storleiksorden $\sqrt{x}$. Dette stemmer også med likning (28) og likning (29), men strider i mot Cramér-modellen, som seier at den gjennomsnittlege storleiksorden til $\Delta'(x)$ er $\sqrt{x}\log(x)$. Det kan vera nyttig å nemna eit par skilnadar mellom avviket funne av Maier (teorem 14) og avviket ein får frå $\Delta'(x)$ (teorem 16):

- Avviket mellom Cramér-modellen og teorem 14 er, kvantitativt sett, større enn avviket mellom Cramér-modellen og teorem 16.
- Avviket mellom Cramér-modellen og teorem 14 er semi-globalt, medan avviket mellom Cramér-modellen og teorem 16 er globalt.
- Avviket mellom Cramér-modellen og teorem 14 gjeld berre for nokre få x-verdiar, medan avviket mellom Cramér-modellen og teorem 16 gjeld for alle x større enn 0.

Desse skilnadane vil vera viktige når vi no skal forsøka å korrigera Cramér-modellen.

8 Moglege forbetringar av Cramér-modellen

Teorem 14 og 16 sår tvil om Cramér-modellen er ei rimeleg etterlikning av primtala i globale og semi-globale tilfelle. Det blir no vist at Cramér-modellen kan modifiserast på ein slik måte at den stemmer overeins med teorem 14. Før vi modifiserer modellen introduserer vi litt nyttig notasjon.

Definisjon 7. Gitt ein $z \in \mathbb{R}$ lét vi

$$P(z) = \prod_{p \leq z} p.$$

Vi definerer vidare

$$Q(z) = \{x \in \mathbb{N} \mid \gcd(x, P(z)) = 1\}.$$

Mengda $Q(z)$ vert kalla mengda av z-kvasiprimtal.

Provet for teorem 14 baserer seg, grovt sett, på at når ein avgrensar til eit intervall av typen $(x, x + \log(x)^\lambda)$ har primtala og z-kvasiprimtala dei same irregularitetane. Dei naturlege tala i det same intervallet vil ikkje ha desse irregularitetane, og det er derfor Cramér-modellen ikkje tilfredsstiller teorem 14. Ved å endra på modellen slik at den berre inneheld variablar indekserte med eit z-kvasiprimtal vil også modellen ha dei naudsynte irregularitetane. Dette er tankegangen bak den korrigerte Cramér-modellen.

Definisjon 8 (Den korrigerte Cramér-modellen)**.** Vi lét $z = z(x)$ vera ein parameter som tilfredsstiller

$$\lim_{x \to \infty} \frac{z}{\log(x)^c} = \infty \qquad \text{og} \qquad z \leq x^d \quad c > 0, d < 1/2.$$

For $n = 3, 4, ...$ definerer vi den diskrete stokastiske variabelen ζ_n ved

$$\zeta_n = 0, \tag{30}$$

dersom $n \notin Q(z)$, og ved

$$\text{Prob}(\zeta_n = 1) = \prod_{p \leq z} (\frac{p}{p-1}) * \frac{1}{\log(n)},$$

$$\text{Prob}(\zeta_n = 0) = 1 - \prod_{p \leq z} (\frac{p}{p-1}) * \frac{1}{\log(n)}$$

dersom $n \in Q(z)$.
Ein kan tilsvarande som i kapittel 3 definera følgja $(\tilde{P}_\nu)$ og funksjonen $\tilde{\Pi}(x)$ ved

$$\tilde{P}_{\nu+1} = m \qquad \text{dersom} \sum_{n=3}^{m} \zeta_n = \nu \text{ og } \sum_{n=3}^{m-1} \zeta_n = \nu - 1,$$

$$\tilde{\Pi}(x) = \sum_{n=3}^{x} \zeta_n.$$

Merk. Med $z = 1$ er dette den tradisjonelle Cramér-modellen, og dersom $z \geq \sqrt{x}$ vil $Q(z) = \mathcal{P} \cap [0, x]$, som krev at ein kjenner alle primtala mindre x, som vert ein umogleg føresetnad når vi lét x gå mot uendeleg.

Ein til dels uventa konsekvens av definisjonen er at ζ_p er konstant lik 0 for alle primtal mindre enn z, sidan $p \notin Q(z)$. Dette kan i utgangspunktet virka som eit problem, men med tanke på at det berre er endeleg mange primtal mindre enn z er det i praksis ikkje noko problem å gje slepp på desse primtala. Dette er fordi vi i hovudsak fokuserer på primtala større enn $z(x)$, anten på intervallet (x,∞) eller på intervallet $(x, x + y]$, og primtala i desse intervalla vil alle vera i $Q(z)$.
Ein umiddelbar fordel med den korrigerte modellen er at fleire av dei openbare problema i den klassiske modellen forsvinn. Det er til dømes ikkje lenger jamne probabilistiske primtal i modellen.

Teorem 17. *I den korrigerte Cramér-modellen er det endeleg mange jamne probabilistiske primtal.*

Prov. Gitt ein parameter z som i definisjonen, vil $P(z)$ vera større enn 2 når $x \to \infty$. Då har vi at $2|P(z)$, og dermed vil alle partal e oppfylla $\gcd(k,P(z))\geq 2$. Det følgjer at $e \notin Q(z)$ for alle partal e. Dermed vil vi også, for alle e, få

$$\zeta_e = 0.$$

Og dermed er $\text{Prob}(\zeta_e = 1) = 0$ for alle partal e. □

Viktigare enn dette er sjølvsagt at den korrigerte modellen samsvarer med teorem 14.
Det ser derimot ikkje ut til at den korrigerte Cramér-modellen stemmer overeins

med teorem 16. Ein kan visa at gitt ein parameter $z = x^{\alpha}$, med $\alpha < \frac{1}{2}$ kan ein, på same måte som for dei vekta variablane ξ_n visa at

$$\mathrm{Var}\left(\sum_{n=3}^{x} \zeta_n\right) \sim g(\alpha)x * \log(x),$$

der g er ein positiv, monotont minkande funksjon slik at

$$\lim_{\alpha \to \frac{1}{2}} g(\alpha) = 0.$$

Vi kan prøva å modifisera modellen vår ytterlegare i håp om at avviket mellom modellen og teorem 16 forsvinn. Dette fører oss til følgjande definisjon.

Definisjon 9 (Den generaliserte Cramér-modellen)**.** Gitt ein $x \in \mathbb{Z}$, og der $\mathcal{P}_x = \{p \text{ primtal} \mid p \leq x\}$, lét vi $S_x \in [1, x] \cap \mathbb{Z}$ vera ei mengde slik at

$$\mathcal{P}_x \subseteq S_x.$$

Då kallar vi S_x for ei mengde av potensielle primtal. Vi definerer vidare dei stokastiske variablane η_n ved

$$\eta_n = 0$$

dersom $n \notin S_x$, og ved

$$\mathrm{Prob}(\eta_n = 1) = \frac{x}{|S_x|}\frac{1}{\log(n)}, \qquad \mathrm{Prob}(\eta_n = 0) = 1 - \frac{x}{|S_x|}\frac{1}{\log(n)}$$

dersom $n \in S_x$.
Dette er den generaliserte Cramér-modellen.

Dette er den mest generelle sannsynsmodellen som bruker ideen til Cramér, men som vi skal sjå, er dei einaste vala av potensielle primtal S_x slik at modellen vår tilfredsstiller teorem 16 mengder S_x som er essensielt like $\mathcal{P}_x$. Meir formelt har vi følgjande teorem:

Teorem 18. *Vi lét x vera eit stort partal, $I = (\frac{x}{2}, x] \cap \mathbb{Z}$, og lét S_x vera dei potensielle primtala slik at*

$$\mathcal{P}_x = \mathcal{P} \cap I \subseteq S_x \subseteq I, \qquad A = \frac{|I|}{|S_x|},$$

For $n \in S_x$ er då

$$\mathrm{Prob}(\eta_n = 1) = \frac{A}{\log(n)} \qquad \mathrm{Prob}(\eta_n = 0) = 1 - \frac{A}{\log(n)}$$

Under desse antakingane gjeld følgjande implikasjon:

$$\mathrm{Var}\left(\sum_{n \in I} \eta_n\right) = O\left(\frac{x}{\log(x)^2}\right) \implies |S_x \setminus P_x| = O\left(\frac{x}{\log(x)^2}\right). \qquad (31)$$

Merk. Sidan vi ikkje nytter vekta variablar blir variansen $O(x/\log(x)^2)$. Dersom vi hadde nytta vekta variablar ville vi, som i førre kapittel, fått at variansen var $O(x)$.

Prov. For $n \in S_x$ er $\mathrm{E}(S_x) = \frac{A}{\log(x)}$. Vi får dermed

$$\begin{aligned}\mathrm{Var}\left(\sum_{n\in I}\eta_n\right) &= \sum_{n\in I}\mathrm{Var}(\eta_n) = \sum_{n\in S_x}(\mathrm{E}(\eta_n^2) - \mathrm{E}(\eta_n)^2)\\ &= \sum_{n\in S_x}\left(\frac{A}{\log(n)} - \frac{A^2}{\log(n)^2}\right)\\ &= \frac{|I|}{\log(x)^2}\left(\log(x) + O(1) - A\left(1 + O\left(\frac{1}{\log(x)}\right)\right)\right). \qquad (32)\end{aligned}$$

Sidan $\mathrm{Var}\left(\sum_{n\in I}\eta_n\right) = O(\frac{x}{\log(x)^2})$ må også (32) vera $O(\frac{x}{\log(x)^2})$. Dette skjer viss og berre viss $A = \log(x) + O(1)$. Frå dette får vi at

$$|S_x| = \frac{|I|}{\log(x)}\left(1 + O\left(\frac{1}{\log(x)}\right)\right) \iff |S_x \setminus \mathcal{P}_x| = O\left(\frac{x}{\log(x)^2}\right).$$

□

Teoremet over seier at ingen generaliserte Cramér-modellar tilfredsstiller teorem 16, med mindre dei potensielle primtala essensielt sett er like dei faktiske primtala. Det let til at så lenge vi antar at variablane i modellen er uavhengige, vil vi ikkje vera i stand til å modifisera Cramér-modellen på ein slik måte at teorem 16 er tilfredsstilt.

9 Konkluderande merknadar

Som vi har sett, kan bruk av Cramér-modellen gje oss formodingar om primtala som samsvarer, og i somme tilfelle, forbetrar dei resultata vi kjenner til i dag. Ein burde jamvel merka seg at det er avvik i modellen som det ikkje ser ut til at vi kan unngå utan å fjerna antakinga om uavhengigheit. Cramér-modellen, med sine uavhengige variablar, kan derfor ikkje etterlikne primtala på ein fullstendig tilfredsstillande måte. Det let til at for å modellera primtala på ein slik måte at vi unngår avvik er vi nøydde til å gje opp antakinga om uavhengige variablar. Ein modell med avhengige variablar vil samsvara betre med primtala, som på ingen måte er uavhengige. Denne modellen vil derimot vera mykje vanskelegare å arbeida med. Primtala er kompliserte, og det ser ut til at ein tilfredsstillande stokastisk modell av primtala vil vera like komplisert.

10 Appendiks: Funksjonane ζ og ψ

Formoding 4 (Riemann-hypotesen)**.** *Vi lét $\zeta : \mathbb{C} \to \mathbb{C}$ vera definert ved*

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \; primtal} \frac{1}{1 - p^{-s}}. \tag{33}$$

Dersom $\zeta(s) = 0$ og $0 \leq Re(s) \leq 1$, så er $Re(s) = 1/2$.

Definisjon 10. Vi definerer funksjonen $\psi(x)$ ved

$$\psi(x) = \sum_{p^m \leq x} \log(p).$$

Som i [5] kan det visast at

$$\psi(x) = x - \sum_{\zeta(\rho)=0} \frac{x^\rho}{\rho} + O(1)$$

der summen er over alle nullpunkt til funksjonen ζ frå formoding 4.

Lemma 7. *Gitt ein $T \in (0, \infty)$, lét vi Γ vera mengda*

$$\{\gamma \in (0, T) \mid \exists\, \beta \in (0, 1) \text{ slik at } \zeta(\beta + i\gamma) = 0\}.$$

Då har vi at

$$\sum_{\gamma \in \Gamma} \frac{1}{\gamma} = O(\log(T)^2).$$

Prov. Sjå Davenport [5]. □

Referansar

[1] R. C. Baker, G. Harman og J. Pintz. "The difference between consecutive primes. II". I: *Proc. London Math. Soc. (3)* 83.3 (2001), s. 532–562.

[2] A.A. Borovkov. *Probability theory.* Universitext. Springer, London, 2013, s. 568–572.

[3] H. Cramér. "On the order of magnitude of the difference between consecutive prime numbers". I: *Acta Arith.* 2.1 (1936), s. 23–46.

[4] H. Cramér. "Some theorems concerning prime numbers". I: *Arkiv för Mat. Astr. o Fys.* 15.5 (1920), s. 1–32.

[5] H. Davenport. *Multiplicative number theory.* Third. Bd. 74. Graduate Texts in Mathematics. Springer-Verlag, New York, 2000, s. 104–114.

[6] P. X. Gallagher. "On the distribution of primes in short intervals". I: *Mathematika* 23.1 (1976), s. 4–9.

[7] H. von Koch. "Sur la distribution des nombres premiers". I: *Acta Math.* 24.1 (1901), s. 159–182.

[8] J.E. Littlewood. "Sur la distribution des nombres premieres". I: *Comptes Rendus Acad. Sci. Paris* 158.5 (1914), s. 1869–1872.

[9] H. Maier. "Primes in short intervals". I: *Michigan Math. J.* 32.2 (1985), s. 221–225.

[10] J. Pintz. "Cramér vs. Cramér. On Cramér's probabilistic model for primes". I: *Funct. Approx. Comment. Math.* 37.part 2 (2007), s. 361–376.

NTNU

Kunnskap for en bedre verden