TG-pusleri #64: hvor mange kort, til vi ser et ess?

Nils Lid Hjort

Department of Mathematics, University of Oslo

8. juli 2026

Abstract. Hvor mange kort må jeg trekke, før jeg finner det første ess? Jeg finner fordelingen til denne ventetiden T, samt dens pene forventning $\mathrm{E}\,T = (N+1)/(n+1)$, der N er antall kort og n antall ess; altså $53/5 = 10.6$ for en standard kortstokk. I et par-tre randbemerkninger finner jeg også denne formelen via en Beta-approksimasjon. Jeg ser også på fordelinger og forventninger for andre, tredje, fjerde gang jeg finner mine ess, der det ligger en hel Dirichlet-fordeling og venter på oss, med en kobling til uniforme ordningsobservatorer. Dessuten viser jeg hvordan man kan estimere antall kort N, med full konfidensfordeling, i anvendelser der man kjenner antall ess n men ikke N; og delvis analogt hvordan man kan inferere på et ukjent antall ess n når N er kjent. Om du har 1000 mennesker i et rom, og må spørre 11 av dem, før du finner den første venstrehendte, hvor mange venstrehendte er det i rommet – her trenger vi både estimat og usikkerhet.

Key words: Beta- og Dirichlet-approksimasjon, ess, konfidenskurve, kortstokk, Manhattan Project, ventetider mellom funn

Fotnoten først

Dette har jeg arbeidet *&* kost meg med i løpet av to balkongsommerdager i kalenderåret 2026 e.Kr., der menneskeheten er ved ett av sine veivalg – og der valget i praksis er tatt; vi er alle med på The Manhattan Project of Our Time, om biter av våre sjeler og hoder så blir utslettet, slik eller slik, eller ei. Den tid på jord er forbi da Lillestøl kunne gi oss TG-puslerier som ikke en hvilken som helst laptop kan finne svaret på etter et lite minutt, eventuelt med litt mellomprompting. Nei, jeg har *ikke* brukt AI på dette, det behager meg inntil videre ikke, «and the authors have refrained from using any AI», som vi anfører i vårt *Disclosure Statement* i Hjort og Stoltenberg (2026a, og med lengre forklaring om det samme i 2026b; vi lurte på om vi skulle bruke «abstained»). Men jeg vil nesten tro at AI vil mestre oppgaven «skriv en 6-7-siders løsning av Lillestøls Pusleri #64, i professor N.L. Hjorts stundom noe blomstrede stil, med flere innfallsvinkler, kanskje med noe Bayesian Nonparametrics og konfidens, og gjerne med en indirekte referanse til historie eller litteratur», innen kort tid. Fantastisk og eventyrlig. Det er håpløst, og vi gir oss ikke.

1 Tiden til det første ess

Jostein Lillestøls *pusleri #64*, i den imponerende utholdende og langvarige serie med slike, fra TG juni 2026, er som følger:

«Ta en vanlig kortstokk med $N = 52$ kort, herav de fire Ess. Stokk den godt og legg stokken på bordet med billedsiden ned. Ta så kort, ett etter ett, fra toppen av stokken, snu kortet og

legg det ut på bordet med billedsiden opp. Fortsett til det første Ess er kommet opp. Hva er forventet antall snudde utlagte kort? Utfordring: Finn flere alternative løsningsmetoder.»

La meg starte med julestrømpe-formelen, som er kjekk å ha ombord for det følgende. Den sier

$$\begin{aligned}\binom{3}{3}+\binom{4}{3}+\cdots+\binom{50}{3}+\binom{51}{3}&=\binom{52}{4},\\ \binom{4}{4}+\binom{5}{4}+\cdots+\binom{51}{4}+\binom{52}{4}&=\binom{53}{5},\end{aligned}\tag{1.1}$$

osv. Dette er en kombinatorisk formel med ulike navn: skriv opp de første rader i Pascals trekant; summér i vei langs en diagonal; og vips-vops svaret befinner seg et skritt til høyre i neste rad (så «ishockeykølle-formelen» er også tilstrekkelig æppt). Sett Pascal-briller på din nese og sjekk at f.eks. $1+4+10+20=35$, og summerer man lenger gir julestrømpen den første summen over. For den andre summen bruker man den fjerde nedoverdiagonal, med $1+5+15+35+\cdots$, osv.

```
                1
              1   2   1
            1   3   3   1
          1   4   6   4   1
        1   5  10  10   5   1
      1   6  15  20  15   6   1
    1   7  21  35  35  21   7   1
Pascals trekant, med de syv begynnende rader.
```

Formelen kan vises per induksjon, eller direkte kombinatorisk. En fortelling der, som verifiserer formel nr. to over, er som følger. Du skal plassere *fem ess*, blant $N+1=53$ kort, lagt på bordet langs plassene 1 til 53, hvilket kan ordnes på $\binom{53}{5}$ måter. Hvis den største er 5: resten kan ordnes på $\binom{4}{4}=1$ måter. Hvis den største er 6: resten kan ordnes på $\binom{5}{4}$ måter – osv. Den generelle versjonen sier

$$\binom{k}{k}+\binom{k+1}{k}+\cdots+\binom{m-1}{k}+\binom{m}{k}=\binom{m+1}{k+1}.$$

Jeg vender tilbake til kortstokken (tenk at sønnen til faren til Bayesian Nonparametrics er verdens beste pokerspiller), og lar T være antall trekk frem til dette første ess. Med N kort, hvorav n er ess og $m=N-n$ er ikke-ess, er det ikke vanskelig å finne punktsannsynlighetene $f_j=P(T=j)$, ved å tenke aha!, ett kort om gangen, og suksessive betinginger:

$$\begin{aligned}f_1&=\frac{n}{N}=\frac{4}{52},\\ f_2&=\frac{m}{N}\frac{n}{N-1}=\frac{48}{52}\frac{4}{51},\\ f_3&=\frac{m}{N}\frac{m-1}{N-1}\frac{n}{N-2}=\frac{48}{52}\frac{47}{51}\frac{4}{50},\\ f_4&=\frac{m}{N}\frac{m-1}{N-1}\frac{m-2}{N-2}\frac{n}{N-3}=\frac{48}{52}\frac{47}{51}\frac{46}{50}\frac{4}{49},\end{aligned}$$

osv. Med litt styr for fakultetene blir den generelle formelen

$$\begin{aligned} f_j &= \frac{m}{N}\frac{m-1}{N-1}\cdots\frac{m-(j-2)}{N-(j-2)}\frac{n}{N-(j-1)} \\ &= \frac{m!}{N!}\frac{(N-j+1)!}{(m-j+1)!}\frac{n}{N-j+1} \\ &= \binom{N-j}{n-1}\Big/\binom{N}{n}, \end{aligned} \tag{1.2}$$

for $j = 1, 2, \ldots, m+1$. Som en liten sjekk kan vi jo merke oss

$$f_{47} = \frac{48}{52}\frac{47}{51}\cdots\frac{3}{7}\frac{4}{6}, \quad f_{48} = \frac{48}{52}\frac{47}{51}\cdots\frac{3}{7}\frac{2}{6}\frac{4}{5}, \quad f_{49} = \frac{48}{52}\frac{47}{51}\cdots\frac{3}{7}\frac{2}{6}\frac{1}{5}\frac{4}{4}.$$

Summen av sannsynlighetene skal jo bli 1, hvilket stemmer med

$$\sum_{j=1}^{m+1}\binom{N-j}{n-1} = \binom{n-1}{n-1} + \binom{n}{n-1} + \cdots + \binom{N-2}{n-1} + \binom{N-1}{n-1} = \binom{N}{n},$$

eller $\binom{3}{3} + \cdots + \binom{51}{3} = \binom{52}{4}$, den første julestrømpe i (1.1). Dette er et faktisk et uavhengig sannsynlighetsfordelingsbevis for julestrømpeformelen.

2 Og hva blir så forventningen?

Jostein spør altså om forventningen. La oss sjekke med

$$\begin{aligned} \mathrm{E}\,(N+1-T) &= \sum_{j=1}^{m+1}(N+1-j)f_j \\ &= \sum_{j=1}^{m+1}(N+1-j)\frac{m!}{N!}\frac{(N-j+1)!}{(m-j+1)!}\frac{n}{N-j+1} \\ &= n\frac{m!\,n!}{N!}\sum_{j=1}^{m+1}\frac{(N-j+1)!}{(m-j+1)!\,n!} \\ &= n\frac{1}{\binom{N}{n}}\Big[\binom{n}{n} + \binom{n+1}{n} + \cdots + \binom{N}{n}\Big] = n\frac{N+1}{n+1}, \end{aligned}$$

som for kortstokken bruker $\binom{4}{4} + \cdots + \binom{52}{4} = \binom{53}{5}$, julestrømpe nr. to i (1.1). For forventningen $\xi = \mathrm{E}\,T$ leder dette til $N + 1 - \xi = n(N+1)/(n+1)$ og til den riktig smukke formelen

$$\mathrm{E}\,T = \xi = \frac{N+1}{n+1}, \quad \text{som for kortstokken betyr} \quad \frac{53}{5} = 10.6. \tag{2.1}$$

En annen løsning, korrelert med andre innsikter, benytter det forsåvidt kjente knep, at

$$\mathrm{E}\,T = f_1 + 2f_2 + 3f_3 + \cdots = S_1 + S_2 + \cdots,$$

der $S_j = P(T \geq j) = p_j + p_{j+1} + \cdots$, overlevelsessjansen, som det heter i survival analysis. Og disse S_j er pene nok,

$$S_1 = 1, \quad S_2 = \frac{m}{N} = \frac{48}{52}, \quad S_3 = \frac{m}{N}\frac{m-1}{N-1} = \frac{48}{52}\frac{47}{51}, \quad S_4 = \frac{m}{N}\frac{m-1}{N-1}\frac{m-2}{N-2} = \frac{48}{52}\frac{47}{51}\frac{46}{50},$$

osv. Den generelle formelen blir

$$S_j = \frac{m}{N}\frac{m-1}{N-1}\cdots\frac{m-(j-2)}{N-(j-2)} = \frac{m!\,n!}{N!}\frac{(N+1-j)!}{(m+1-j)!\,n!} = \binom{N+1-j}{n}\Big/\binom{N}{n},$$

altså

$$S_j = P(T \geq j) = \binom{53-j}{4}\Big/\binom{52}{4}, \quad \text{for } j = 1, 2, \cdots, 49$$

for vår faste kortstokk. Da gjenstår det å summere – hvilket blir smukt, idet vi igjen bruker julestrømpen:

$$\mathrm{E}\,T = \binom{53}{5}\Big/\binom{52}{4} = \frac{53}{5} = 10.6.$$

Den generelle formelen, for andre kortstokker og andre antall ess, er som i (2.1).

Vi merker oss at T blir stokastisk større i N og mindre i n, som får konsekvenser for konfidenskonstruksjonene under.

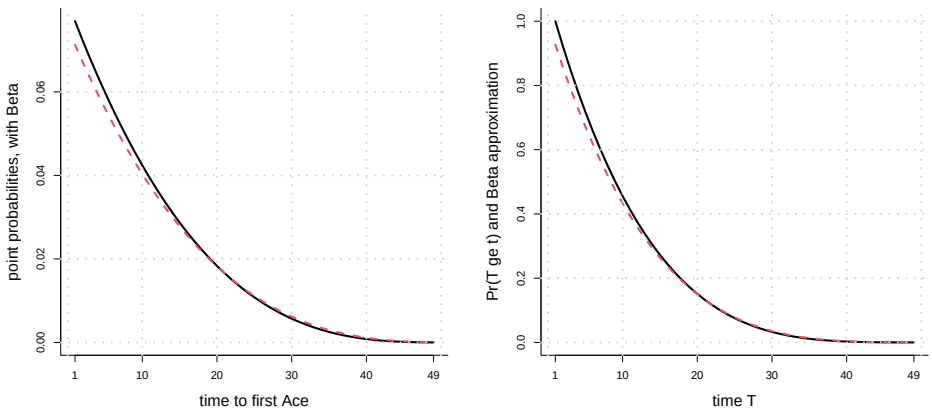

Figure 1: Den eksakte fordelingen til T, tidspunktet for første ess (sort fulle kurve), mot Beta-approksimasjonen (rød stiplet kurve). Venstre: punktsannsynlighetene $f_j = P(T = j)$; høyre: overlevevelsessannsynlighetene $S_j = P(T \geq j)$. Forventningen for approksimasjonen er identisk med den eksakte, $53/5 = 10.6$.

3 En Beta-approksimasjon

For passende store kortstokker, som for vårt hjemlige $N = 52$, er situasjonen «nesten kontinuerlig». Se på $B_N = T/(N+1)$, den normaliserte ventetid. Da har vi

$$P(B_N \geq x) = P(T \geq (N+1)x) = \binom{N+1-(N+1)x}{n}\Big/\binom{N}{n}.$$

Men denne, kan man vise, med hammeren fra Stirling 1730 eller uten, konvergerer mot $(1-x)^n$, for $x \in (0,1)$. Dette viser konvergens i fordeling,

$$B_N \to_d \mathrm{Beta}(1, n),$$

altså en Beta$(1,4)$, med forventning $1/5$, for våre $n=4$ ess. Figuren 1 viser at approksimasjonen er riktig god, for $(N,n)=(52,4)$. Til venstre ser vi den eksakte f_j, mot approksimasjonen

$$f_j^* = P\{(N+1)\,\mathrm{Beta}(1,n) \in [j \pm \tfrac{1}{2}]\} = \mathrm{be}(j/(N+1),1,n)/(N+1),$$

der $\mathrm{be}(x,1,n) = n(1-x)^{n-1}$ er Beta-tettheten, med kumulativ $\mathrm{Be}(x,1,n) = 1-(1-x)^n$. Til høyre har vi den eksakte S_j, mot approksimasjonen

$$P((N+1)\,\mathrm{Beta}(1,n) \geq j) = \{1 - j/(N+1)\}^n.$$

Siden vi alle kan forventningen i en Beta-fordeling, får vi svaret

$$\mathrm{E}\,T \approx (N+1)\,\mathrm{E}\,\mathrm{Beta}(1,n) = \frac{N+1}{n+1},$$

en gang til; forventningen til approksimasjonen er den korrekte forventning. Vi kan også høste ut en god approksimasjon til variansen til T, fra $T=(N+1)\,B_N$ og Beta-fordelingen.

4 Kan vi gjette hva N eller n er, etter å ha sett T?

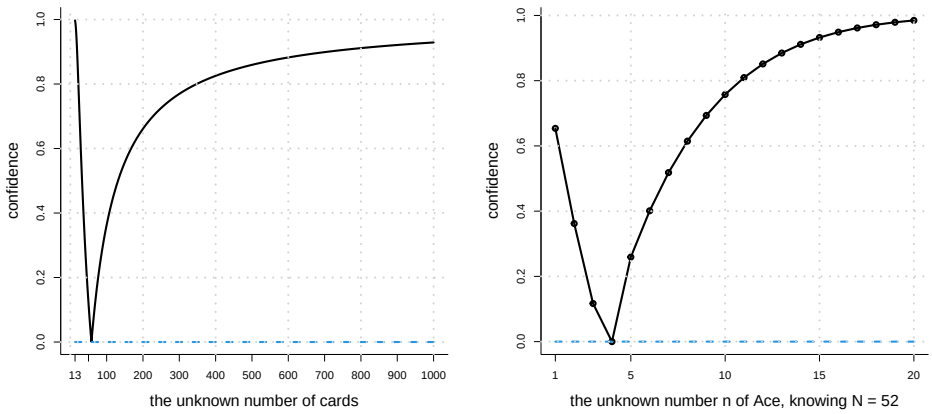

Figure 2: To konfidenskurver, etter å ha trengt $T = t_{\mathrm{obs}} = 10$ trekk før mitt første ess. Venstre: jeg vet det er $n=4$ ess, men estimerer N, antall kort. Punktestimatet er $\widehat{N}=58$, og konfidensen er meget skjev. Høyre: jeg vet det er $N=52$ kort, men estimerer n, antall ess. Punktestimatet er $\widehat{n}=4$.

Det er situasjoner, utenfor kasinoer og kortsalonger, der man leter blant individer eller gjenstander inntil man finner den første av n gitte interessante eller tilstrekkelig aktverdige – men der man ikke vet N. Tilsvarende er det situasjoner der man vet hvor stor en kortstokk eller en populasjon er, med N individer, men der man ikke vet hvor mange Spesielle Kort eller individer det er. Hvordan kan vi estimere disse ukjente tall, med konfidens?

Først: gitt n ess, men antall kort N ukjent. Anta $T = t_{\mathrm{obs}} = 10$, i et slikt enkelteksperiment, og at vi vet at det er $n=4$ ess (hvem kan glemme EM 1976, EM 1977). Momentmetoden setter $T = t_{\mathrm{obs}}$ lik foventningen $(N+1)/(n+1)$, med svar $\widetilde{N} = (n+1)t_{\mathrm{obs}} - 1 = 49$; godt.

En bedre metode, viser det seg, ut fra forventet presisjon, er via Schweder og Hjort (2016, Chs. 3–4) og deres konfidensfordelinger (CD-er). Vi kjører

$$C_A(N) = P_N(T \geq t_{\rm obs} = 10) = \binom{N+1-t_{\rm obs}}{n} \Big/ \binom{N}{n} \quad \text{for } N \geq N_{\rm lower} = t_{\rm obs} + n - 1 = 13$$

som kumulativ konfidens (den er monotont voksende i N), og fra denne konfidenskurven

$$\mathrm{cc}_A(N) = |1 - 2\,C_A(N)|.$$

Denne er vist i Figur 2, venstre panel. Medianestimatet er $\widehat{N} = 58$, og kurvens dramatiske høyreskjevhet gjenspeiler den store usikkerheten, basert på ett enkelteksperiment. Alle konfidensintervaller, f.eks. av grad 90% eller 80%, lar seg greit avlese.

Så: gitt N kort, men antall ess n ukjent. Metodikken pekt på over, fra Schweder og Hjort (2016), kan brukes også her, men gir andre typer kurver. Konfidensfordelingen her, som virker «motsatt vei», bruker at T er stor når n er liten, og tar formen

$$C_B(n) = P_n(T \leq t_{\rm obs} = 10) = \binom{N+1-t_{\rm obs}}{n} \Big/ \binom{N}{n} \quad \text{for } n = 1, 2, 3, \ldots.$$

Denne er monotont voksende i n. Igjen med $T = t_{\rm obs} = 10$, som illustrasjon, gir konfidenskurven $\mathrm{cc}_B(n) = |1 - 2\,C_B(n)|$, vist frem i Figur 2, høyre panel, med medianestimat $\widehat{n} = 4$.

Metoder relatert til dem jeg kort har illustrert her, for ess i en kortstokk, kan bli brukt til å estimere dyrebestander, samt til å telle de utellede, som i faktisk flere av Statistical Stories i Hjort og Stoltenberg (2026): les og arbeid deg gjerne gjennom Story #31 (How many Clethrionomys glareoli); Story #32 (How many deer in the forest); Story #42 (How many Abel envelopes in 1902); Story #66 (How many were killed in Guatemala, 1978–1995).

5 Andre gang, tredje gang, fjerde gang

Josteins fokus er T_1, første gang vi ser et ess, og vi har funnet punktsannsynligheter $f(t, N, n)$ fra (1.2) for $t = n, n+1, \cdots, m+1$. Men hva med T_2, andre gang vi ser et ess? Gitt $T_1 = t_1$ er situasjonen som tidligere, for gapet $T_2 - T_1$, modulo de endrede parametere, som nå er $N - t_1$ kort med $n - 1$ ess. Altså har vi

$$\mathrm{E}\,\{(T_2 - T_1)\,|\,t_1\} = \frac{N+1-t_1}{n},$$

som per dobbeltforventning medfører

$$\mathrm{E}\,(T_2 - T_1) = \frac{N+1}{n} - \frac{N+1}{n+1}\frac{1}{n} = \frac{N+1}{n}\Big(1 - \frac{1}{n+1}\Big) = \frac{N+1}{n+1},$$

altså samme ξ som for hovedformelen (2.1). Det samme finner vi for $T_3 - T_2$, som for gitt $(T_1, T_2) = (t_1, t_2)$ må ha fordelingen $f(g_3, N - t_2, n - 2)$. Vi kan slå fast, faktisk uten å ha regnet ordentlig på sannsynlighetsfordelingene for T_2, T_3, T_4, at gapene

$$G_1 = T_1, \quad G_2 = T_2 - T_1, \quad G_3 = T_3 - T_2, \quad G_4 = T_4 - T_3, \ldots$$

har samme forventning ξ. For den faste kortstokk har vi

$$(\mathrm{E}\,T_1, \mathrm{E}\,T_1, \mathrm{E}\,T_3, \mathrm{E}\,T_4) = (53/5, 106/5, 159/6, 212/5) = (10.6, 21.2, 31.2, 42.4),$$

altså kortposisjonene $1, 2, \ldots, 52, 53$ delt i fem like store biter.

Gapene $G_2, G_3, \ldots$ må faktisk ha presis samme marginale fordeling som det første, $G_1 = T_1$. De er ikke uavhengige, men ombyttbare («exchangeable», som de Finetti ville sagt, når han snakket engelsk). Med notasjon $f(t, N, n)$ for $P(T_1 = t)$, med full kortstokk N og alle ess n på plass, som i (1.2), kan man altså vise, for $G_2 = T_2 - T_1$ marginalt, at

$$f(g_2) = P(T_2 - T_1 = g_2) = \sum_{t_1} f(t_1, N, n) f(g_2, N - t_1, n - 1) = f(g_2, N, n).$$

Vi kan altså se på ess-tidene som

$$T_1 = G_1, \quad T_2 = G_1 + G_2, \quad T_3 = G_1 + G_2 + G_3, \quad T_4 = G_1 + G_2 + G_3 + G_4,$$

der $G_1, \ldots, G_4$ har samme fordeling (men de er innbyrdes negativt korrelerte). Dette passer jo perfekt med forventningsverdiene $(\xi, 2\xi, 3\xi, 4\xi)$.

6 Den kontinuerlige modell (når kortstokken er stor)

Vi har klart å analysere situasjonen godt nok, via de eksakte fordelinger for ess-tidene $T_1, T_2, \ldots$, osv. Svarene er en anlese kronglete, men forståelige og anvendbare, som vist med konfidenskurvene $C_A(N)$ og $C_B(n)$ over. Hele situasjonen blir penere, og lettere å analysere, matematisk og statistisk, når vi skrider til en kontinuerlig grense. Approksimasjonene som følger av det under vil være gode, så lenge N er passende stor (52 holder sikkert) og n ikke er stor i forhold til denne N.

Nøkkelen ligger i å oversette ess-tidene $T_1 < T_2 < T_3 < \cdots$ til brøkdeler

$$B_1 = T_1/(N+1), \quad B_2 = (T_2 - T_1)/(N - T_1), \quad B_3 = (T_3 - T_2)/(N - T_2), \ldots.$$

Hver av disse er uavhengige, gitt de foregående, og av samme karakter, men med stadig mindre kortstokker. Siden vi allerede har skjønt at $B_1 \sim \mathrm{Beta}(1, n)$, forstår vi også hvordan det går med de andre, gitt de foregående; når N vokser vil

$$B_1 \sim \mathrm{Beta}(1, n), \; B_2 \,|\, T_1 \sim \mathrm{Beta}(1, n-1), \; B_3 \,|\, (T_1, T_2) \sim \mathrm{Beta}(1, n-2), \ldots,$$

der B_j gitt de foregående er en ny uavhengig Beta. Dette kan også transformeres til de skalerte gap

$$\begin{aligned}
D_1 &= T_1/(N+1), \\
D_2 &= (T_2 - T_1)/(N+1), \\
D_3 &= (T_3 - T_2)/(N+1), \ldots \\
D_n &= (T_n - T_{n-1})/(N+1), \\
D_{n+1} &= (N - T_n)/(N+1),
\end{aligned}$$

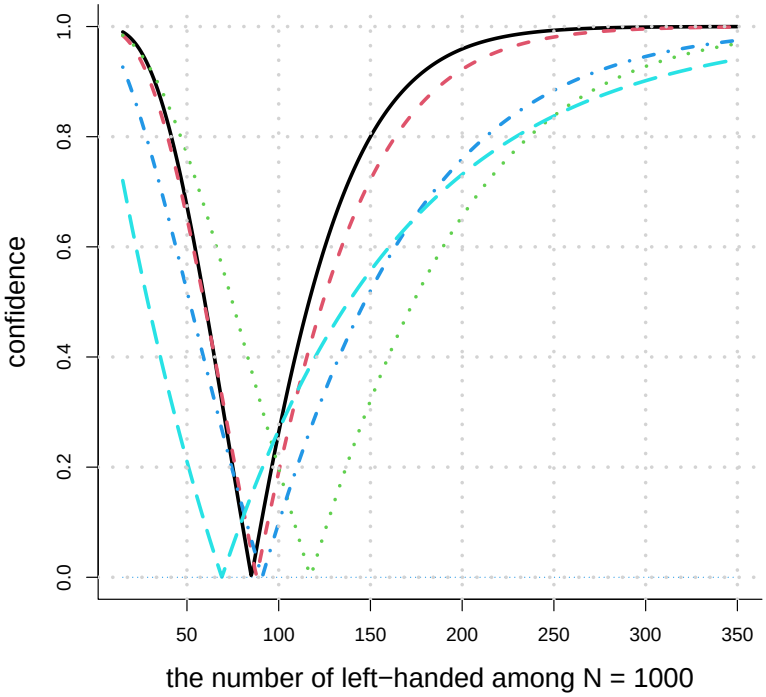

Figure 3: Hvor mange kjevhendte er det, av $N = 1000$ personer, gitt at jeg hilser på én om gangen, og finner de kjevhendte som nummer $(T_1, T_2, T_3, T_4, T_5) = (10, 18, 22, 39, 50)$? Her er konfidenskurver 1, 2, 3, 4, 5, etterhvert som mer informasjon kommer inn, og der den femte er mest informativ (sort kurve), med $\widehat{n} = 85$, og 90 prosent intervall $[33, 172]$.

mer $D_1 + \cdots + D_n + D_{n+1} = 1$. Det er da en fin øvelse med transformasjonssetningen, komplett med en Jacobi-determinant, å vise fra Beta-bitene over at

$$(D_1, D_2, \ldots, D_n, D_{n+1}) \sim \mathrm{Dir}(1, 1, \ldots, 1, 1),$$

en flat Dirichlet-fordeling. Men dette er sannelig identisk med fordelingen til $U_1, U_2, \ldots, U_n, 1 - U_n$, ordningsobservatoren for et i.i.d.-sample av størrelse n fra den uniforme fordelingen på enhetsintervallet. Altså har vi arbeidet oss frem til den forståelse at

$$(T_1, \ldots, T_n) \sim (N + 1)(U_1, \ldots, U_n),$$

skalerte ordningsobservatorer, og dermed «vet vi det meste» om hele ess-plasseringsprosessen.

Vi har for eksempel at $U_k \sim D_1 + \cdots + D_k \sim \mathrm{Beta}(k, n + 1 - k)$, med

$$\mathrm{E}\, U_k = \frac{k}{n+1}, \quad \mathrm{Var}\, U_k = \frac{k}{n+1}\Big(1 - \frac{k}{n+1}\Big)\frac{1}{n+2},$$

som straks gir (gode approksimasjoner til) forventning og varians for $T_k = (N + 1)U_k$. Spesielt, enda en gang, siden Jostein vil ha flere løsningsforslag, har vi $\mathrm{E}\, T_k = k(N + 1)/(n + 1)$.

Dette kan vi jo utnytte, på ymist vis, også for estimering og inferens for diverse «hvor mange er det» spørsmål. Anta vi har $N = 1000$ personer i et rom, og jeg hilser på én av gangen

for å gjette meg frem til n, antallet kjevhendte. Anta at de kjevhendte jeg etterhvert treffer er $(T_1, T_2, T_3, T_4, T_5) = (10, 18, 22, 39, 50)$ (hvorpå jeg stopper å sjekke med de resterende 950). Hva er et godt estimat på antall kjevhendte, og et godt konfidensintervall?

Jo, for det første er tettheten til de fem første uniforme ordningsobservatorer lik

$$h(u_1, \ldots, u_5) = 1 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \binom{n}{5} (1 - u_5)^{n-5} \quad \text{for } u_1 < \cdots < u_5 < 1.$$

Av dette ser vi at U_5 er suffisient. Skalerer vi tilbake til de 1000 personer i rommet lærer vi at $T_5 = 50$ er suffisient informasjon. Vi lager dernest den suffisiensbestemte konfidensfordelingen

$$C(n) = P_n(T_5 \leq 50) = P((N+1)\,\mathrm{Beta}(5, n-4) \leq 50) = \mathrm{Be}(50/(N+1), 5, n-4),$$

med den kumulative Beta. Figur 3 viser konfidenskurven $\mathrm{cc}(n) = |1 - 2\,C(n)|$, med medianestimatet $\widehat{n} = 85$, og med nokså skjevt 90 prosent intervall $[33, 172]$.

Mine studenter i de ulike kurs har måttet vende seg til «ekstra randbemerkninger», og her er en slik. Vi kan skrive

$$U_1 = B_1, \quad U_2 = U_1 + (1 - B_1)B_2, \quad U_3 = U_2 + (1 - B_1)(1 - B_2)B_3, \ldots,$$

som med en pinnebrekkingsalgoritme (vi brekker en pinne i to; så den gjenværende delpinne i to; osv.), med uavhengige $B_1 \sim \mathrm{Beta}(1, n)$, $B_2 \sim \mathrm{Beta}(1, n-1)$, $B_3 \sim \mathrm{Beta}(1, n-2)$, osv. Så

$$U_k = D_1 + \cdots + D_k, \quad \text{med } D_i = (1 - B_1) \cdots (1 - B_{i-1})B_i,$$

som betyr at vi kan representere U_k, og dermed våre venner T_k fra kortstokkproblemet, ved hjelp av produkter av uavhengige Beta-variable. Dette er en slags endelig-sum-produkt-representasjon for en Dirichlet-vektor versjon av Sethuramans uendelig-sum-produkt-representasjon av Dirichlet-prosessen, kjent & kjært fra Bayesian Nonparametrics.

Referanser

N.L. Hjort, C.C. Holmes, P. Müller, S.G. Walker (2010). *Bayesian Nonparametrics.* Cambridge University Press.

N.L. Hjort (2008). Korrelasjonen mellom gjennomsnitt og median: essay om Jostein Lillestøls pusleri #23 i Tilfeldig Gang nr. 1. *Tilfeldig Gang* nr. 2, 22–25.

N.L. Hjort (2026). Counting the uncounted: How many were killed in Guatemala, 1978-1995? *Proceedings of the International Workshop on Statistical Modelling*, Oslo, July 2026, invited talk.

N.L. Hjort, E.Aa. Stoltenberg (2026a). Probability Proofs for Stirling (and More): The Ubiquitous Role of $\sqrt{2\pi}$. *American Statistician*, **80**, 405–412.

N.L. Hjort, E.Aa. Stoltenberg (2026b). *Statistical Inference: 600 Exercises and 100 Stories.* Cambridge University Press.

T. Schweder, N.L. Hjort (2016). *Confidence, Likelihood, Probability.* Cambridge University Press.