Om bingospil

Frank Bengtson

Nygårdterrasserne 295 B
DK-3520 Farum
frank.bengtson@get2net.dk

Indledning

Bingospil er et tallotteri, der afsluttes med en eller flere vindere. Hver spiller har en eller flere plader med forskellige tal i felter. Tallene udtrækkes et ad gangen, og spillerne markerer de udtrukne tal på pladerne. Den der først får udtrukket en plades tal, eller et specielt mønster, har vundet, men det kan ske, at flere mangler det samme tal. Vinderen er så den, der først råber *BANKO!*, hvis der spilles lokalt, ellers er der flere vindere. Ofte spilles flere spil i serie.

En nærmere undersøgelse af spil og plader kan byde på en del interessante og underholdende mærkværdigheder, hvis man er til den slags. I artiklen søges følgende spørgsmål besvaret ved beregninger og ved en computermodel:

- Hvordan afhænger antal træk til banko af antal plader i spillet?
- Hvor mange plader er der i alt?
- Kan pladerne laves, så antal træk til banko ændres?
- Hvor mange plader har de samme tal?
- Hvor ofte udtrækkes mere end en plade?

Danske Spil vurderer, at der i Danmark hver uge spilles bingo for 25 mill. kroner af 250.000 spillere. Man kunne observere lokale spil med op til nogle tusinde plader, for at få et indtryk af sammenhængen mellem antal plader og antal træk til bingo. Der afholdes også TV-bingoshows med over ti millioner plader, der kun kan afvikles ved hjælp af computere, hvor pladerne er lagret.

1 RBINGO rektangulære plader

En plade har tre rækker og ni søjler, der skal være fem tal i hver række, mindst et tal i hver søjle, og søjletallene er i rækkefølge. Tal i søjler: 1: 1-9, 2: 10-19, 3: 20-29, 4: 30-39, 5: 40-49, 6: 50-59, 7: 60-69, 8: 70-79 og 9: 80-90. Søjletallene kan så vælges ud af: 9,10,10,10,10,10,10,10,11 tal. Pladen her har *søjletype* 111322122.

	10	28	31	47	55			
			34	49	58		71	83
6			36			63	73	85

Observation 1 *Ses bort fra ordningen af søjlerne, har plader fire mulige søjletyper 1,1,1,2,2,2,2,2,2; 1,1,1,1,2,2,2,2,3; 1,1,1,1,1,2,2,3,3 og 1,1,1,1,1,1,3,3,3. Plader har forskellige tal for forskellige søjletyper og søjleplaceringer. Der er tale om et plademønster og et antal mulige taludfyldninger for hvert mønster.*

1.1 Antal plader med forskellige tal

Der kan laves op til seks plader uden fælles tal. Tallene må bruges igen, når flere end seks plader skal laves. Har blot to plader et tal fælles, kan begge være udfyldt, når det fælles tal trækkes. 15 tal kan vælges ud af 90 på $\binom{90}{15} \approx 4.6 \times 10^{16}$ måder, når der ikke stilles specielle krav til de valgte tal. I det følgende regnes med ti tal i hver søjle, ønskes mere præcise beregninger, skal der regnes med ni tal i første søjle og elleve tal i sidste. Niels Andersen: *Hvor mange bankoplader er der?* [A] s.3 har beregnet tallet til 6080.082602.343750.

Antal tal i søjler og tilsvarende antal plader er ca.

$$
\begin{array}{lcll}
111222222 & : & 10^3\binom{10}{2}^6\binom{9}{6} & \approx & 6.98\ 10^{14} \\
111122223 & : & 10^4\binom{10}{2}^4\binom{10}{3}\binom{9}{4}\binom{5}{4} & \approx & 3.10\ 10^{15} \\
111112233 & : & 10^5\binom{10}{2}^2\binom{10}{3}^2\binom{9}{5}\binom{4}{2} & \approx & 2.20\ 10^{15} \\
111111333 & : & 10^6\binom{10}{3}^3\binom{9}{6} & \approx & 1.45\ 10^{14} \\
\text{sum} & : & & \approx & 6.1\ 10^{15}
\end{array}
$$

Dvs. ca. 13% af antal talsæt med 15 tal uden restriktioner kan optræde som plader. Antallet er så stort, at selv om pladerne laves af et program en efter en, er det meget lidt sandsynligt med to ens plader. Man kommer ikke i nærheden af det største antal plader, heller ikke i det virkelige spil. Man kunne have flere ens plader i et spil, men det ville nok være vanskeligt at finde vinderen, hvis alle plader var ens.

1.2 Seks plader med alle tal *liniebingo* eller *Generalized Line Bingo*

Liniebingo er et rektangulært skema f.eks. 15×6 med alle tal, det svarer til seks plader med alle tal. Det er ligegyldigt, hvordan tallene er fordelt, alle skal bare være med. Antal træk til en plade er udtrukket kan ikke blive mere end 85, idet der kan mangle et tal fra hver plade, og når det 85.tal trækkes, vil en plade være udfyldt.

Hos D.M. Bloom & K. Suman: *Generalized Line Bingo, American Mathematical Monthly* [BS] , findes et ret enkelt udtryk for beregning af middel antal træk, til en eller flere søjler er udtrukket. Middel antal træk til j søjler

$$E_j = mn \prod_{i=j}^{n-1} mi/(mi+1)$$

n søjler

m rækker

1	16	31	46	61	76
2	17	32	47	62	77
.	.	.	.	.	.
14	29	44	59	74	89
15	30	45	60	75	90

Middel antal træk til første fem tal, en række ud af 18

$$E_1 = (18 \times 5)\frac{5}{6}\frac{10}{11}\cdots\frac{85}{86} \quad = 46.7$$

Middel antal træk til første 15 tal, en plade ud af seks

$$E_1 = (6 \times 15)\frac{15}{16}\frac{30}{31}\frac{45}{46}\frac{60}{61}\frac{75}{76} \quad = 77.5$$

1		21	33			61		80	4	17		30	41		62		
	11			40	50		71	85		19	20			51		72	89
	12	22		46		69	74				29		49	55	65	77	
	14	23	31			63		81			24	32			64	76	82
2			35	45	52		73		7	16		34		53			86
3	15				58	67	79				27		42		66	78	87
5		25	36				75	83	8	13	28				60		84
		26	38	43	54			88	9			37	44	57		70	
6	10			47	56	68				18		39	48	59			90

I en computermodel giver seks plader med alle tal 77.5, og seks vilkårlige plader giver 78.3.

1.3 Plader med forskellige tal og mønstre

Hvis plader regnes for forskellige, når de samme tal rykkes op eller ned, skal der regnes med mønstre. Her laves en optælling, en programmeringsløsning kan også laves, og hos [A] findes en beregning med flere elegante anvendelser af frembringende funktioner.

1. For plader med tre søjler med tre tal skal der i seks søjler være mindst et tal. Man har $\binom{6}{2}\binom{4}{2} = 15 \times 6 = 90$, men for at bruge princippet i tælleformlen

$$\left|\begin{array}{cccccc|ccc} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{array}\right|,$$

$$\binom{6}{2}^3 - \binom{6}{5}\binom{5}{2}^3 + \binom{6}{4}\binom{4}{2}^3 - \binom{6}{3}\binom{3}{2}^3 + \binom{6}{2}\binom{2}{2}^3 = 90.$$

   De tre søjler kan vælges på $\binom{9}{3} = 84$ måder, dvs. $90 \times 84 = 7560$ i alt.

2. Med to søjler med tre tal har man

$$\left|\begin{array}{ccccc|cc|cc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{array}\right|,$$

$$\binom{7}{3}^3 - \binom{7}{6}\binom{6}{3}^3 + \binom{7}{5}\binom{5}{3}^3 - \binom{7}{4}\binom{4}{3}^3 + \binom{7}{3}\binom{3}{3}^3 = 5670,$$

   men syv gange vil der være en søjle med tre, som sammen med de to til højre giver tre søjler med tre. Der er så $5670 - 7 \times 90 = 5040$ med to søjler med tre til højre, og de to søjler kan vælges ud af ni søjler, dvs. $5040 \times \binom{9}{2} = 5040 \times 36 = 181440$ i alt.

3. Med en søjle med tre

$$\left|\begin{array}{cccc|cccc|c} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \end{array}\right|,$$

$$\binom{8}{4}^3 - \binom{8}{7}\binom{7}{4}^3 + \binom{8}{6}\binom{6}{4}^3 - \binom{8}{5}\binom{5}{4}^3 + \binom{8}{4}\binom{4}{4}^3 = 87570,$$

men der vil være en og to søjler ud af otte, som sammen med søjlen til højre giver to og tre søjler med tre. Der må være $87570 - \binom{8}{1}5040 - \binom{8}{2}90 = 87570 - 8 \times 5040 - 28 \times 90 = 87570 - 40320 - 2520 = 44730$. Søjlen med tre kan vælges ud af ni, så der er $\binom{9}{1} \times 44730 = 402570$.

4. Af tælleformlen får man det samlede antal med mindst et tal i hver søjle

$$\left|\begin{matrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{matrix}\right|,$$

$$\binom{9}{5}^3 - \binom{9}{8}\binom{8}{5}^3 + \binom{9}{7}\binom{7}{5}^3 - \binom{9}{6}\binom{6}{5}^3 + \binom{9}{5}\binom{5}{5}^3 = 735210$$

Antal 111222222 er så 735210-(7560+181440+402570)=143640.

Tabel 1: Antal plademønstre for søjletyper

søjle typer	antal mønstre	antal søjleplac.	antal mønstre pr. søjleplac.
111222222	143640	$\binom{9}{3} = 84$	143640/84=1710
111122223	402570	$\binom{9}{1}\binom{8}{4} = 630$	402570/630=639
111112233	181440	$\binom{9}{2}\binom{7}{2} = 756$	181440/756=240
111111333	7560	$\binom{9}{3} = 84$	7560/84=90
i alt	735210		

Femten kan vælges ud af syvogtyve på $\binom{27}{15} = 17383860$ måder, dvs. $735210/17383860 = 4.2\%$ kan optræde som plademønstre.

Antal plademønstre med tal og antal plader med samme tal middel

$$\begin{array}{lcll}
111222222 & : & 143640 \times 10^3 \binom{10}{2}^6 & \approx \; 1.19\ 10^{18} \\
111122223 & : & 402570 \times 10^4 \binom{10}{2}^4 \binom{10}{3} & \approx \; 1.98\ 10^{18} \\
111112233 & : & 181440 \times 10^5 \binom{10}{2}^2 \binom{10}{3}^2 & \approx \; 5.29\ 10^{17} \\
111111333 & : & 7560 \times 10^6 \binom{10}{3}^3 & \approx \; 1.31\ 10^{16} \\
\text{sum} & : & & \approx \; 3.7\ 10^{18} \\
\text{middel med samme tal} & : & 3.7\ 10^{18}/6.1\ 10^{15} & \approx \; 604
\end{array}$$

For helt præcise tal skal regnes med ..., hos [A] s.6 er *...antallet af lovlige banko-plader...* beregnet til 3.669688.706217.187500.

1.4 Plader med samme tal

To rækker kan ikke have fælles tal. Det er oplagt med en eller flere søjler med tre tal.

To af 90 RBINGO plader med samme tal, 15 øverste rækker, 45 forskellige rækker, 15 øverste rækkepar og 45 rækkepar i alt.

	12		34			65	70	81
1					58	68	75	87
		27		42		69	77	90

1			34			65	70	81
	12				58	68	75	87
		27		42		69	77	90

For søjletype 112222221 er 126 rækker øverst og nederst. Der er 1452 mulige rækker $\binom{6}{5}2^5 + \binom{6}{4}\binom{3}{1}2^4 + \binom{6}{3}\binom{3}{2}2^3 + \binom{6}{2}\binom{3}{3}2^2 = 192 + 720 + 480 + 60 = 1452$ og dermed 1200 i midten. F.eks. kan rækken 4,21,33,61,82 som er på pladerne nedenfor kun være foroven.

$$\left|\begin{array}{ccccccccc} 4 & - & 21 & 33 & - & - & 61 & - & 82 \\ - & 11 & 22 & - & 40 & 50 & - & 71 & - \\ - & - & - & 37 & 46 & 53 & 69 & 74 & - \end{array}\right|\left|\begin{array}{ccccccccc} - & 11 & - & - & 40 & 50 & - & 71 & - \\ 4 & - & 21 & 33 & - & - & 61 & - & 82 \\ - & - & 22 & 37 & 46 & 53 & 69 & 74? & - \end{array}\right|,$$

Tallet 74 kan ikke placeres. Der er 1452 forskellige rækker til 1710 plader med tre på hver, dvs. hver række skal i middel bruges $1710 \times 3/1452 = 5130/1452 \approx 3.5$ gang.

Seks af 1710 RBINGO plader med samme tal, 1452 forskellige rækker, 126 øverst og nederst og 1200 i midten. 126 øverste rækkepar og 1452 rækkepar i alt.

Rækker foroven og forneden bruges to gange (15 plader), seks gange (60 plader), tyve gange (45 plader) og halvfjerds gange (6 plader). Rækker i midten en gang (810 plader) to gange (270 plader) og tre gange (120 plader). Spil med disse 1710 plader vil være som spil med en plade, 126 øverste rækker og 1452 rækker og rækkepar.

4		21	33			61		82
	11	22		40	50		71	
			37	46	53	69	74	

4		21	33			61		82
	11		37	40	50		71	
		22		46	53	69	74	

4		21	33			61		82
	11			40	50	69	71	
		22	37	46	53		74	

4		21	33			61		82
		22		40	50	69	71	
	11		37	46	53		74	

4		21	33			61		82
		22	37	40	50		71	
	11			46	53	69	74	

4		21	33			61		82
			37	40	50	69	71	
	11	22		46	53		74	

1.5 Antal rækker, rækkepar og plader

I forbindelse med antal træk til banko er det kun tal på pladen, der har betydning, og det aktuelle antal plader er $6.1\ 10^{15}$. Det kan nu overvejes, hvordan pladerne kan laves i modellen. Laves plader en ad gangen, vælges en ud af $3.71\ 10^{18}$, og der er i middel 604 med ens tal. Der vil derfor næppe være to plader med samme tal for de pladeantal, der kan køres i modellen eller for virkelige spil.

Der er fem tal i hver række og de vælges ud af ni søjler med ca. ti i hver. Antal rækker vil så være ca. $\binom{9}{5}10^5 = 126 \times 10^5 \approx 1.3\ 10^7 = 13000000$, præcist 12565000. Der kan højst være $\binom{90}{5} = 43949268 \approx 4.4\ 10^7$, dvs. ca. 30% vil optræde som rækker.

Man kunne lave et globalt spil, hvor beboerne på jorden har plader med forskellige tal. Regnes med syv milliarder beboere skal bruges $7\ 10^9 \times 3 = 2.1\ 10^{10}$ rækker, dvs. rækker skal i middel bruges 1671 gange.

Hvis man ville lave et spil, hvor alle plader var med, skulle hver jordbo administrere $6.1\ 10^{15}/7\ 10^9 \approx 871000$ plader. Rækker skulle i middel bruges 485 mill. gange, og sandsynligheden ville være 13% for første plade på 15 taltræk. Med alle $3.7\ 10^{18}$ plader skulle man klare $3.7\ 10^{18}/7\ 10^9 \approx 530$ mill plader. I middel skal hver række bruges ca. 880 mia. gange, så spørgsmålet er, om spillerne ser pladerne som forskellige. Sandsynligheden vil være 13% for 90, 240, 639 eller 1710 plader på første femten træk.

1.6 Antal plader med bingo i samme spil

Øverst er antal plader med bingo i samme spil optalt i en computermodel, og nederst er tal fra et TV-bingospil konsulentfirma `http://www.tv-bingo.com`.

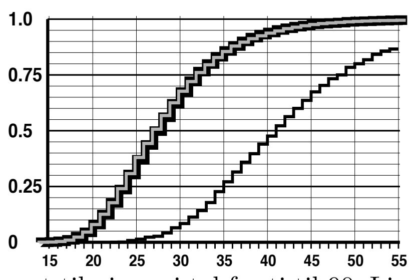
egen computermodel	antal plader			
antal plader med bingo	10	1.000	10.000	100.000
1	0.89	0.86	0.83	0.81
2	0.10	0.13	0.16	0.17
3 eller flere	0.01	0.01	0.01	0.02

	antal plader				
ant. pl.med bingo	500.000	1.000.000	3.000.000	5.000.000	10.000.000
1	0.79	0.77	0.73	0.76	0.75
2	0.16	0.17	0.19	0.15	0.17
3 eller flere	0.05	0.06	0.08	0.09	0.08

1.7 Valg af tal til plader

Tal til plader var ca. 13% af mulige talsæt med 15 tal. I både computer- og sandsynlighedsmodel kan det overvejes, om man kan se bort fra det specielle valg af tal til plader. Kan man blot vælge 15 forskellige tal? Kan man endnu enklere blot vælge 16 eller 17 tal med tilbagelægning?

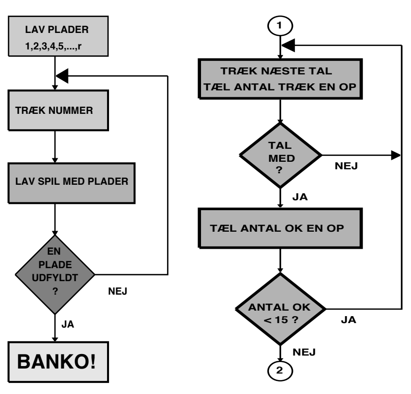
Fordeling for antal plader kraftig graf og antal gange 15 vilkårlige tal med tilbagelægning skal vælges tynd graf, for at alle tal fra 1 til 90 er med, er ens. Vælges 27 plader, er sandsynligheden ca. 50% for, at alle tal er med. Vælges 41 plader, er sandsynligheden ca. 95% for, at alle tal er med. **Pladetallene er kun lidt specielle.**



Højre graf svarer til, at der vælges seks tal fra et til ni og ni tal fra ti til 90. Ligner udvidet *coupon collector* problem.

1.8 Computermodel

I det virkelige spil behandles pladerne samtidigt eller parallelt. Med mange plader er samtidigheden besværlig at programmere. Laver man f.eks. et spil med ti plader kan man i stedet lave ti spil med en plade ad gangen og samme træknummerliste. Antal træk til banko er så det mindste af de ti antal træk til banko. Det er langt enklere at programmere en seriel behandling af pladerne.



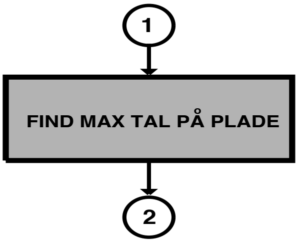
Figur 1: Rutediagrammer for hele spillet og spil med en plade. Herudfra kan ret let laves programmer.

For spil med en plade kan en bestemt plade vælges og taltræklisterne varieres, eller en fast taltrækliste vælges og pladen varieres. Både plader og taltræklister kan selvfølgelig også varieres. For at bestemme middelværdi og evt. (model)fordeling, beregnes middelværdien for mange spil.

For at lave et spil med r plader laves r spil med forskellige plader og en fast taltrækliste, det mindste antal træk for r spil findes ved at sammenligne den aktuelle værdi med den hidtil mindste. Hvis der er r plader og antal træk til banko på hver plade betegnes med $X_1, \ldots, X_r$, er antal træk til banko for r plader $\min\{X_1, \ldots, X_r\}$, og middelværdien for en plade $\overline{x} = (X_1 + \cdots + X_r)/r$. Med en plade og r forskellige taltræklister får man samme udtryk. I det første tilfælde vælges r gange q tal til plader, i det andet vælges r taltræklister (r gange n tal), hvilket er mere besværligt.

1.9 Stor og lille model med fast taltrækliste og minmax

Med en fast plade vil forskellige træklister i middel give samme ventetid. Vælges i stedet en fast trækliste, f.eks. den særlig pæne $1, 2, 3, \ldots, n$, kan pladetallene vælges tilfældigt. Her forsvinder plademønsteret afsn. 1 s. 80, idet ordningen af tallene kan give søjler uden tal.



Sætning 1.1 *Ses bort fra plademønsteret, vil pladens q tal vælges tilfældigt ud af n tal. Antal træk for spil med en plade og r plader kan bestemmes som*

$$\begin{aligned} \textit{antal træk for spil med en plade} &= \max\{T\} \\ \textit{antal træk for spil med } r \textit{ plader} &= \min\{\max\{T_1\}, \ldots, \max\{T_r\}\} \end{aligned}$$

Lille model kan let programmeres, stor model er lidt mere besværlig. Der ser ikke ud til at være forskel på stor og lille model op til 10^6 plader. Med pladeantal over 10^6 bliver tidsforbruget stort. Da der kun er 12565000 rækker kan hele området køres igennem. Middelantaltræk for r plader kan ca. beregnes som

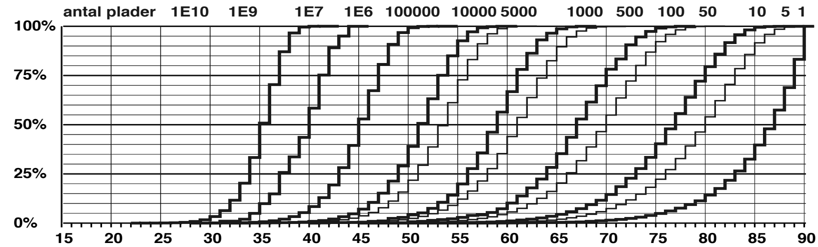
$$m = 85.3\ r^{-0.056} \approx 85\ r^{-0.05}\ \text{, middelantal træk plader,}$$

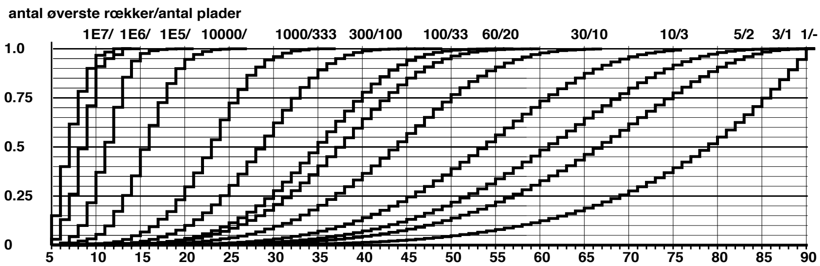
og er lidt mindre end medianerne, da venstre hale er længst for de mindre pladeantal.

1.10 Observerede fordelinger for plader og rækker fra computermodel

1.11 Sandsynlighedsfunktion og fordeling

Ethvert spil vil have et antal træk mellem q og n, og sandsynligheden for et antal træk $P(X = m)$, $q \leq m \leq n$ kan beregnes. Det er en stokastisk variabel med





en *ventetid* eller *livslængde*, hvor sandsynligheden for udtrækning vokser stærkt med antal træk. Antallet af delmængder med q tal valgt ud af m tal er $\binom{m}{q}$, og i forhold til det samlede antal delmængder $\binom{n}{q}$ med q tal ud af n, er det den klassiske definition med *antal gunstige/antal mulige*. Man får så umiddelbart fordelingen [1]

$$(1) \qquad F(m) = \binom{m}{q} \Big/ \binom{n}{q} = \frac{m-q+1}{m+1} F(m+1) = \frac{m(m-1)\cdots(m-q+1)}{n(n-1)\cdots(n-q+1)}$$

$$F(n) = \binom{n}{q} \Big/ \binom{n}{q} = 1; \; F(q) = 1 \Big/ \binom{n}{q}$$

og sandsynlighedsfunktionen

$$P(X=n) = f(n) = \binom{n-1}{q-1} \Big/ \binom{n}{q} = \frac{q}{n}$$

$$P(X=n-1) = f(n-1) = \binom{n-2}{q-1} \Big/ \binom{n}{q} = P(X=n)\frac{n-q}{n-1}$$

$$P(X=q) = f(q) = 1 \Big/ \binom{n}{q}$$

Man har

$$(2) \qquad P(X=m) = \tfrac{q}{m} P(X \leq m)$$

[1] Kendte fordelinger for max for 15 ud af 90, giver straks sandsynlighedsfunktion, middelværdi og varians.

1.12 Middelværdi og varians for en plade

$$
\begin{aligned}
E(X) &= q\,P(X=q)+(q+1)P(X=q+1)+\cdots+n\,P(X=n) \\
&= q\ P(X\le q)+\cdots+q\ P(X\le n) && \text{, af (2)} \\
&= q\ \left(\binom{q}{q}+\cdots+\binom{n}{q}\right)\Big/\binom{n}{q} && \text{, af (1)} \\
&= q\ \binom{n+1}{q+1}\Big/\binom{n}{q} && \text{, af sumformel} \\
(3)\qquad &= q\ \tfrac{n+1}{q+1}
\end{aligned}
$$

Med $n=90$ og $q=15$ er $E(X)=15\frac{90+1}{15+1}=85.31$ og medianen eller 50% fraktilen er 86. For øverste (eller en række) er $q=5$ og $E(X)=5\frac{90+1}{5+1}=75.83$, medianen 78. Ved at regne løs får man

$$Var(X)=\frac{q(n+1)(n-q)}{(q+1)^2(q+2)}=E(X)\frac{n-q}{(q+1)(q+2)}$$

Med $n=90$ og $q=15$ er $\sigma(X)=\sqrt{Var(X)}=4.85$.

1.13 Spil med flere plader

Med r plader stopper spillet, når en eller flere pladers tal er udtrukket. Dette sker med et antal træk fra *antal tal på plade* q til *antal tal i alt* n.

$$Y_r=(Y_r)_{\text{serie}}=(Y_r)_{(1)}=\min\{X_1,\ldots,X_r\}$$

Det forudsættes, at X_k'erne er uafhængige identisk fordelte. Forudsættes X_k'erne uafhængige, vil det svare til r spil med en plade, hvor pladerne vælges med tilbagelægning. En forudsætningen er, at pladerne ikke har et mønster, dvs. der udtrækkes femten forskellige tal.

$$P(Y_r\ge m)=\{P(X_1\ge m),\ldots,P(X_r\ge m)\}=\prod_{k=1}^{r}P(X_k\ge m)=(P(X\ge m))^r$$

$$P(Y_r=m)=P(Y\ge m)-P(Y\ge m+1),\ m+1\le n$$

$$P(Y_r=q)=1-\left(1-1\Big/\binom{n}{q}\right)^r,\ P(Y_r=n)=\left(1-\frac{\binom{n-1}{q}}{\binom{n}{q}}\right)^r=\left(\frac{q}{n}\right)^r$$

$$F_{Y_r}(m)=P(Y_r\le m)=1-(1-F_X(m))^r=1-\left(1-\frac{\binom{m}{q}}{\binom{n}{q}}\right)^r \tag{4}$$

Ved beregning af middelværdien for Y har man

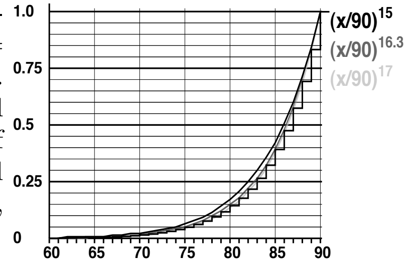
$$E(Y)=q+\left(1-\frac{\binom{q}{q}}{\binom{n}{q}}\right)^r+\cdots+\left(1-\frac{\binom{n-2}{q}}{\binom{n}{q}}\right)^r+\left(1-\frac{\binom{n-1}{q}}{\binom{n}{q}}\right)^r \tag{5}$$

Alle parenteser er mindre end 1, så middelværdien nærmer sig q. Man skal op på ret store r-værdier, før første parentes begynder at blive mindre end 1.

1.14 Approksimation med potensfunktion

For at finde en enklere beregning af middelværdien forsøges med simple approksimationer.

Svarende til (2) må findes en differentiabel udgave af fordelingen, som skal opfylde $G'(x) = g(x) = (q/x)G(x),\ q \le x \le n,\ G(n) = 1$, dvs. $G(x) = (x/n)^q$. Modellen svarer til, at 15 tal vælges et ad gangen med tilbagelægning ud af 90. For at få 15 forskellige ud af 90 skal i middel vælges (coupon collector problem) $n \ln \frac{n+0.5}{n-q+0.5}$, med de aktuelle værdier $90 \ln \frac{90+0.5}{75+0.5} \approx 16.3$.



1.15 Middelværdi og varians for approksimation med potensfunktion

$$G(x) = 1 - \left(1 - \left(\frac{x}{n}\right)^q\right)^r \text{ og } g(x) = \frac{rq}{n}\left(1 - \left(\frac{x}{n}\right)^q\right)^{r-1}\left(\frac{x}{n}\right)^{q-1},$$

$$\begin{aligned}
E(Y) &= \int_q^n x\left(\frac{rq}{n}\left(1 - \left(\frac{x}{n}\right)^q\right)^{r-1}\left(\frac{x}{n}\right)^{q-1}\right) dx \\
&\approx \cdots \\
&= r\ n\ B(r, 1+1/q), \text{ hvor B er } \textit{Betafunktionen} \\
&= r\ n\ \frac{\Gamma(r)\Gamma(1+1/q)}{\Gamma(r+1+1/q)} \qquad , (\Gamma(1) = 1, \Gamma(n+1) = n!, \Gamma(x+1) = x\Gamma(x), x \in \mathbb{R}_+) \\
&= n \frac{q}{q+1}\frac{2q}{2q+1}\cdots\frac{(r-1)q}{(r-1)q+1}\frac{rq}{rq+1} \\
&= n \frac{1}{1+\frac{1}{q}}\frac{1}{1+\frac{1}{2q}}\cdots\frac{1}{1+\frac{1}{(r-1)q}}\frac{1}{1+\frac{1}{rq}}
\end{aligned}$$

Middelværdien er udtrykt ved ret enkle regneudtryk med parametrene q (antal tal på plade), n (antal tal i alt) og r (antal plader). Der er approksimationer undervejs.
Justeres eksponenten q til $q + 2$ får man

$$E(Y) = 90 \frac{17}{18}\frac{34}{35} \times \cdots \times \frac{17(r-1)}{17(r-1)+1}\frac{17r}{17r+1}$$

Fra generaliseret liniebingo afsn. 1.2 s. 81 havde man for seks plader med alle tal (et eksakt resultat), og med de sidste udtryk

$$\begin{aligned}
\text{fra liniebingo :} &\quad 90\frac{15}{16}\frac{30}{31}\frac{45}{46}\frac{60}{61}\frac{75}{76} = 77.5 \\
\text{ikke justeret eksponent :} &\quad 90\frac{15}{16}\frac{30}{31}\frac{45}{46}\frac{60}{61}\frac{75}{76}\frac{90}{91} = 76.7 \\
\text{justeret eksponent :} &\quad 90\frac{17}{18}\frac{34}{35}\frac{51}{52}\frac{68}{69}\frac{85}{86}\frac{102}{103} = 78.1
\end{aligned}$$

Seks plader med alle tal giver i stor model 77.5 og kan selvfølgelig ikke laves i lille model. Med seks vilkårlige plader giver både lille og stor model 78.3. Med approksimationer har man næsten samme approksimation som fra computermodellen. [2]

$$E(Y_r) \approx 90\ \Gamma(1 + 1/17)\ r^{-\frac{1}{17}} \approx 90 \times 0.9693\ r^{-0.059} = 87.2\ r^{-0.059} \approx 87\ r^{-0.06}$$

[2]

$$E(Y_r) \approx n\ r^{-\frac{1}{q}}\ \Gamma(1 + 1/q), 1 < q < n,\ (n+1)/n < 1 + 1/q < 2,$$

Variansen kan beregnes til

$$Var(Y_r) = 90^2\ r^{-\frac{2}{17}} \left(\Gamma\left(1+2/17\right) - \left(\Gamma\left(1+1/17\right)\right)^2\right) \approx 40\ r^{-0.1176}, 5 < r$$

Sammenlignes middelværdi og varians med Weibull fordelingens middelværdi og varians

$$E(W) = n\ r^{-\frac{1}{q}}\ \Gamma\left(1+1/q\right) = a + b\ \Gamma\left(1+1/c\right)$$

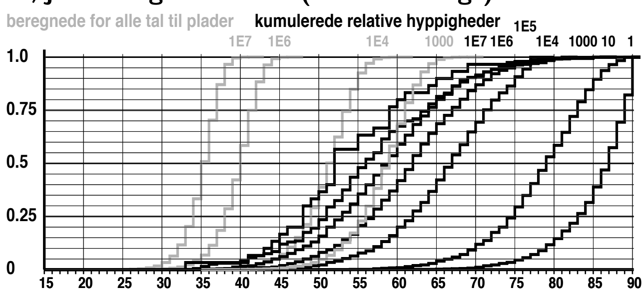
$$Var(W) = n^2 r^{-\frac{2}{q}} \left(\Gamma\left(1+2/q\right) - \left(\Gamma\left(1+1/q\right)\right)^2\right) = b^2 \left(\Gamma\left(1+2/c\right) - \Gamma^{\ 2}\left(1+1/c\right)\right)$$

får man $a = 0$, $b = n\ r^{-\frac{1}{q}}$ og $c = q$. Med de aktuelle værdier $b = 90\ r^{-\frac{1}{15}}$ og $c = 15$, eller justeret $b = 90\ r^{-\frac{1}{17}}$ og $c = 17$ ($b \sim \beta$ er scale og $c \sim \alpha$ er shape). Sandsynlighedsmodellen $Y_r \sim Weibull(\alpha, \beta)$ passer rimeligt op til $r = 100000$. Det er en af de klassiske *livslængdefordelinger* med overlevelses funktionen $S(y) = \exp(-(y/\beta)^\alpha)$ og fordeling $F(y) = 1 - S(y) = 1 - \exp(-(y/\beta)^\alpha)$.

For større pladeantal er $Y_r \sim extremvalue(a, c)$ med fordeling $1 - \exp(-\exp((x-a)/c))$ og overlevelses funktion $S(y) = \exp(-\exp((y-a)/c))$. Hvis $X \sim extremvalue(0, 1)$ er $E(X) = -\gamma = -0.5772 \approx -0.6$ og $Var(X) = \pi^2/6 \approx 5/3$. Er $Y = a + bX$ er $Y \sim extremvalue(a, b)$ og $E(Y) = a + b\ E(X) = a - 0.6\ b$ og $Var(Y) = (5/3)b^2$ Medianen er $a + b\ln(\ln(2))$, se f.eks. [D], [S].

1.16 RBINGO plader med søjletal valgt ud af tre (færrest muligt)

Når hver søjles tal vælges ud af tre tal, er antal plader ca. 7.6 10^6.
Med 1000 plader er medianen 67, medens den normalt er 59. Minimumfordelig er nær venstre graf. Stabil max 80 til 85.



2 KBINGO plader

Hver plade har forskellige lodrette L4 og L5, vandrette V4 og V5, og diagonaler D4

- $L5_1, L5_2, L5_3, L5_4$ ud af $\binom{15}{5} = 3003$
- $L4$ ud af $\binom{15}{4} = 1365$
- $V4, D4_1, D4_2$ ud af $15^4 = 50626$
- $V5_1, V5_2, V5_3, V5_4$ ud af $15^5 = 759375$

8	25	43	50	75
4	18	40	58	73
6	30	ANDERS BINGO BANKO	60	65
5	28	45	51	72
9	24	39	56	62

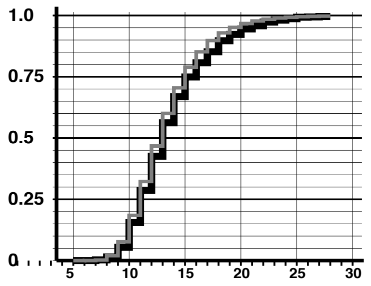
ventetiden er $\min\{L5_1, \ldots, V5_4\}$. En plade er som spil med 12 plader $L5_1, \ldots, V5_4$, hvor pladetallene er valgt ud af 24 tal, bruges fra to til tre gange og er sammenvævet i et mønster, dvs. der er ret stor afhængighed.

$$\text{fire-bingo:} \begin{cases} \text{række og diagonaler V4,D4: } 15^4 & = & 50625 \\ \text{søjle L4 : } \binom{15}{4} & = & 1365 \\ \text{i alt} & = & 51990 \approx 5.2\ 10^4 \\ \text{højst } \binom{75}{4} & = & 1\ 215\ 450 \approx 1.2\ 10^6, 4.3\% \end{cases}$$

Jo flere plader i et spil, jo flere med bingo på samme antal træk. I tabellen er de optalte relative hyppigheder.

ant plader	antal plader med bingo										
r	1	2	3	4	5	6	7	8	9	10	$\bar{x}$
100	0.84	0.13	0.02	0							1.2
1000	0.67	0.20	0.08	0.03	0.01	0.01					1.6
10000	0.18	0.17	0.14	0.10	0.08	0.07	0.05	0.05	0.03	0.02	5.3
100000	0.0	0.0	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.03	47.7

Den øverste tynde trappekurve er fordeling for antal gange 24 tal skal vælges for at alle er med. Den nederste kraftige trappekurve er antal plader, der skal vælges, for at alle tal er med. **Der er nogen afhængighed mellem tal til K-BINGO plader.**



2.1 Beregning

Er $P(B_m)$ sandsynligheden for bingo med færre eller m taltræk, og klassedeles efter antal hit, som her kan være fra fire til 24, har man med k og m valgt passende

$$P(B_m) = P(B_4|4 \text{ hit})P(4 \text{ hit}) + \cdots + P(B_m|\text{k hit})P(\text{k hit}),$$

Der er fire mulige *fire-bingoer* med fire hit $T_4 = 4$, dvs.

$$P(B_m|4 \text{ hit}) = \frac{4}{\binom{24}{4}},\ P(4 \text{ hit}) = \frac{\binom{24}{4}\binom{51}{m-4}}{\binom{75}{m}}, 4 \leq m \leq 55$$

Med fem eller flere hit kan der være fire *fire-bingoer* og otte *fem-bingoer*. Med fem hit er der $T_5 = 4\binom{20}{1} + 8\binom{19}{0} = 88$ muligheder for bingo

$$P(B_m|5 \text{ hit}) = \frac{88}{\binom{24}{5}} = \frac{T_5}{\binom{24}{5}},\ P(5 \text{ hit}) = \frac{\binom{24}{5}\binom{51}{m-5}}{\binom{75}{m}}, 5 \leq m \leq 56$$

Tabellen er lavet af et program, der tæller antal bingo i forhold til antal hit. Med f.eks. fire hit køres alle mønstre igennem, men det ses umiddelbart, at der kun er fire mulige.
Med 20 hit er der mindst en bingo.

antal hit	T_n ant bingo	antal mulige	sand
$0,\dots,3$	0	...	0
4	4	$K(24,4) = 10626$	0.0004
5	88	$K(24,5) = 42504$	0.0021
6	912	$K(24,6) = 134596$	0.0068
7	5928	$K(24,7) = 346104$	0.0171
8	27102	$K(24,8) = 735471$	0.0369
9	92520	$K(24,9) = 1307504$	0.0715
10	244092	$K(24,10) = 1961256$	0.1245
11	507696	$K(24,11) = 2496144$	0.2034
12	841100	$K(24,12) = 2704156$	0.3110
13	1113360	$K(24,13) = 2496144$	0.4460
14	1174620	$K(24,14) = 1961256$	0.5989
15	981424	$K(24,15) = 1307504$	0.7506
16	644445	$K(24,16) = 735471$	0.8762
17	331056	$K(24,17) = 346104$	0.9595
18	133428	$K(24,18) = 134596$	0.9913
19	42480	$K(24,19) = 42504$	0.9994
20	10626	$K(24,20) = 10626$	1
21	2024	$K(24,21) = 2024$	1
22	276	$K(24,22) = 276$	1
23	24	$K(24,23) = 24$	1
24	1	$K(24,24) = 1$	1

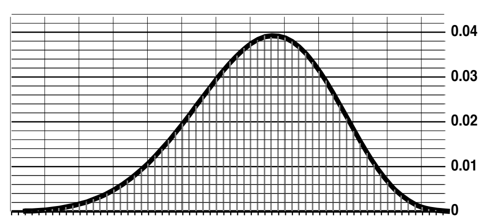
For m udtrukne numre, hvor $4 \leq m \leq 24$ kræves for mindst en bingo, at der er fra 4 til m hit.

$$\begin{aligned} P(B_m) &= \frac{\binom{24}{4}\binom{51}{m-4}}{\binom{75}{m}} \frac{T_4}{\binom{24}{4}} + \cdots + \frac{\binom{24}{m}\binom{51}{0}}{\binom{75}{m}} \frac{T_m}{\binom{24}{m}} \\ &= \left(\binom{51}{m-4}T_4 + \cdots + \binom{51}{0}T_m\right) \Big/ \binom{75}{m} \end{aligned}$$

og for $25 \leq m \leq 71$ kræver mindst en bingo fra 4 til 24 hit, $\binom{51}{k} = 0$ for $k > 51$.

$$\begin{aligned} P(B_m) &= \frac{\binom{24}{4}\binom{51}{m-4}}{\binom{75}{m}} \frac{T_4}{\binom{24}{4}} + \cdots + \frac{\binom{24}{24}\binom{51}{m-24}}{\binom{75}{m}} \frac{T_{24}}{\binom{24}{24}} \\ &= \left(\binom{51}{m-4}T_4 + \cdots + \binom{51}{m-24}T_{24}\right) \Big/ \binom{75}{m} \end{aligned}$$

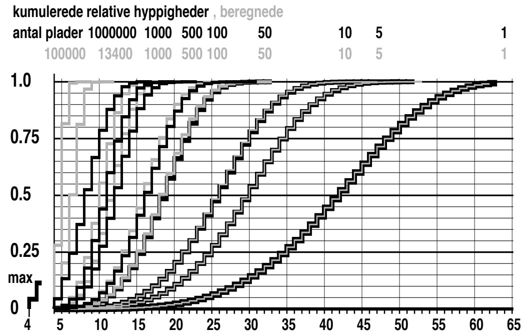
Alle bingospil må have en mindste fordeling, når alle plader er med. Den optæder næppe i praksis, den må begynde i antal tal på plade, og have værdien *(antal muligheder for bingo)/(antal mulige talsæt)*. Hvis alle talsæt er med, dvs. der er intet plademønster, er det en lodret linie. I R-BINGO er så mange plader, at computermodellen ikke kan køres med alle plader, men for KBINGO kan de ca. 800000 bingomuligheder køres igennem. Her beregnes fordelingen for flere plader, som om de var uafhængige, se evt. afsn. 1.13 s.88. Afhængigheden viser sig som en *spredning* mellem beregnet og målt fordeling, de udgår fra samme punkt på 1.aksen. Med 10^6 plader er medianen for beregnet fordeling 5 og for computermodellen 8.



Figur 2: I fig. er kraftig graf beregnet sandsynlighed, stolper er relativ hyppighed fra computermodel.

Middelværdier og fordelinger for op til 1000 plader kan beregnes ud fra sandsynlighederne for en plade som i afsn. 1.13 s.88, hvor der forudsattes uafhængighed. Det gøres lettest i samme program, som beregner T_i. For større pladeantal er afvigelserne ret store.

David B. Agard & Michael W. Shackleford: *A New Look at Probabilities in Bingo* [AS] beregner sandsynlighederne lidt anderledes og hyppighederne tælles op i computermodel for spil med 1, 10 og 50 plader og ligner resultaterne her.



Figur 3: På fig. er beregnede fordelinger lys grå (forudsat uafhængige), og målte fordelinger mørkere grå (kumulerede relative hyppigheder).

2.2 Netbrevkassespørgsmål KBINGO, netbingospil

Eksempel 2.1 Fra `wizardofodds.com/askthewizard/numbered/askthewizard109.html`:

> At bingogala.com they offer a $500 prize for a coverall within 54 calls. You told me earlier that the probability of that at least 1 card in 600 will get a coverall in 54 calls is 3.21%. So in 380 days (to date) at 8 sessions per day they should have 97.58 $500 winners, right? However I counted only 76 winners on their home page. When I brought up my question about this in chat my husband and I were both banned from the site which really sent my antenna up? Sorry to be a pest but if they are running an unscrupulous site I want to know how to figure it out so that I can shout it far and wide with facts. Thank you for any help you can give me in this matter.

Se evt. svar samme sted fra Sept.24 2002.

Relativ hyppighed for coverall har været $76/(380 \times 8) = 0.025$, og der er udbetalt \$12.5 pr. spil. Middelantal træk for coverall for en plade er $24(75+1)/(24+1) = 72.96 \approx 73$ (3) s.88. Af (4) s.88 har man

$$P(Y \leq 54) = 1 - \left(1 - \frac{\binom{54}{24}}{\binom{75}{24}}\right)^{600} \approx 0.03212$$

for hvert spil. Gentages 8 gange pr. dag i 380 dage, og vælges den oplagte binomialmodel, er det en binomialfordeling med sandsynlighed p=0.0321 og antal n=3040.

$$P(X \leq 76) = \sum_{k=0}^{76} \binom{3040}{k} 0.0321^k (1-0.0321)^{3040-k} = 0.0127$$

Hændelsen vil observeres færre end eller 13 gange ud af 1000 tilsvarende spil (1000 gange 3040), så det er rimeligt, at spørgeren undrer sig. Computermodellen giver 0.0319 for 50000 spil, som kun ændrer lidt ved sandsynligheden.

Som det nævnes i svaret, er det ikke sikkert, om der har været 600 plader i hvert spil. I nogle spil kan have været færre plader f.eks. 560, hvilket giver

$$P(Y \leq 54) = 1 - \left(1 - \frac{\binom{54}{24}}{\binom{75}{24}}\right)^{560} \approx 0.0300$$

Pladerne kan have været valgt specielt, hvilket giver lidt længere venstre hale og mindre hældning, dvs. større middelværdi og større spredning. Sandsynligheden bliver mindre for et bestemt udfald.

Er $P(Y \leq 54) = 0.03$, har man

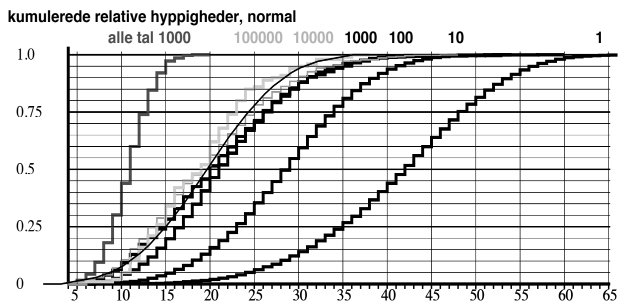
$$P(X \leq 76) = \sum_{k=0}^{76} \binom{3040}{k} 0.03^k (1-0.03)^{3040-k} = 0.056$$

som er 5.6% af tilfældene. Det er derfor vanskeligt, at udtale sig om rimeligheden af antal gange coverall er forekommet uden at vide, hvor mange plader der mindst har været i hvert spil, og hvordan pladerne er lavet.

Ved netspil laves pladerne formentlig af et program eller vælges fra et lager. Spørgsmålet er, om tallene på pladerne er jævnt fordelt. For coverall vælges 24 tal ud af 75, så afhængigheden er lille og uden betyning. Mange forhold påvirker antal træk. Det er nok rimeligt at antage, at spilentreprenøren er klar over, at specielle pladevalg giver flere træk til coverall. Til gengæld øges antal plader udtrukket på samme antal træk, men pladerne skal være meget specielle, før dette område nås.

2.3 Plader med færre tal

Med en computermodel kan man let undersøge, hvordan specielle pladevalg påvirker antal træk. Vælges tallene til hver søjle ud af seks tal er der ca. 9000 bingomuligheder.



Fordelingerne bliver mindre stejle og middelværdierne større, de ligner normalfordelinger. Højre gren ser ud til at ligge fast i intervallet [35,40], medens venstre gren bevæger sig mod mindsteværdien 4.

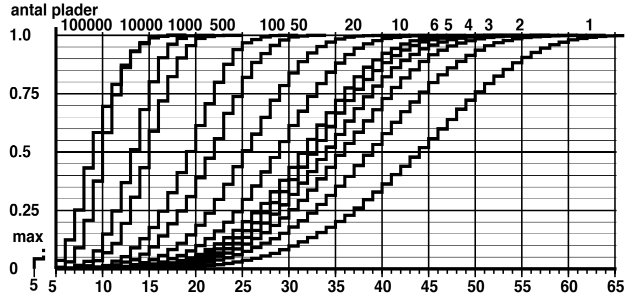
2.4 SKBINGO

Kvadratiske plader uden slettet center kaldes ofte svensk bingo. For det meste er 12 muligheder for bingo, $L_1, \ldots, L_5, V_1, \ldots, V_5, D_1, D_2$, hvor pladetallene er valgt ud af 25 tal, bruges fra to til tre gange og er sammenvævet i et mønster.

B	I	N	G	O
14	26	34	59	70
9	30	36	60	61
2	29	39	56	66
8	22	43	57	65
6	27	31	50	72

- L_1, L_2, L_3, L_4, L_5 ud af $\binom{15}{5} = 3003$
- $V_1, V_2, V_3, V_4, V_5, D_1, D_2$ ud af $15^5 = 759375$

Antal muligheder for bingo er 774390, og det samlede antal talsæt med fem tal



er $\binom{75}{5} \approx 1.7\ 10^7$, dvs. ca. 4.5% kan optræde som bingomuligheder. Observerede hyppigheder fra computermodel er tegnet som trappekurver. Spillet ligner KBINGO, men fordelingerne er rykket lidt til højre, hvilket passer med, at muligheden for bingo på fire træk ikke er med. Antal plader udtrukket på samme antal træk, er en del mindre.

ant plader	antal plader med bingo										
r	1	2	3	4	5	6	7	8	9	10	$\bar{x}$
100	0.84	0.13	0.02	0.01							1.2
1000	0.74	0.18	0.05	0.02	0.01	0.00					1.4
10000	0.41	0.23	0.12	0.08	0.06	0.04	0.02	0.01	0.01	0.00	2.7
100000	0.07	0.05	0.06	0.07	0.03	0.03	0.04	0.03	0.03	0.01	21.2

3 Konklusion

Computermodeller og beregninger passer pænt sammen, når der ikke er for stor afhængighed mellem plader og pladetal. Resultater fra computermodeller for spil med plader med færre tal regnes derfor for troværdige. Det er praktisk taget umuligt at afgøre, om pladerne i et spil er lavet specielt. Der kan laves millioner af plader, hvor et eller flere tal mangler.

Litteratur

[AS] David B. Agard & Michael W. Shackleford 2002. *A New Look at Probabilities in Bingo* in College Mathematics Journal, Volume 33, Number 4, 301-305.

[A] Nils Andersen, 2003. *Hvor mange bankoplader er der?* Normat nummer 3, 2003, side 102, `http://www.diku.dk/ nils/Notes/bankoplader.pdf`

[BS] D. M. Bloom, Kenneth Suman, GCHQ Problems Group 1999. *Generalized Line Bingo: 10565[1997,68]*, American Mathematical Monthly, Vol. 106, No. 1 (Jan., 1999), p. 72.

[D] H.A.David 1970. *Order Statistics*, John Wiley & Sons, Inc, 1970.

[S] Peter J. Smith 2002. *Analysis of failure and survival data.* Chapman & Hall/CRC.