Løsningsforslag til Stokastik 1.-10. klasse

Bemærk, at vi benytter betegnelsen øvelser som en meget bred betegnelse. Derfor er der også nogle af vores øvelser, der nærmer sig kategorien 'undersøgelser', dem giver vi som oftest ikke løsningsforslag til, ligesom svar til kategorien 'overvej-diskuter' ikke giver megen mening.

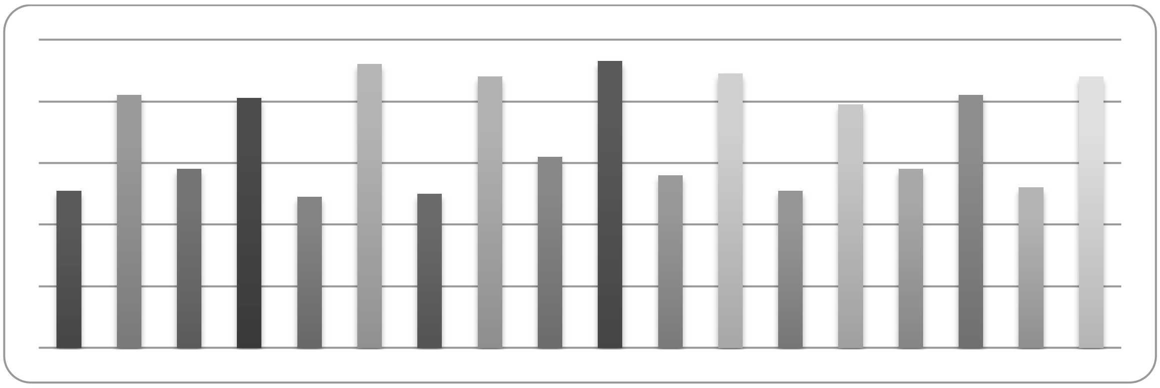
Kapitel 1 Åben udforskende dataanalyse

Øvelse 3

Det kan være praktisk at have data tilgængelige i en elektronisk udgave for nemmere at kunne eksperimentere med dem:

Dag 1	51	82	58	81	49	92	50	88	62	93	56	89	51	79	58	82	52	88
Dag 2	86	78	71	77	76	94	75	50	83	82	72	77	75	65	79	72	78	77
Dag 3	65	89	49	88	51	78	85	65	75	77	69	92	68	87	61	81	55	93

I mangel af en bedre idé prøver vi at lave et søjlediagram over tallene den første dag:



Her falder det i øjnene, at der er et ganske regulært skift mellem korte og lange søjler, som om en lang ventetid efterfølges af en kortere. Måske skulle vi så prøve at opdele materialet i de kortere og de lange hver for sig – altså vælge hver anden. Vi fortsætter med at gøre det for den første dag, så vi kan bruge de to næste dage til at afprøve den indsigt, vi tror at have vundet ud fra en analyse af førstedagens materiale.

Hver anden / korte:	51	58	49	50	62	56	51	58	52
Hver anden /lange:	82	81	92	88	93	89	79	82	88

Vi kan karakterisere hver talserie ved hjælp af fx minimum, kvartilsæt og maksimum for at se, om det afslører noget interessant (i et regneark kan det klares i en håndevending med funktionen kvartil):

deskriptor	Minimum	nedre kvartil	median	øvre kvartil	Maksimum
hjælpetal	0	1	2	3	4
korte	49	51	52	58	62
lange	79	82	88	89	93

Vi observerer, at de lange og de korte ventetider er klart adskilte. Og vi kunne nok forvente, at de korte også i fremtiden ville ligge i 50'erne, mens de lange ville ligge i 80'erne, måske med enkelte afvigelser.

Fortsæt selv beskrivelsen og undersøgelsen.

Øvelse 4

Igen kan det være behageligt at have data repræsenteret i elektronisk form:

Østerby	x		Vesterby	y
4	1		2	4
2	2		3	2
1	3		1	1
3	3		1	0
2	4		1	1
5	1		3	2
3	4		4	5
1	2		5	4
2	2		3	1
3	4		2	2

Vi ved ikke, om der er spillet mod de øvrige hold i samme rækkefølge, så det vil ikke give mening at sammenligne kamp for kamp, selvom data er stillet sådan op i rækker.

Det drejer sig først og fremmest om at konstruere et kvalitetsmål ud fra sådanne kampdata. Antal vundne kampe vil naturligt indgå, ligesom antallet af tabte. Måske kunne (vundne – tabte) være et godt groft mål, men så er der også spørgsmålet om, hvor meget de har vundet med. Måske siger den samlede måldifference noget vigtigt. Sportsfolk blandt læserne vil selv kunne finjustere dette kvalitetsmål og måske lave et regneark, der kan håndtere sådanne sammenligninger mere generelt. Hertil kommer en historisk og kulturelt bestemt faktor, der siger noget om stemningen i kampe mellem Østerby og Vesterby. Det kan være svært at håndtere matematisk, men måske kan statistikken fra de tidligere års indbyrdes kampe sige noget.

Øvelse 5

1) Det kan godt være at svømmeveste er en god sikkerhed mod drukning, men argumentet holder ikke, hvilket afsløres, hvis man skriver noget andet i stedet for svømmeveste, fx "havde hat på":
"I det forløbne år druknede 35 personer ved bådulykker. Kun fem af dem havde hat på; ingen af de andre bar hat. Bærer de selv hat, når de er ude at sejle?"

Men det mest kritiske spørgsmål er: "Hvor mange folk til søs bærer egentlig svømmevest? Hvis kun 3% bærer svømmevest og en 'stikprøve' på 35 falder i vandet og drukner, så ville vi selv ved virkningsløse svømmeveste kun forvente, at en enkelt af disse havde svømmevest på. Når der nu var 5 af de 35, der havde svømmevest på, så synes det at tyde på, at svømmevestene får folk til lettere at falde over bord og drukne. De er altså direkte farlige, hvis kun 3% bærer svømmevest.

Overvej, hvor stor en procentdel af befolkningen der normal skal bære svømmeveste til søs, før argumentet i annoncen kan siges at holde?

2) *Den snigende død i mineralvand*
På grafen kan man se, at hvis alle folk i en mellemstor by på 25.000 hver dag drikker 20 liter mineralvand fra Perrier, så vil antallet af kræfttilfælde inden for de næste 40 år stige med 1 i byen. Hertil kommer forureningen fra de lastbiler der hver dag skal bringe de nødvendige 1½ millioner flasker ind til byen. Regeringen overvejer at nedsætte det størst tilladelige indtag til 5 liter pr. dag, hvilket ca. vil halvere risikoen og i al fald lette på trafikken.

I de resterende øvelser i dette kapitel, må læseren klare sig uden forfatternes assistance.

Kapitel 2 Beskrivende statistik

Øvelse 4

4, 4, 4, 7, 7, 10, 12, 12, 12 er en mulighed med to typetal. Men 4, 4, 4, 7, 7,… kan godt suppleres med fem karakterer, så kravene opfyldes og kun 4 er typetal.

Øvelse 5

I udgangpositionen var vi et observationssæt $\{x_1, x_2, x_3, ...x_n\}$ med median < typetal < middeltal.

1) Hvis vi nu ser på observationssættet, hvor alle observationer er ganget med 10, så vil alle de tre centrale deskriptorer selvfølgelig også blive ganget med 10, så den indbyrdes beliggenhed bliver uforandret: median < typetal < middeltal.

Hvis fx 9 er typetal, altså den oftest forekommende observation, bliver 90 selvfølgelig den oftest forekommende observation, hvis der sættes et 0 bag hver observation, da alle 9-taller bliver til 90. Hvis 8 er den midterste observation (eller median), så bliver 80 selvfølgelig den midterste i det 10 gange større observationssæt, fordi 80 vil få akkurat samme placering i dette observationssæt, som 8 havde i det oprindelige.

Hvis 11 er middeltallet (altså gennemsnittet), så er tallene i det 10 gange større observationssæt 10 gange større hver for sig og derfor også i gennemsnittet eller i middeltal.

2) Hvis vi ganger alle observationerne med -1 så vendes størrelsesrækkefølgen om, og de nye deskriptorer: middeltal, median og typetal bliver lig de gamle med et minus foran. Da multi-plikation med -1 vender ulighedstegn fås - middeltal < - typetal < - median eller nyt middeltal < nyt typetal < ny median.

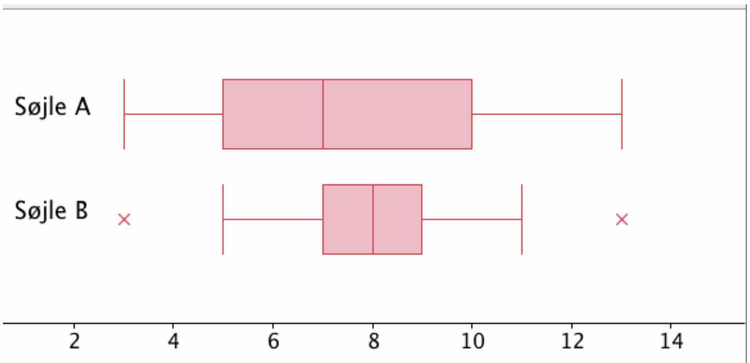
Rækkefølgen af de tre centrale deskriptorer vendes altså om, når observationssættet ganges med -1. Denne øvelse kan udnyttes i den følgende undersøgelse.

Øvelse 6

2,7,7,7,7,7,7,7,7,7,12 er en mulighed med kvartilafstand 0.

Øvelse 8

Her er en relevant repræsentation af data (box-plots) til at fortolke ud fra, men to hyppighedsfordelinger vil afsløre samme principielle forskel på de to hold både med hensyn til midten af holdet og holdets spredning.



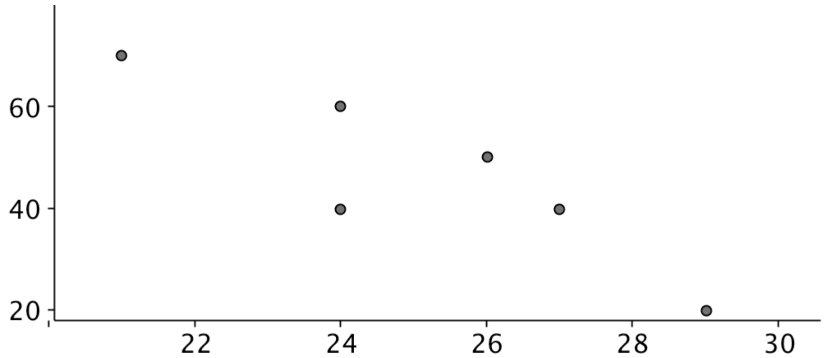
Statistik

	n	Mi…	σ	s	Min	Q1	Me…	Q3	Maks
Søjle A	50	7.26	2.…	2.…	3	5	7	10	13
Søjle B	50	8.02	1.…	1.…	3	7	8	9	13

Kapitel 3 På sporet af sammenhænge

Øvelse 2

Her er et scatterplot af data, som kan benyttes til at fortolke ud fra. Umiddelbart fortæller plottet en anden historie end den forventede, idet det synes, som om der er mere ro i klassen, jo større den er!



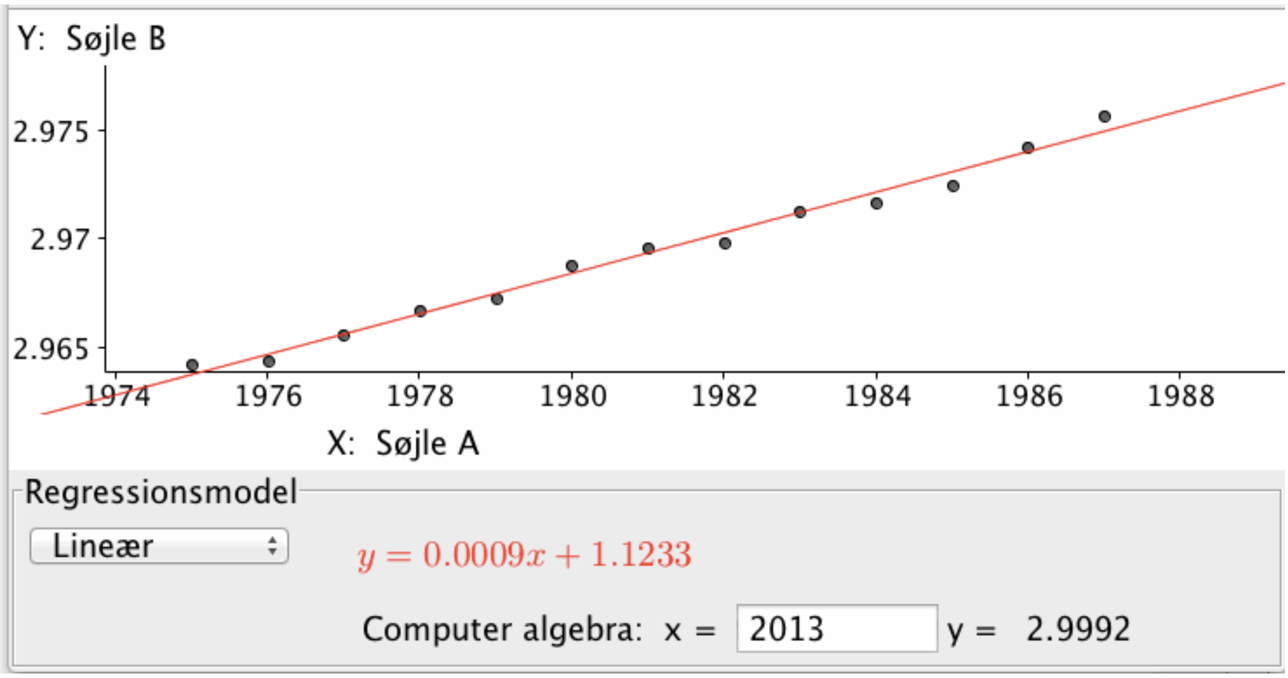
Øvelse 3

For at finde passende funktion for hver gruppe, kan det være praktisk at have måledata i elektronisk form:

Gruppe 1		Gruppe 2		Gruppe 3			
Højde	Vægt		Højde	Vægt		Højde	Vægt
170	72		190	87		170	98
192	92		180	78		192	114
167	69		167	67		167	95
160	64		186	83		160	90
185	85		166	66		185	51
192	92		185	82		192	55
168	70		186	83		168	73
193	93		181	79		193	78
170	72		181	79		170	81
163	66		195	91		163	80
167	69		160	61		167	78
164	67		164	64		164	86
161	64		195	91		161	57
167	69		160	61		167	56
182	82		174	73		182	109

Øvelse 9

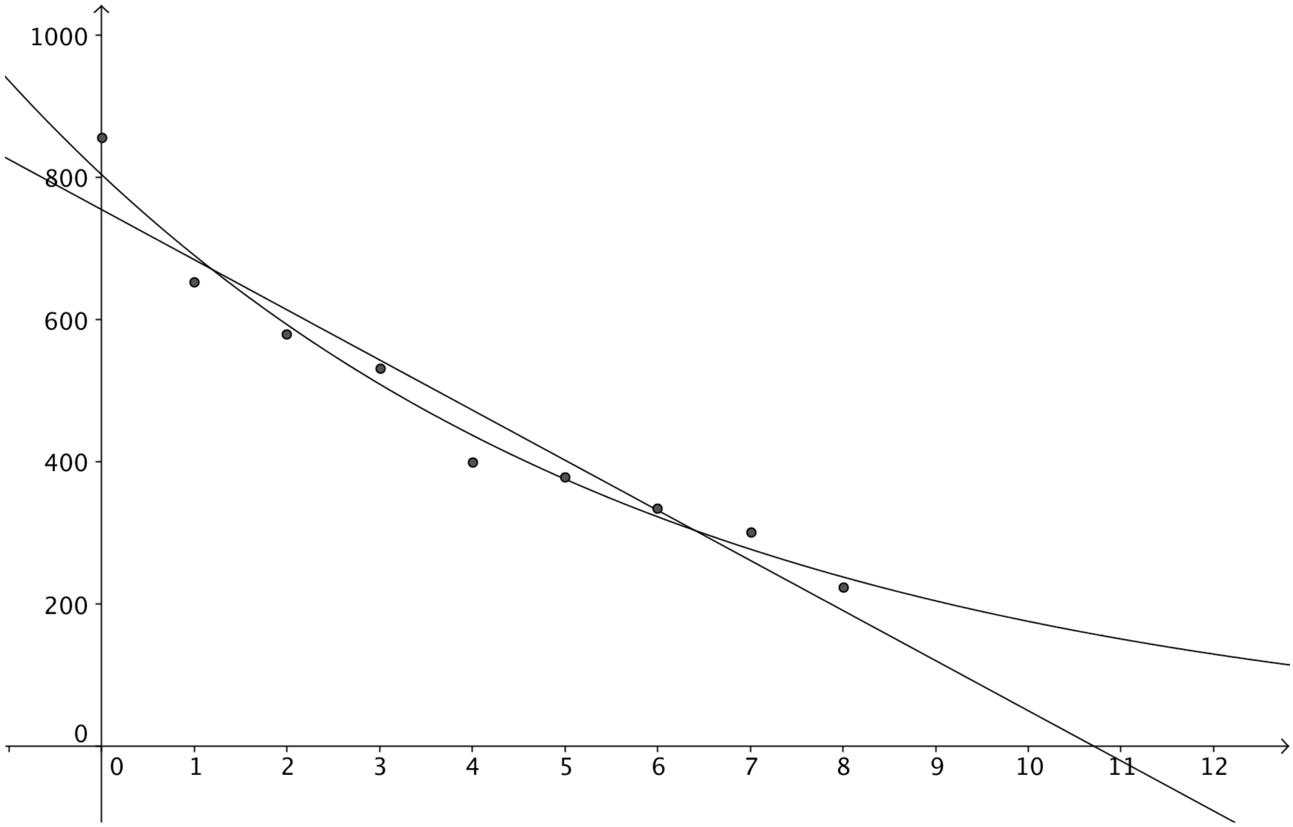
Den bedste rette linje bestemmes i GeoGebra



Modellen giver en forventet skævhed i 2013 på 2,9992 meter

Øvelse 10

Nedenfor ses to forskellige modeller for udviklingen i antallet af 'Camillaer'. En lineær og en eksponentiel model. x er antal år efter 1996.



Modellerne er hhv $f(x) = 803.6369 \cdot 0.8589^x$ og $g(x) = -70.4667x + 754.7556$.
Passer de også i dag? Tjek fx http://www.dst.dk/da/Statistik/emner/navne/navne-til-nyfoedte.aspx

Øvelse 14

1) Da odds kan antage ethvert positiv talværdi, så kan odds ratio – altså forholdet mellem to odds også antage enhver positiv værdi – og faktisk også 0, da odds 0 også er en mulighed.

Hvis odds ratio havde været lig 1 i eksemplet med deltidsarbejde for mænd/kvinder, så ville det betyde, at andelen af folk på deltidsarbejde var den samme for de to køn. Hvis den var 0,5, så ville kvinders tendens til at være på deltidsarbejde være den halve af mændenes. Hvis derimod odds ratio var meget stor, ville kvinder i langt større grad være på deltidsarbejde end mænd.

Selvfølgelig bliver enhver kort beskrivelse af betydningen lidt forsimplet. Hvor vi ovenfor skrev: "ville kvinder i langt større grad være på deltidsarbejde end mænd", har vi fx ikke påstået, at der er flere kvinder på deltidsarbejde, end der er mænd på deltidsarbejde. Hvis der er få kvinder på arbejdsmarkedet i det hele taget, kunne det fx være, at de næsten alle sammen var på deltidsarbejde, men alligevel udgjorde en langt mindre gruppe end mændene på deltidsarbejde. Tilsvarende tolkning for meget lille odds ratio.

2) Odds ratio for at være på fuld tid i en sammenligning mellem kønnene:

$$\frac{\frac{21}{9}}{\frac{22}{1}} = \frac{21 \cdot 1}{9 \cdot 22} = \text{det omvendte af før } = \text{cirka } 0{,}11$$

Også hvis man bytter om på de to køn, byttes der om på tæller og nævner, og resultatet bliver ca. 9,43.

Det kan synes lidt kaotisk, men der er kun to mulige resultater her 9,43 og 0,11. Og hvis man ved, at tendensen for deltidsarbejde er størst for kvinder, så behøver man ikke at blive forvirret af, at der er fire forskellige odds ratios, som man kan udregne.

Øvelse 17

store

	succes	ikke succes	odds		
Test A	192	71	2,7	Odds Ratio	1,2
Test B	55	25	2,2		

lille

	succes	ikke succes	odds		
A	81	6	13,5	Odds Ratio	2,1
B	234	36	6,5		

Begge

	succes	ikke succes	odds		
A	273	77	3,55	Odds Ratio	0,75
B	289	61	4,74		

Odds Ratio for behandlingerne på begge sygehuse tilsammen er “modsat” OR for hver af de to sygehuse for sig.

Hvordan kan det gå til?

Kapitel 4 Elementer af stokastikkens didaktik

Øvelse 2

$\frac{3}{8}$ kan forklares ved tænkningen: Der er tre mulige rækkefølgerdpp, pdp, ppd, hver af disse har sandsynlighed $\frac{1}{8}$.

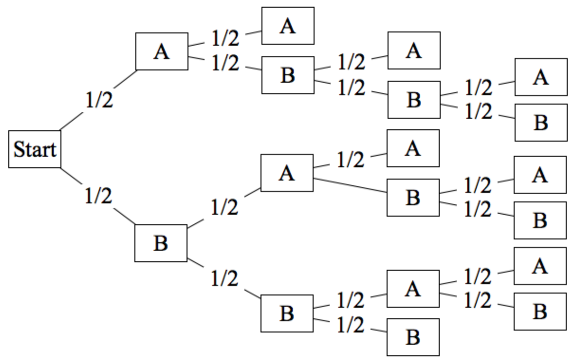
$\frac{1}{4}$ fremkommer formentlig ved tænkningen: Der er fire muligheder: 0 drenge, 1 dreng, 2 drenge og 3 drenge. Vi leder efter den ene af dem; derfor svaret $\frac{1}{4}$

Øvelse 3

Fermat/Pascal problemet er lidt vanskeligere at få hold på. En løsning – den som peger fremad mod mere systematisk sandsynlighedsregning – er at se på, hvad der kunne være sket, hvis ikke spillerne var blevet afbrudt og dele puljen i forhold til de to spilleres sandsynligheder for at vinde puljen. Stillingen er 8 – 7, og man skulle have spillet til én spiller havde vundet 10 gange. Derfor ville der aldrig skulle spilles mere end 4 spil mere for at finde vinderen. Vi viser i skemaet en systematisk analyse af, hvad der kunne være sket i disse højst 4 spil.

A	A			A har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$
A	B	A		A har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}$
A	B	B	A	A har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{16}$
A	B	B	B	B har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{16}$
B	A	A		A har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}$
B	A	B	A	A har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{16}$
B	A	B	B	B har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{16}$
B	B	A	A	A har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{16}$
B	B	A	B	B har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{16}$
B	B	B		B har vundet - sandsynlighed $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}$

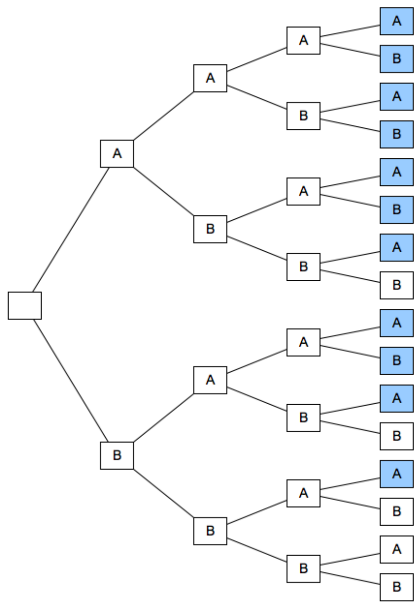
For at være sikre på, at vi har dækket alle muligheder, summerer vi lige alle sandsynlighederne for at sikre os, at de tilsammen giver 1. Det gør de.
I et chancetræ kan det se således ud:



Hvert spil har 50 % chance for at gå til A eller til B, og vi anvender multiplikationsprincippet for at beregne sandsynligheden for de forskellige forløb.

A har sandsynligheden $\frac{11}{16}$ for at vinde puljen, hvis spillerne havde fortsat. Det er derfor retfærdigt at dele puljen i forholdet 11:5, hvilket netop giver A $\frac{11}{16}$ af puljen.

I et tælletræ kan det se således ud



De slutsituationer, der er markeret med blåt, er de forløb hvor A har vundet hele puljen. Det sker i 11 ud af 16 situationer. Det er derfor retfærdigt at dele puljen i forholdet 11:5.
Som der fremgår af teksten forekommer der mange andre svarforslag til pointproblemet. Prøv, om du kan forklare motivationen bag hvert enkelt af dem.

Kapitel 5 Tre slags sandsynlighed

Øvelse 9

Svarene er i rækkefølge: $\frac{1}{52}, \frac{4}{52} = \frac{1}{13}, \frac{12}{52} = \frac{1}{4}, \frac{8}{13}$

Øvelse 10

1) begge kort er røde? $\frac{26 \cdot 26}{52 \cdot 52} = \frac{1}{4}$, idet vi tænker på at det samlede udfaldsrum har $52 \cdot 52$ lige sandsynlige udfald.

2) begge kort er klør? $\frac{13 \cdot 13}{52 \cdot 52} = \frac{1}{16}$, idet vi igen tænker på at det samlede udfaldsrum har $52 \cdot 52$ lige sandsynlige udfald.

3) de har samme farve (rød eller sort)? $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, idet vi benytter svaret på 1) og indser at sandsynligheden for at begge er sorte også er ¼.

4) ¼

5) Det må give ½, da sandsynligheden for samme farve var ½ ifølge 3). Det er jo 100 procent sikkert at enten har de to kort samme farve eller forskellig farve.

6) Selv om teksten endnu ikke har gennemgået formlen for modsat hændelse, så synes det rimeligt, at sandsynligheden for forskellige kulør = 1 – sandsynligheden for samme kulør = 1 – ¼ = ¾

Kapitel 6 Kombinerede sandsynligheder, chancetræer og spil

Øvelse 1

Vi skal finde ud af, om nogle af disse principper kan anvendes i de efterfølgende eksempler:

1) $P(A \cup B) = P(A) + P(B)$, hvis A og B er disjunkte (det additive princip)

2) $P(\text{både } A \text{ og } B) = P(A) \cdot P(B)$, hvis A og B er uafhængige (multiplikationsprincippet)

3) $P(\text{både } A \text{ og } B) = P(A \mid B) \cdot P(B)$, gælder altid (det generelle princip)

1) Hændelserne er her: A = billedkort i klør, B = billedkort i spar. Da hændelserne er disjunkte kan vi benytte princip 1, så $P(A \text{ eller } B) = P(A) + P(B)$.

2) Her må hændelserne være: A = første kort er klør, B = andet kort er klør. Hændelserne er hverken disjunkte (andet kort kan godt være klør, selv om det første er det) eller uafhængige (hvis første kort er klør, nedsætter det sandsynligheden for, at det andet bliver det, da der tages fra samme spil kort), så vores eneste chance er at benytte 3). Den er da også velegnet fordi $P(\text{både } A \text{ og } B)$ netop bliver P(begge klør). Vi får P(begge klør) = P(det andet er klør|første er klør) $\cdot P$(første er klør) = $\frac{12}{51} \cdot \frac{13}{52}$

3) Kan beregnes som i delopgave 2), men da vi har uafhængighed, kan vi også bruge multiplikationsprincippet, og få svaret $\frac{1}{2} \cdot \frac{1}{2}$, idet A = første mønt viser krone, B = anden mønt viser krone.

4) Her er det umiddelbart svært at se, at der er to hændelser A og B involveret, men det må selvfølgelig igen være noget med første terning og anden terning. Faktisk kan man kun få et ulige produkt, hvis begge terningerne viser noget ulige. Så det ville være hensigtsmæssigt at definere

A = første terning viser ulige, B = anden terning viser ulige,

så P(ulige) = P(både første og anden viser ulige) = $\frac{1}{2} \cdot \frac{1}{2}$, hvor vi har anvendt princip 2, der er tilladt, da terningerne ikke påvirker hinanden.

Øvelse 2

1) Der er uafhængighed mellem de tre terningekast, så sandsynligheden er $\left(\frac{1}{6}\right)^3$

2) Der er uafhængighed mellem de timøntkast, så sandsynligheden er $\left(\frac{1}{2}\right)^{10}$

3) Hvis kortene trækkes samtidig (eller efter hinanden, uden tilbagelægning - hvilket er det samme) er der ikke uafhængighed mellem hændelserne.
P(3. billedkort| 1. og 2. er billedkort) = $\frac{10}{50}$ da der så er 10 billedkort blandt de resterende 50 kort.
P(1. og 2. billedkort) = P(2. er billedkort| 1. er billedkort)·P(1. er billedkort) = $\frac{11}{51}\cdot\frac{12}{52}$
P(1., 2. og 3 er billedkort) kan nu bestemmes som
P(3. billedkort| 1. og 2. er billedkort)·P(1. og 2. billedkort) = $\frac{10}{50}\cdot\frac{11}{51}\cdot\frac{12}{52}\approx 0{,}012$

4) Den rigtige position for hver af de seks "pinde" må anses for at være uafhængige af hinanden. Da der er 3 muligheder for hver pind bliver sandsynligheden $\left(\frac{1}{3}\right)^6$

5) Tegnene i de enkelte kampe må anses for at være uafhængige af hinanden. Der er tre muligheder i hver kamp, så sandsynligheden bliver $\left(\frac{1}{3}\right)^{13}$

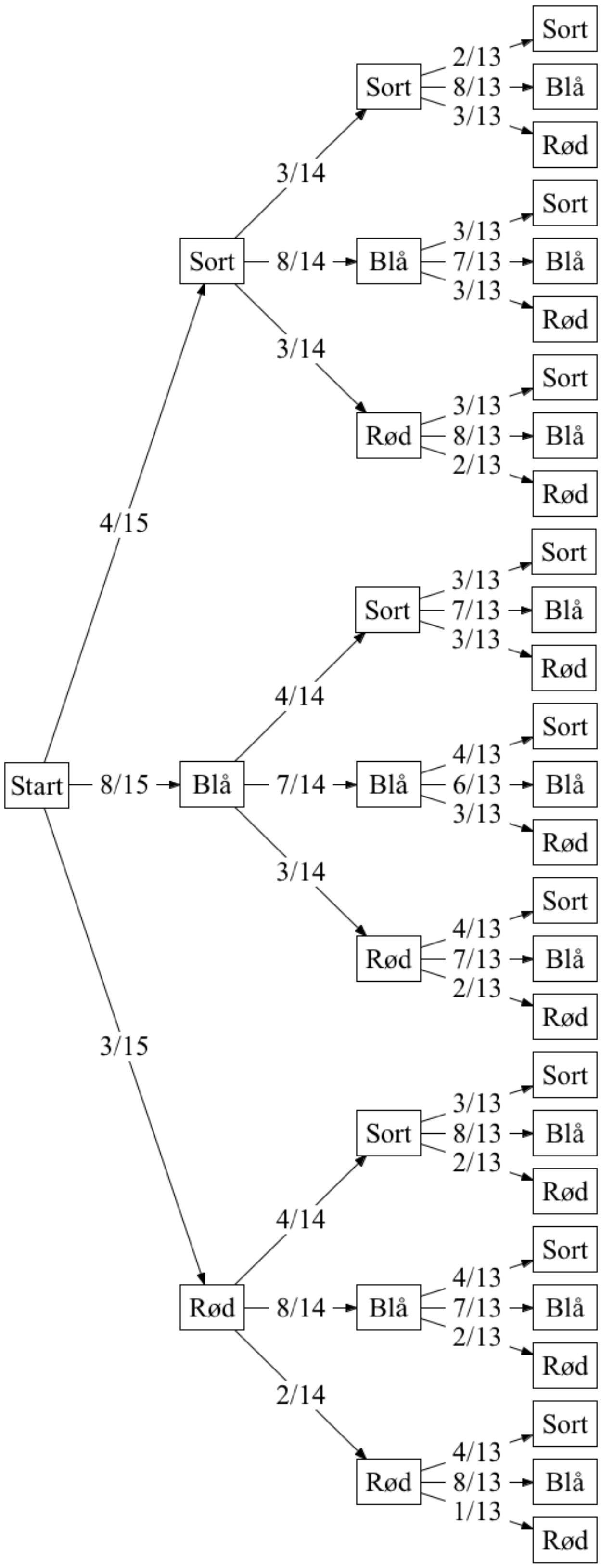
Øvelse 4

Hvis udtrækningen foregår uden tilbagelægning, vil det påvirke sandsynligheden. Lav selv en ny version af chancetræet i figur 3 og se, at der nu er sandsynligheder $\frac{6}{10}$, $\frac{4}{10}$, $\frac{6}{9}$, $\frac{5}{9}$, $\frac{4}{9}$ og $\frac{3}{9}$ langs grenene.

Hvis man udtrækker med tilbagelægning, er sandsynlighederne de samme i begge situationer.

Øvelse 5

I et chancetræ ser mulighederne således ud



1) Tre blå findes kun et sted i træet. Sandsynlighed $\frac{8}{15}\cdot\frac{7}{14}\cdot\frac{6}{13}\approx 0{,}1231$

2) To blå og en rød findes tre steder i chancetræet RBB, BRB, BBR. Sandsynligheden er
$\frac{3}{15}\cdot\frac{8}{14}\cdot\frac{7}{13}+\frac{8}{15}\cdot\frac{3}{14}\cdot\frac{7}{13}+\frac{8}{15}\cdot\frac{7}{14}\cdot\frac{3}{13}\approx 0{,}1846$

3) Alle tre forskellige farve findes seks steder
SRB, SBR, RSB, RBS, BRS, BSR.
Sandsynligheden bliver $\frac{4}{15}\cdot\frac{8}{14}\cdot\frac{3}{13}\cdot 6\approx 0{,}2110$

4) Man kan bestemme sandsynligheden for 1 eller 2 blå til 0,8000 og sandsynligheden for ingen blå til 0,0769.

Den forventede udbetaling i et enkelt spil er $5\cdot 0{,}1231 + 2\cdot 0{,}8000 + 10\cdot 0{,}0769 = 2{,}9845$.

Hvis spillet koster 3 kroner, vil man i gennemsnit tjene 1,55 øre pr. spil som spiludbyder, hvilket ikke kan siges at være overvældende, da der også kan være dage hvor du faktisk taber mange penge ved en tilfældighed.

Øvelse 6

Sandsynligheden for 13 rigtige er $\left(\frac{1}{3}\right)^{13}\approx 0{,}000000627$

Sandsynligheden for at få 12 rigtige er 26 gange så stor. Tænk på et chancetræ for spillet. Der vil være 13 stier med én forkert og 12 rigtige. Langs grenene med 12 rigtige er der 12 steder en sandsynlighed på $\frac{1}{3}$ og et sted (ved den forkerte) en sandsynlighed på $\frac{2}{3}$

I alt $13\cdot\frac{2}{3}\left(\frac{1}{3}\right)^{12}$

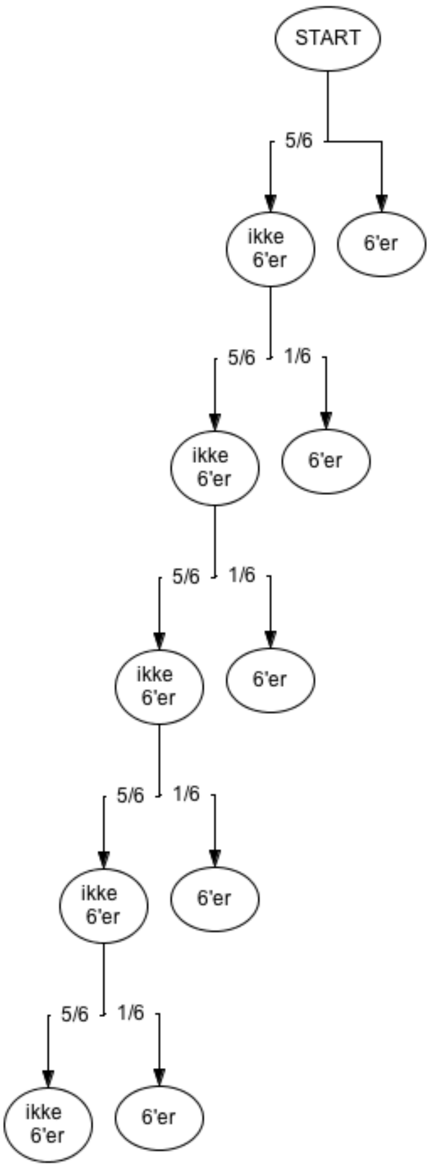
Øvelse 7

Her er en idé til at komme i gang. Sandsynligheden for at fx 5 personer alle har forskellig fødselsdag er $\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \frac{361}{365}$, idet vi ikke har problemer med den første person, men må kræve at den anden ikke har fødselsdag på samme dag som den første, hvilket giver sandsynligheden $\frac{364}{365}$, hvorefter sandsynligheden for at den tredje ikke har fødselsdag samme dag som en af de to første er $\frac{363}{365}$, osv. Vi ganger sandsynlighederne sammen med hjemmel i multiplikationsprincippet.

Hvis der er 10 personer er sandsynligheden for at alle har forskellig fødselsdag $\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{357}{365} \cdot \frac{356}{365}$.

Fortsæt denne tankegang til klassen med de 30 elever. For at gennemføre den faktiske beregning kan det være praktisk at tage et regneark i brug.

Øvelse 9
Situationen kan vises i et chancetræ



Sandsynlighed for at den kommer i 1. kast: $\frac{1}{6}$

Sandsynlighed for at den kommer i 2. kast: $\frac{5}{6} \cdot \frac{1}{6}$

Sandsynlighed for at den kommer i 3. kast: $\left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$

Sandsynlighed for at den kommer i *n*. kast: $\left(\frac{5}{6}\right)^{n-1} \cdot \frac{1}{6}$

Øvelse 10

Vi benytter det generelle multiplikationsprincip for flere hændelser efter hinanden, og vi skriver op i omvendt orden for at få det til at stemme med den tidslige udvikling:

P(1. klør, 2. spar, 3. hjerter og 4. ruder) = P(1. klør) $\cdot P$(2. spar|1. klør) $\cdot P$(3. hjerter|1. klør og 2. spar) $\cdot P$(4. ruder|1. klør og 2. spar og 3. hjerter) = $\frac{13}{52}\cdot\frac{13}{51}\cdot\frac{13}{50}\cdot\frac{13}{49}$, hvor den faldende nævner afspejler, at vi har færre kort at tage af efterhånden, mens den konstante tæller svarer til, at vi hver gang tager hul på en 'frisk' kulør, hvorfra vi ikke før har taget noget.

2) P(alle fire kort har forskellig kulør) = $\frac{13}{52}\cdot\frac{13}{51}\cdot\frac{13}{50}\cdot\frac{13}{49}\cdot 4\cdot 3\cdot 2\cdot 1$

eller direkte: $1\cdot\frac{39}{51}\cdot\frac{26}{50}\cdot\frac{13}{49}$, idet første kort kan vælges frit, mens det andet skal vælges blandt de 39 kort af en anden kulør osv.

Øvelse 11

A: 47 %, B: 10 %, AB: 6 % O: 37 %

Vi antager, at udvælgelse af de to personer sker tilfældigt, således at vi kan betragte deres blodtyper som uafhængige af hinanden.

1) P(begge A) = P(den ene A) $\cdot P$(den anden A) = $0{,}47\cdot 0{,}47$.

2) P(A og O **eller** O og A) [**eller** giver anledning til at bruge det additive princip]

= P(A **og** O) + P(O **og** A) [**og** giver anledning til at bruge det multiplikative princip]

dvs. vi får $0{,}47\cdot 0{,}37 + 0{,}37\cdot 0{,}47 = 0{,}3478$.

3) P(begge samme type) [vi bruger igen det additive princip]

= P(A,A) + P(B,B) + P(AB,AB) + P(O,O) [og nu det multiplikative princip]

= $0{,}47 \cdot 0{,}47 + 0{,}10 \cdot 0{,}10 + 0{,}06 \cdot 0{,}06 + 0{,}37 \cdot 0{,}37 = 0{,}3714$.

4) 1 – P(begge har samme type) = 0,6286.

Øvelse 12

Lav et chancetræ med 5 niveauer og faste sandsynligheder for "lus/ikke lus" i hver forgrening.

Øvelse 14

1) Uden tilbagelægning.

Simulering i *Kugle123* – sorte kugler er de røde i simuleringen. Modellen har 2 røde og 1 hvid kugle, og man udvælger 2 uden tilbagelægning.

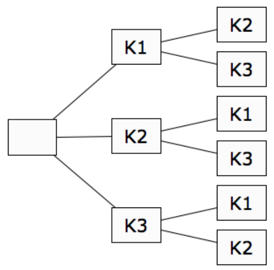
Antal røde kugler	Andel røde kugler	Antal eksperimenter	Antal i %	Kumuleret i %
1	0.500	1356	67.80	67.80
2	1.000	644	32.20	100.00

Udført i alt: 2000 eksperimenter

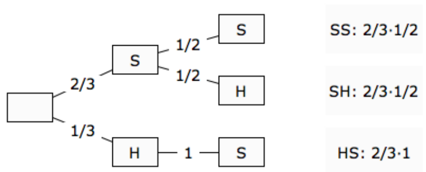
Gennemsnit: Antal røde kugler pr. eksperiment: 1.32
Andel - - - - - : 0.661
Sandsynligheden er 32,2% for at få 2 sorte.

Udfaldsrum. Hvis de to sorte kaldes K1 og K2 og den hvide kaldes K3, er der følgende muligheder (K1,K2), (K2,K1), (K1,K3), (K3,K1), (K2,K3), (K3,K2). Der er 2 ud af 6 muligheder, som giver to sorte Sandsynlighed $\frac{1}{3}$.

Tælletræ



Chancetræ

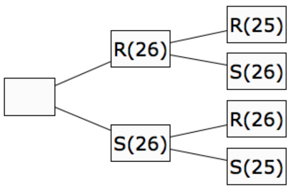


Spillet kunne også spilles med tilbagelægning.

3) Dette spil er uden tilbagelægning
Det kan simuleres med *Kugle123* med en model, hvor der er 52 kugler, hvor de 26 er røde. Man udtrækker 3 uden tilbagelægning.

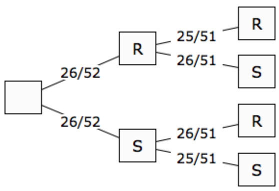
Det ureducerede tælletræ bliver uhåndterbart.

Reduceret tælletræ (vi viser det for et udtræk på to kugler og læseren kan selv supplere med en sidste forgrening for den tredje kugle):



Der er således 26·25 + 26·26 + 26·26 + 26·25 = 2652 muligheder i alt. Af disse er 26·25 = 650 med to røde.

Chancetræ (igen viser vi for udtræk på to kugler og læseren kan selv supplere med den sidste forgrening):



Sandsynligheden for 2 røde er så $\frac{26}{52} \cdot \frac{25}{51} = \frac{650}{2652}$
(Sandsynligheden for 3 røde som i opgaven i bogen giver 11,8 %)

Kapitel 7 Sandsynlighedsfordelinger

Øvelse 3

Antallet af veje er 1, 5, 10, 10, 5, 1

Øvelse 4

Hvis vi fx ser på antallet af muligheder for at få 3 seksere i 5 kast med en terning, kan disse deles op i de muligheder, hvor der kommer en sekser i første slag og i dem, hvor der ikke kommer en sekser i det første slag.

Hvis der er en sekser i det første slag, skal de fire sidste slag resultere i præcis 2 seksere. Dette kan pr. definition ske på K(4,2) måder.

Hvis der ikke er en sekser i første slag, skal der komme 3 seksere i de sidste 4 slag. Dette kan ske på K(4,3) måder.

I alt er der K(5,3) = K(4,2) + K(4,3) måder at få præcis 3 seksere i 5 slag med en terning.

Øvelse 5

Hvis vi fx ser på antallet af muligheder for at få $r \geq 1$ seksere i n kast med en terning, kan disse deles op i de muligheder hvor der kommer en sekser i første slag og i dem, hvor der ikke kommer en sekser i det første slag.

Hvis der er en sekser i det første slag, skal de $n - 1$ sidste slag resultere i præcis $r - 1$ seksere. Dette kan pr. definition ske på $K(n-1,r-1)$ måder.

Hvis der ikke er en sekser i første slag, skal der komme r seksere i de sidste $n - 1$ slag. Dette kan ske på $K(n-1,r)$ måder.

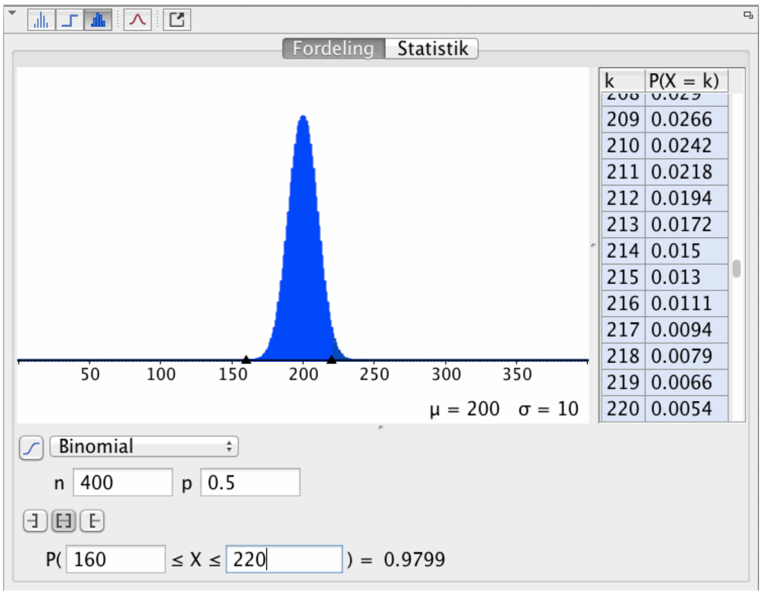
I alt er der $K(n,r) = K(n-1,r-1) + K(n-1,r)$ måder at få præcis r seksere i n slag med en terning.

Undersøgelse 1

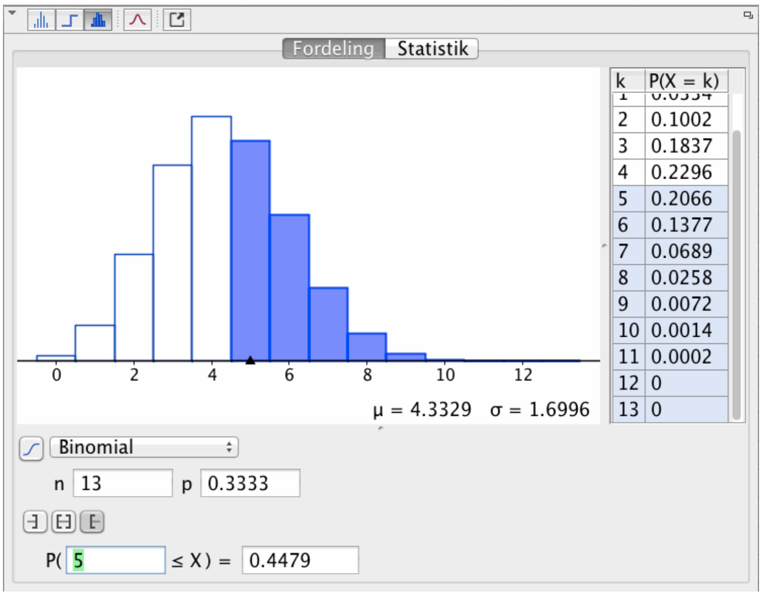
Summen i rækkerne bliver potenser af 2. Man kan argumentere for dette ved at observere, at et tal i en række kommer med to gange i den næste række pga. den måde, man opbygger trekanten på, hvilket betyder, at summen i den række nu bliver det dobbelte af summen i rækken over. Da den allerførste række (som her vil blive kaldt den 0'te række) er et 1-tal, bliver summen i næste række $2 \cdot 1 = 2$. I næste række igen får man $2 \cdot 2 = 2^2$, så bliver det 2^3 og så videre.

Øvelse 8

Med GeoGebras *Sandsynligheds Lommeregner* får man en sandsynlighed på ca. 98%.



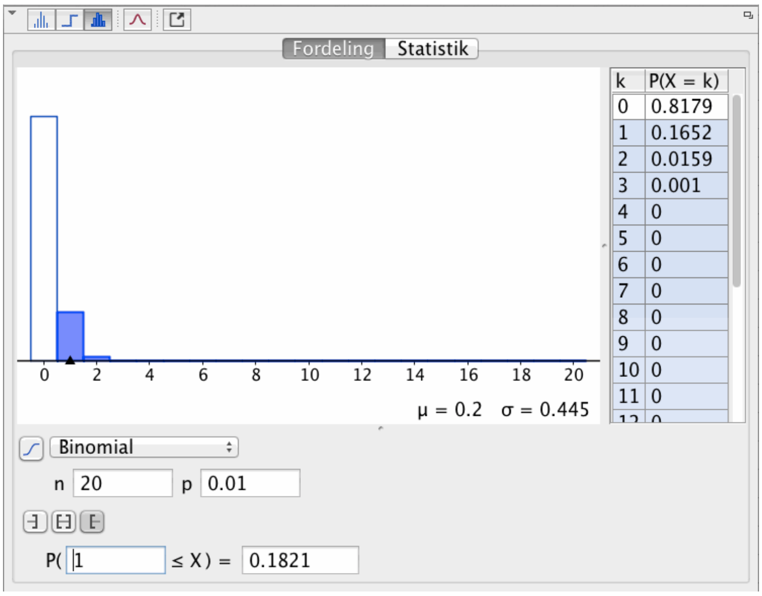
Øvelse 9



Sandsynligheden for mindst 5 rigtige er 0,4479.

Alle sandsynlighederne kan ses til højre i billedet.

Øvelse 10



Sandsynligheden for mindst et inficeret æg er 18,21%

Sandsynligheden for ingen inficerede æg er 81,79%

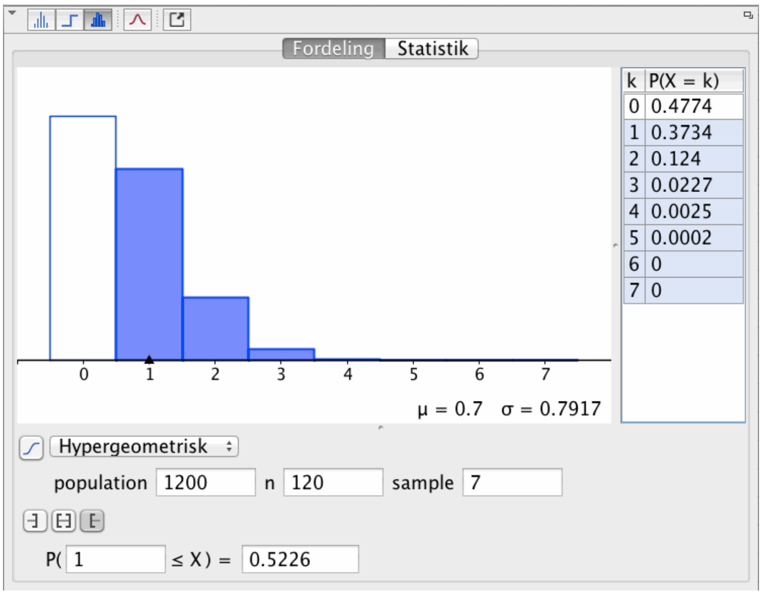
Øvelse 12

	Hypergeo-metrisk	Binomial	Procentvis afvigelse
0	0,58994	0,59049	-0,09%
1	0,32896	0,32805	0,28%
2	0,07270	0,07290	-0,28%
3	0,00796	0,00810	-1,79%
4	0,00043	0,00045	-4,31%
5	0,00001	0,00001	-7,90%

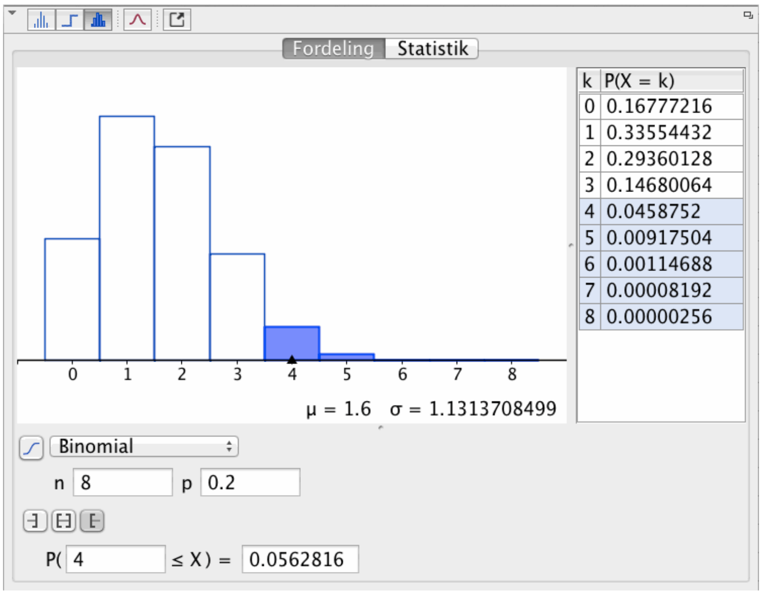
Der er ikke ret store absolutte afvigelser i tallene. Det ser ikke ud til at betyde ret meget, om man tænker på situationen med eller uden tilbagelægning, med mindre den procentvise afvigelse betyder noget.

Dette skyldes, at sandsynligheden for at få gevinst ikke ændrer sig ret meget, når man fjerner et lod. Hvis man fjerner et gevinstlod i første udtrækning, er gevinstchancen i næste udtrækning $\frac{119}{1199} = 0,0992$. Hvis man fjerne en nitte i første udtrækning, er gevinstchancen i næste udtrækning $\frac{120}{1199} = 0,100008$. Ingen af de to sandsynligheder er væsentligt forskellige fra 0,1.

Ved at prøve sig frem, kan man opdage at sandsynligheden for mindst én gevinst er 46,9% hvis man køber 6 lodder, og 52,3% hvis man køber 7. Man skal derfor mindst købe 7 lodder for at have 50% chance for mindst en gevinst.



Øvelse 13

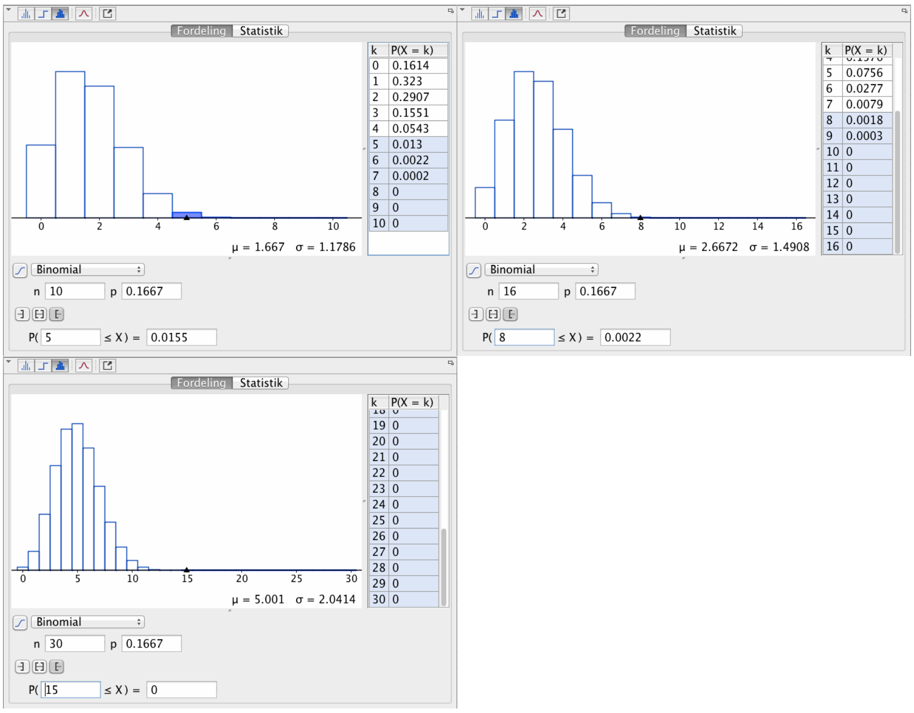


Sandsynlighed for 4 eller flere er 0,0563

Sandsynlighed for 0 er 16,8%

Sandsynlighed for 8 er 0,0003%

Øvelse 14



Det bliver sværere og sværere at få halvdelen til at være seksere. Det burde det også blive, da situationerne afviger mere og mere fra, at vi i gennemsnit forventer at 1/6 af kastene giver en sekser. I vores eksperimentelle fortolkning af sandsynlighed forventer vi jo at gennemsnittet ved mange kast skal være tæt på 1/6, hvorimod vi godt ved at, hvis der kun er seks kast, kan vi ofte forvente 0 eller 2 seksere og altså også sommetider 3 seksere eller halvdelen.

Øvelse 15
Som fjerdemand er du interesseret i at vide, hvor stor sandsynligheden er for at én af de tre første 3 trækker 'nitten'.

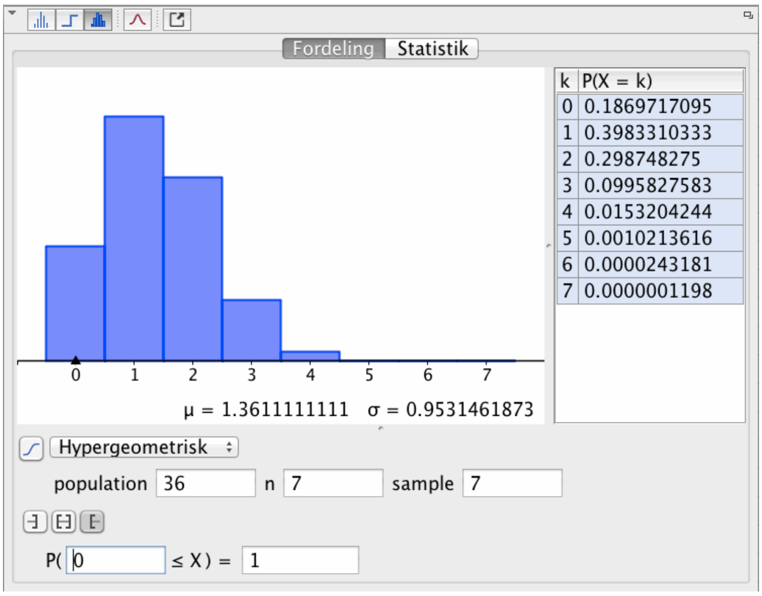
Uden tilbagelægning
Sandsynligheden for at ingen af de første tre trækker nitten er $\frac{3}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}=\frac{1}{4}$ Der er derfor 75 % chance for at én af de første tre trækker nitten. Til gengæld kan du så være sikker på at trække den, og din risiko bliver altså 25 %.

Med tilbagelægning
Sandsynligheden for at ingen af de første tre trækker nitten er $\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{3}{4}=\frac{27}{64}$ Der er derfor $1-\frac{27}{64}\approx 57,8\%$ chance for at én af de første tre trækker nitten. Men i $\frac{27}{64}$ af tilfældene bliver du altså tvunget til at trække. Fordelen er imidlertid, at risikoen for at trække nitten kun er $\frac{1}{4}$, så din risiko for at få nitten i første runde er $\frac{27}{256}$ eller 10,5 %.
Til gengæld risikerer du, at ingen de andre trækker nitten i anden runde, så du bliver tvunget til at trække igen. Sandsynligheden for dette er imidlertid ret lille nemlig $\left(\frac{3}{4}\right)^7\approx 13,3\%$, og herefter er din risiko for at trække nitten kun en fjerdedel. I alt er risikoen for, at du får nitten i anden runde kun $\left(\frac{3}{4}\right)^7\cdot\frac{1}{4}$ eller ca. 3 %.

Selvfølgelig risikerer du så får nitten i tredje runde eller i en af de uendeligt mange teoretisk mulige følgende runder. Men det, at risikoen for at få nitten i første omgang plus sandsynligheden for overhovedet at komme til at trække i en senere omgang er 10,5 % + 13,3 % og altså mindre end 25 %, fortæller dig, at du alligevel er bedst stillet med at vælge lodtrækningen med tilbagelægning.

Øvelse 16

Hvis man anskuer lottospillet som om vindertallene var givet på forhånd, ville det svare til, at 7 af de 36 tal havde en særlig egenskab. Antallet af rigtige på en kupon, handler så om, hvor mange af de særlige tal, man har valgt på sin kupon. Der er altså 36 kugler; 7 har en særlig egenskab og man vælger tilfældigt 7 blandt de 36. Antal rigtige er antallet af de specielle, der er med blandt de 7 man har valgt.



Kapitel 8 Stikprøver og estimation

Øvelse 1

En simulering med *Kugle123* og Kugle-Model K2(100,1,ja)

Antal udtagne kugler	Antal eksperimenter	Antal i %	Kumuleret %
1	26	1.30	1.30
2	17	0.85	2.15
3	13	0.65	2.80
4	20	1.00	3.80
5	12	0.60	4.40

Så der er ca. 4,5% chance for, at hændelsen sker inden for de første 5 år.

I forberedelsesmaterialet til eksamen for mellem- og sluttrin i maj 2013 kan man læse den følgende regel, som kan kaste lys over sjældne hændelser.

> 40-15 reglen
> Der er ca. 40 % chance for, at den sjældne hændelse indtræffer allerede inden for den første halvdel af den forventede periode.
> Der er ca. 15 % chance for, at den sjældne hændelse først indtræffer efter forløbet af det dobbelte af den forventede periode.

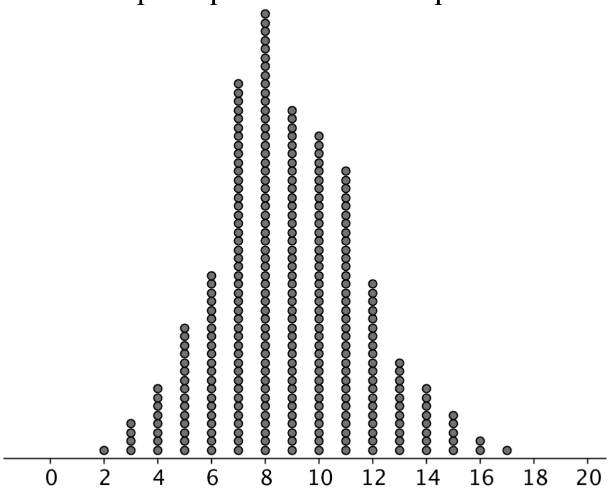
Prøv selv med *Kugle123,* om du kan se disse procentsatser.

Øvelse 5

At man ikke kan forventet at ramme præcis 9 i prøverne, er samme fænomen, der gjorde, at fx Kerrich (side 112 - 113) ikke ramte 50% plat i sine forsøg. Der er statistisk usikkerhed.

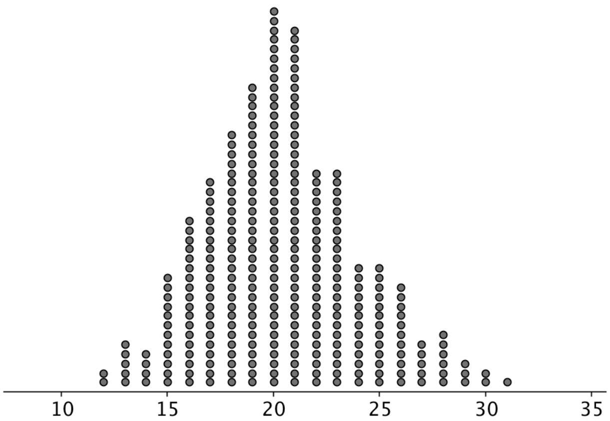
Øvelse 6

Her ses et punktplot over 300 stikprøver á 50 kugler.



Øvelse 7

Her ses et punktplot over 300 stikprøver á 100 kugler.



Stikprøverne giver altså bud på tilslutningen til partiet, der ligger mellem 12 % og 31 %.
Det er meget usikkert, hvad man egentlig kan sige om vælgertilslutningen ud fra stikprøverne.

Øvelse 8
Usikkerheden burde gå kraftigt ned.. Ved en simulering af stikprøver på 1000 "vælgere" bør andelen ligge mellem ca. 15 % og 25 %.

Øvelse 10

Ja, uanset om man regner med 1000 eller 1200 adspurgte, ligger den øvre grænse af intervallerne på over 35 % – og den rigtige andel burde være indfanget.

Øvelse 11

Intervallet bliver fra 53,3 % til 56,7 %.

Øvelse 12

Simulering med modellen Kugle 1 i *Kugle123* overlades til læseren.

Hvis man har en konstant p, og man så firdobler n, så ser man, at

$$\sqrt{\frac{p\cdot(1-p)}{4\cdot n}}=\frac{1}{2}\cdot\sqrt{\frac{p\cdot(1-p)}{n}}$$

Da kvadratroden af 4 er lig med 2.

Med andre ord halveres det tal, der angiver usikkerheden, når stikprøvestørrelses firdobles. Dette gælder helt generelt og derfor også i det konkrete taleksempel i opgaven.

Øvelse 14

Det er ikke nok, at Periferidemokraterne får 2% i undersøgelsen, altså 24 stemmer ud af de 1200, fordi der ville 95 %-konfidensintervallet være:

$$\left[0,02-1,96\cdot\sqrt{\frac{0,02\cdot(1-0,02)}{1200}};\ 0,02-1,96\cdot\sqrt{\frac{0,02\cdot(1-0,02)}{1200}}\right]=$$

$$[0,02-0,008;\ 0,02+0,008]=[0,012;\ 0,028].$$

Det ville sige, at den faktiske vælgertilslutning kunne være helt nede omkring 1,2 % – i hvert fald set ud fra denne særlige 95 %-konfidensintervalmetode. I virkeligheden kunne det selvfølgelig være meget værre, men sandsynligheden for det er meget lille.

Nu ville man nok gætte på 3 % og gennemføre beregningen igen:

$$\left[0,03-1,96\cdot\sqrt{\frac{0,03\cdot(1-0,03)}{1200}};0,03-1,96\cdot\sqrt{\frac{0,03\cdot(1-0,03)}{1200}}\right]=[0,03-0,0097;0,03+0,0097]=[0,0203;0,0297]$$

Vi ser, at nu ligger spærregrænsen under vores konfidensinterval, så mon ikke vi skulle anbefale Periferidemokraterne at nå op på 3% i en sådan undersøgelse, før de kan begynde at slappe lidt af i valgkampen. Men igen er der intet sikkert i statistik, men den kan hjælpe til med at træffe beslutninger i en usikker verden.

Man kan påstå, at vi var lidt heldige i vores gæt på 3 % ovenfor. Hvad nu hvis vi havde skudt langt ved siden af, skulle vi så blive ved med at afprøve gæt på gæt? Nej, i sådanne situationer sørger man for at gætte for højt og for lavt og så i midten osv. På den måde tager det ikke så lang tid at ramme målet. Men med et regneark bliver det hele langt lettere, for vi kunne lave en tabelindgang i kolonne A med de forskellige vælgertilslutninger, vi kunne forestille os opinionsundersøgelsen gav som resultat: fra 0,01 til 0,05 med spring på 0,001 og så lade regnearket lave beregningen af konfidensintervallet under anvendelse af 'fyld nedad'. Så ville det ikke tage lang tid før, man havde en god tabel, der kunne danne grundlag for partiledelsens beslutninger.

Kapitel 9 Hypotesetest

Øvelse 1

1) Med $N = 1000$, $n = 40$ og $D = 30$ giver den Hypergeometriske fordeling, at sandsynligheden for at få 3 eller flere defekte i stikprøven er 0,1104. Dette betyder, at der er en risiko på 11,04 % for at vores stikprøve giver så mange defekte, at vi griber ind i produktionsgangen, selvom fejlprocenten faktisk holder sig på 3.

2) Med $N = 1000$, $n = 40$ og $D = 50$ giver den Hypergeometriske fordeling, at sandsynligheden for at få 2 eller færre defekte i stikprøven er 0,6771. Dette betyder, at der er 67,71% risiko for, at vores stikprøve får os til at konkludere, at produktionen kører OK, selvom fejlprocenten er steget til 5. Der er meget stor risiko for, at stikprøven ikke får os til at opdage, at der er noget galt.

3)

Fejlprocent	Sandsynlighed for 2 eller færre defekte i stikprøven
10	0,2170
15	0,0455
20	0,0071

Tallene er atter beregnet ved hjælp af den Hypergeometriske fordeling.

Vi ser, at stikprøven er ret dårlig til at opdage, hvis der er 3 gange for mange defekte skruer.

Øvelse 2

Ud fra binomialfordelingen $b(25;0,2)$ kan man beregne sandsynligheden for, at man tilfældigt kommer under 4 gevinster. Sandsynligheden er 42,07%.

For at få procenten ned, kan man sætte grænsen ned, eller man kan købe flere lodder.

Hvis grænsen sættes ned, bliver det sværere at komme til at lave vrøvl. Til gengæld bliver det også sværere at opdage, hvis der faktisk er for få gevinstlodder.

At købe flere lodder er en sikrere metode – men den koster flere penge.

Øvelse 4

X er antal 'rød' ved 100 spil på en roulette.

j	$P(X \leq j)$
30	0,0001
31	0,0003
32	0,0005
33	0,0011
34	0,0022
35	0,0040
36	0,0072
37	0,0125
38	0,0207
39	0,0332
40	0,0511
41	0,0760

Nulhypotesen er: Rouletten er OK.

Alternativ hypotese: Rouletten giver for få 'rød'.

Vi beslutter ud fra et udbredt ønske om at holde procenten af fejl af første art under 5 %, og altså at nulhypotesen forkastes, hvis der er 39 eller færre 'røde' i 100 spil.

Fejlen af 1. art bliver så 0,0322 dvs. 3,22 %

Hvis rouletten er manipuleret og der fx. kun er 40 % chance for 'rød', er sandsynligheden for at få 40 eller flere 'røde' i 100 spil 53,79 % – dette er fejlen af 2. art i dette tilfælde.

Hvis rouletten er manipuleret mere groft så der fx. kun er 35 % chance for 'rød', er sandsynligheden for at få 40 eller flere 'røde' i 100 spil 7,24 % – dette er fejlen af 2. art i dette tilfælde.

I lyset af disse resultater burde vi nok prøve at forbedre vores test. Vi kunne lade procentdelen af fejl af første art stige for på den måde at få procenten af fejl af anden art til at falde: måske kunne vi ændre grænsen på de 39 til 41 i stedet og så lave en ny gennemregning. Hvis ikke det nytter tilstrækkeligt, kunne vi overveje at lade stikprøven stige til 200 spil eller mere.

3)

X er antal rigtige svar blandt de 40, når man blot gætter:

J	P($X \leq j$)
0	0,0000
1	0,0001
2	0,0010
3	0,0047
4	0,0160
5	0,0433
6	0,0962
7	0,1820
8	0,2998
9	0,4395
10	0,5839
11	0,7151
12	0,8209
13	0,8968
14	0,9456
15	0,9738
16	0,9884
17	0,9953
18	0,9983
19	0,9994
20	0,9998
21	1,0000

Hvis eleven har 15 eller flere rigtige svar, er sandsynligheden for, at hun har svaret tilfældigt på 1 – 0,9456 = 0,0544. Sagt på en anden måde, så vil 1 ud af 20 elever, der intet ved og bare gætter klare en test, der sætter grænsen ved 15 rigtige. Om det er acceptabelt afhænger af formålet med testen. Det ville ikke gå til en teoretisk køreprøve.

Hvis eleven har 16 eller flere rigtige svar, er sandsynligheden for, at hun har svaret tilfældigt på 1 – 0,9738 = 0,0262. Er den lille nok til, at man vil tro på, at eleven ikke blot har gættet?

4) Hvis Jonathan blot gætter, er sandsynligheden for at gætte den første øl korrekt $\frac{1}{3}$. Efterfølende er sandsynligheden $\frac{1}{3}$ for, at han gætter den næste øl korrekt. Hvis de to første øl er gættet korrekt, vil den sidste også blive angivet korrekt. Under forudsætning af at han ikke kan smage forskel, er der derfor en sandsynlighed på $\frac{1}{6}$ for, at han alligevel gætter alle tre øl.
Da vi er tilbøjelige til at tro, at han rent faktisk kan smage forskel, hvis alle tre øl gættes korrekt, er der i hypotesetest-sprog derfor en fejl af 1. art på 16,67 %.

Ved at gentage eksperimentet flere gange, kan vi få fejlen af første art ned.
Hvis vi insisterer på, at Jonathan skal svare korrekt på alle tre øl to gange, er sandsynligheden for dette $\left(\frac{1}{6}\right)^2 \approx 0,0278$ og insisterer vi på, at han skal svare korrekt 3 serier i træk, er sandsynligheden for, at det er tilfældigt blot $\left(\frac{1}{6}\right)^3 \approx 0,0046$. Fejlen af 1. art vil så være under 1 %.

Hvis vi har endnu højere ambitioner med vores øltest, så kan det være at vi vælger denne skrappe udgave af testen, hvor der igen skal smages på 9 glas, men hvor der hver gang er frit valg af, hvilket af de tre mærker der skal iskænkes. Chancen for, at vi i denne situation lader en ignorant slippe gennem, er $\left(\frac{1}{3}\right)^9 \approx 0,0005$, så fejlen af første art har nu en forsvindende størrelse.

Hvis Jonathan faktisk er en god kender af øl og ligger på 95 % rigtige inden for de angivne ølmærker, så ville vi jo på den anden side godt have, at testen lod ham bestå. Men hans chance for at bestå er blevet på $0,95^9 \approx 0,63$, altså 63 %. Så sandsynligheden for, at vi får ham forkastet (laver fejl af 2. art), er på 37 %.

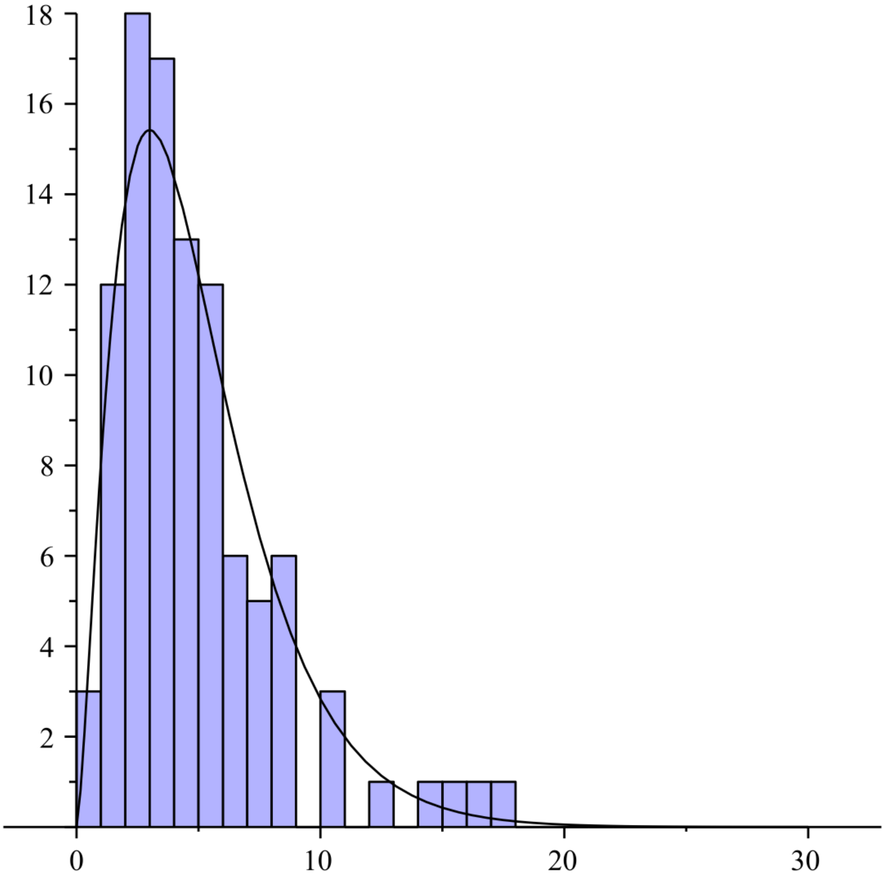
Med mindre noget særligt gør sig gældende, må man sige, at der er meget skæv fordeling mellem fejl af første og anden art i denne situation. Kunne man ikke tillade en lidt større fejl af første art for drastisk at få reduceret fejl af anden art. Hvilken virkning ville det for eksempel have at tillade forsøgspersonen en enkelt fejlplacering ud af de ni skud, altså sætte kriteriet for at gå videre til 8 rigtige ud af 9?

Øvelse 7

$(obs_1 - forv_1) + (obs_2 - forv_2) + ... + (obs_n - forv_n) = obs_1 + obs_2 + ... + obs_n - forv_1 - forv_2 - ... - forv_n$ = summen af de observerede – summen af de forventede = 0.

Øvelse 8

Sådan kan data fra simuleringen se ud. Den fuldt optrukne kurve er den tilsvarende eksakte χ^2 – fordeling.



Øvelse 10

2) Der er 556 planter i alt

Observeret	315	108	101	32
Forventet	$\frac{9}{16}$ af 556= 312,75	$\frac{3}{16}$ af 556= 104,25	$\frac{3}{16}$ af 556= 104,25	$\frac{1}{16}$ af 556= 34,75

Beregningen af χ^2 giver 0,47. Testen har 3 frihedsgrader. Sandsynligheden for at komme op på 0,47 ved tilfældighedernes spil er 92,5%. Det er altså noget, der ofte sker, så vi må konkludere, at data og model passer sammen.

Man kan også sige, at sandsynligheden for at få en så lille værdi for χ^2 er 100 – 92,5 = 7,5 %, så det vil være ret sjældent at få så god en overensstemmelse mellem eksperiment og teori. I forbindelse med videnskabeligt arbejde og for den sags skyld fysikrapporter i gymnasiet vil en så god overensstemmelse mellem teori og eksperiment sommetider motivere en nærmere undersøgelse af, om en eller anden har fiflet ved de eksperimentelle data (altså snydt) for at tilfredsstille formodede krav fra omverdenen.

Øvelse 11

Observeret	11	8	2	2	1
Forventet	41% af 24 = 9,84	43% af 24 = 10,32	7% af 24 =1,68	5% af 24 = 1,20	4% af 24 = 0,96

Beregningen af χ^2 giver 1,25. Testen har 4 frihedsgrader. Sandsynligheden for at komme op på 1,25 ved tilfældighedernes spil er 89,6%, så der er intet, der tyder på, at din klasse skulle være afvigende fra de andre 9. klasser. Vi må konkludere, at data og model passer sammen.

Afsluttende undersøgelser

Ingen løsningsforslag