12. Bevisets stilling

Opgave 4
Hvorfor dur $(x,y) = (1,-3)$ ikke? Det er et spørgsmål, der kan give anledning til forskellige overvejelser. Her beskrives en af dem.

Vi har den mindste sum: 13 + 27 = 40, og den største sum: 28 + 23 = 51.
Vi skal finde to tal, hvor hvert af dem adderet til hver addend i den mindste sum giver en sum, der ligger mellem 40 og 51. Altså et talpar (x,y) der gør følgende sandt:

$$40 < (13 + x) + (27 + y) < 51$$

Vi behøver ikke at gøre det helt så besværligt, vi kan blot tage summen af x og y og addere denne til 40 og hurtigt se, om summen ligger mellem 40 og 51. Altså kontrollere om følgende er sandt:

$$40 < 40 + x + y < 51$$

Talparret (1,-3) giver ved addition til 40: $40 + 1 + (-3) = 38$, hvilket ikke dur.

Vi kan nu se, at det talpar, vi skal addere til den mindste sum, skal give en sum, der er større end nul. Altså har vi

$x + y > 0$ som er ensbetydende med, at $y > -x$

Det grafiske billede heraf giver et område, der ligger over den rette linje $y = -x$ gennem (0,0).

Men kan vi så bruge alle talpar i dette område? Det fører os ind på spørgsmålet: Hvorfor skal summen af x og y være mindre end 11, som der står i teksten over figur 13? Hvad sker der, hvis man sætter $x = 6$ og $y = 7$? Hvad sker der, hvis man vælger $x = -36$ og $y = 44$? Forsøg selv. Det generelle svar følger på side 531.

Opgave 5
Det mærkelige i den oprindelige opgave med at finde summer mellem

$$\begin{array}{r} 24 \\ \underline{+18} \end{array} \text{ og } \begin{array}{r} 37 \\ \underline{+15} \end{array}$$

er, at det indlysende svar, hvor man finder den ene addend mellem 24 og 37 og den anden addend mellem 18 og 15 ikke nødvendigvis giver en sum mellem 42 og 52. Det er formodentlig det, der giver vanskeligheder, idet der ikke ville have været noget problem, hvis de to summer havde været

$$\begin{array}{r} 24 \\ \underline{+15} \end{array} \text{ og } \begin{array}{r} 37 \\ \underline{+18} \end{array}$$

for der ligger en almen sandhed bag, som følger af de enkleste love for ulighedsregning, nemlig:

Hvis $a < x < c$ og $b < y < d$, så er $a + b < x + y < c + d$,

hvilket er noget som elever i små klasser godt kan forstå, forudsat det bliver forklaret med noget i retning af, at hvis man lægger to store tal sammen, så får man noget større, end hvis man lægger to små tal sammen.

Hvis vi vil bruge denne teori på den oprindelige opgave, så får vi

$$\begin{array}{r} a \\ \underline{+b} \end{array} \text{ og } \begin{array}{r} c \\ \underline{+d} \end{array}$$

og vi skal blot sørge for, at vi har additionsstykker, der fra starten opfylder, at $a < c$ og $b < d$.

13. Sprog og definitioner

Øvelse 1

Vi giver et eksempel, hvor begrebets ekstension selvsagt er subjektivt i den forstand, at den eksempelsamling forskellige mennesker råder over er forskellig.

Begrebet*: Geometrisk sted*

Begrebets intension:
Kan her naturligt angives ved dets definition: et geometrisk sted er mængden af punkter i planen, for hvilke det gælder, at alle punkter har en fælles geometrisk egenskab.

Begrebets ekstension omfatter alle eksempler, vi kan kun give nogle få:
Det kan være en cirkelperiferi, hvor alle punkterne 'på blyantstregen', der afgrænser cirklen, altså cirkelperiferien har samme afstand til centrum. Eller det kan være en vinkelhalveringslinje, hvor den vinkelrette afstand til begge vinklens ben, er den samme for ethvert punkt, man vælger på denne. Endelig kunne det også være en midtnormal til to punkter A og B, fordi denne består af en punktmængde, hvor det for et vilkårligt valgt punkt gælder, at afstanden til hhv. A og B er lige stor.

Et andet eksempel har vi i funktionsbegrebet. En funktion i mængden $\{1,2,3,4,5,6,7,8,9,10,11,12\}$ kan have intensionen at vise, hvad klokken er tre timer senere. Hvilket kan udtrykkes med forskriften $f(x) = x + 3(\text{mod } 12)$.

Den samme funktion kan også gives ved dens ekstension jf. det vi har kaldt Kyeds definition s. 437. Funktionens ekstension bliver derefter: $f = \{(1,4);(2,5);(3,6);(4,7);...;(9,12);(10,1);(11,2);(12,3)\}$.

Øvelse 4

I øvelsen står der: "Der har allerede været nævnt to muligheder i starten af kapitlet." Det er desværre en fejl, da det drejer sig om elevernes diskussion af lige og ulige i Mikkels 3.c i kapitel 12. Se evt. starten på: Overvej-diskuter 1 i kapitel 12, side 520.

Svar:
En typisk matematisk definition på de lige tal, er de tal, der kan skrives på formen: $2 \cdot n, \quad n \in \mathbf{N}$.

Og de ulige tal defineres som tal, der kan skrives på formen: $2 \cdot n - 1, \quad n \in \mathbf{N}$.

Hvis eleverne i 3. c skulle have lavet en definition, ville de nok have sagt, at et antal kastanjer er ulige, hvis de ikke kan lægges i to lige lange rækker, og den ene række bliver 1 kortere end den anden. Eller at de ulige tal er dem, der ender på 1, 3, 5, 7 og 9.

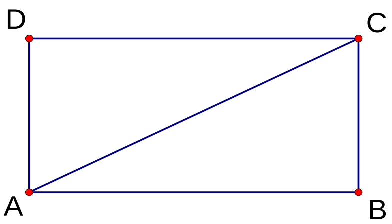
14. Beviser og logik

Øvelse 1
Beviset kan overføres bogstav for bogstav, idet dog definition 1 erstattes med definition rum 1.

Øvelse 2
Hvis man definerer konveksitet i det givne rum således: En mængde K kaldes konveks, hvis der for to vilkårligt valgte punkter A og B i K gælder, at linjestykket $[A,B]$ også ligger i K (altså akkurat det samme som i definition 1).

Vælges denne definition, bliver fællesmængden af konvekse mængder altid konveks. Beviset er det samme som for sætning 1.

Øvelse 3



Vi kalder firkanten $ABCD$ og vil vise, at $|AB|$ er mindre end summen af de tre andre sider. Altså $|AB| < |BC| + |CD| + |DA|$.

Bevis:
Diagonalen AC tegnes, og vi benytter trekantuligheden på $\triangle ABC$, hvilket giver $|AB| < |BC| + |CA|$.

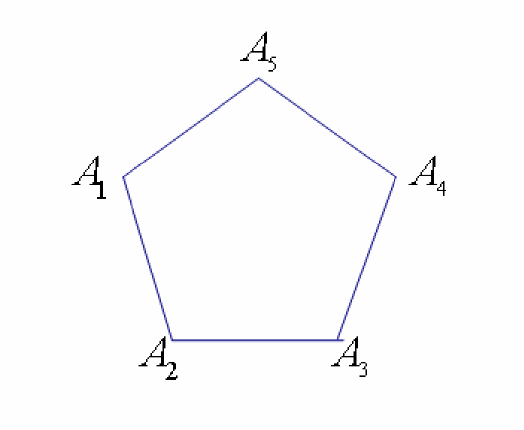
Dernæst benyttes trekantuligheden på $\triangle ACD$, hvilket giver $|AC| < |CD| + |DA|$.

Kombineres dette med den forrige ulighed fås det ønskede $|AB| < |BC| + |CD| + |DA|$.

Opgave 1

Vi skal vise sætning 2:

I enhver n-kant er enhver side mindre end summen af de øvrige $n - 1$ kanter.



Vi har vist, at sætningen gælder for tre- og firkanter. Vi viser den nu for en femkant $A_1A_2A_3A_4A_5$. Vi vil altså vise, at $A_1A_2 < A_2A_3 + A_3A_4 + A_4A_5 + A_5A_1$.

Bevis:

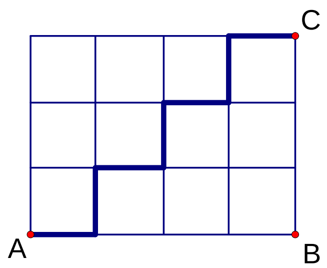
Diagonalen A_1A_3 tegnes. Da $A_1A_3A_4A_5$ er en firkant, gælder ifølge sætning 2 brugt på en firkant, at $A_2A_3 < A_3A_4 + A_4A_5 + A_5A_1$

Benytter vi trekantuligheden på $\triangle A_1A_2A_3$ fås $A_1A_2 < A_2A_3 + A_3A_1$

Kombineres de to uligheder fås det ønskede.

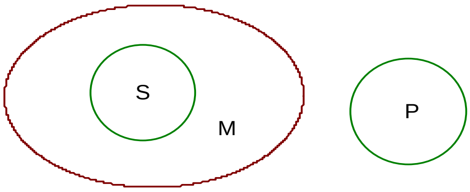
Vi har hermed vist sætningen for en femkant og kan benytte dette som afsæt til at bevise den for en sekskant ved akkurat samme argument osv. Det er dette 'osv.' man kalder et induktionsbevis, men vi skriver det ikke formelt op her.

Øvelse 7 kun 3)



Hvis vi lader linjestykkerne følge vejene gælder trekantuligheden. Men lighedstegnet gælder ofte, og det er altså ikke noget krav, at B ligger på 'stykket AC', jf. figuren, hvor $AC = AB + BC$. Sagen er, at linjestykket AC ikke er veldefineret, og selv om vi ønsker at definere det, så kommer vi let i vanskeligheder.

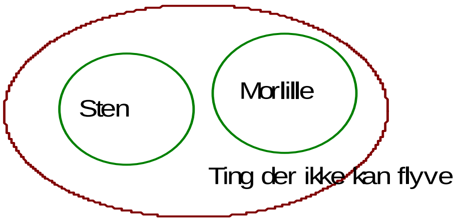
Øvelse 8
Den logiske slutning kan illustreres således:



Det fremgår tydeligt af diagrammet, at ingen S kan være P.

Øvelse 9
Man kan slutte, at alle kvadrater er parallelogrammer.

Øvelse 10
Det fremgår af diagrammet, at Morlille og en sten ikke har noget tilfælles. Dvs. man kan ikke slutte noget fra den ene til den anden.



Øvelse 11
Vi kan slutte, at grunden ikke er rektangulær.

Øvelse 12
Fejlagtige slutninger.
Man ser sommetider følgende slutninger brugt. Hvad er der galt ved dem?

C.	1) Hvis p så q <u>2) q er sand</u> p er sand	D.	1) p eller q <u>2) p er sand</u> q er falsk

C.
Et eksempel på at det kan gå galt, hvis C. anvendes:
Selv om q er sand, behøver p ikke at være det. Fx har vi det sande udsagn: Hvis x er et sammensat tal, så er $3 \cdot x$ et sammensat tal, hvor
p: x er et sammensat tal.
q: $3 \cdot x$ er et sammensat tal.

Men vi kan ikke som i C. slutte fra, at q er sand til at p er sand.
Hvis fx $x = 5$, så er q sand, og p er falsk, idet 15 er et sammensat tal, mens 5 er et primtal. Konklusionen i C. kan være forkert. Altså er selve slutningsformen forkert.

I hverdagslivet ser man dog C. brugt fx om læring og testning. Hvis en elev har lært meget matematik i skolen, så vil denne elev klare en test godt (hvis p så q). Eleven klarer en test godt (q), altså har han lært meget matematik i skolen (p). Konklusionen (p) følger ikke strengt logisk, da det, testen måler, ikke behøver at være i overensstemmelse med 'lært meget matematik'. Det er faktisk svært at udforme tests, der kommer om ad hele det store spektrum, der ligger i 'lært meget matematik'.

D.
Et eksempel på at det kan gå galt, hvis D anvendes ukritisk:
p: figuren er et rektangel.
q: figuren er en rombe.

Hvis i overensstemmelse med præmisserne i D vi har en figur, der enten er et rektangel eller en rombe, og vi yderligere ved, at det er en rombe, så kan figuren faktisk være et rektangel, nemlig hvis figuren er et kvadrat. Påstanden q behøver altså ikke være falsk.
I dagligsproget bruger man nogle gange 'eller' eksklusivt (udelukkende) sådan, at begge påstande ikke må være sande samtidig. Som en statsleder kunne sige: Enten er I med mig, eller også er I imod mig. Men i logikken kan p og q begge være sande i påstanden 'p eller q'.

Opgave 2
Man skal vende E for at undersøge, om der er et lige tal på den anden side, og man skal vende 7 for at sikre sig, at der ikke er en vokal på den anden side. Da man ikke ud fra "Hvis der en vokal på den ene side, så er der et lige tal på den anden" kan slutte noget som helst ud fra et K eller 4, behøver man ikke at vende disse to kort.

Øvelse 13
For at kunne afgøre om de følgende par af udsagn er ækvivalente, hvilket vil sige, at de er sande for de samme kombinationer af p og q, udfylder vi først tavlen herunder. Denne anvendes sammen med sandhedstavlen i bogen.
1) ‘Hvis p så q’ er ækvivalent med ‘¬p eller q’
2) ‘p og q’ er ækvivalent med ‘¬(¬p eller ¬q)’

p	**q**	¬p	¬q	¬p eller q	Hvis p så q	¬p eller ¬q	¬(¬p eller ¬q)
s	**s**	f	f	s	s	f	s
s	**f**	f	s	f	f	s	f
f	**s**	s	f	s	s	s	f
f	**f**	s	s	s	s	s	f

Svar: ‘Hvis p så q’ er ækvivalent med ‘¬p eller q’ og ‘p og q’ er ækvivalent med ‘¬(¬p eller ¬q)’,da de er sande for samme kombinationer af p og q.

Bemærkning: Man kan også forklare disse ækvivalenser sprogligt, men det er svært. Hvis der er tale om mere komplicerede ækvivalenser, så er denne metode med sandhedstavler den eneste sikre. Den kan nemlig udføres rent mekanisk efter ret enkle regler. Prøv fx at lave beregningerne som i skemaet ovenfor i et regneark, hvor den klassiske repræsentaion af s og f er 1 og 0. Så kan ‘p og q’ udregnes som ‘p gange q’.

Øvelse 17
1) Ud fra definition 3 skal vi vise, at:

$7 \mid 91$, hvilket er sandt, da $91 = 7 \cdot 13$, kvotienten er 13.

$13 \mid 143$, hvilket er sandt, da $143 = 13 \cdot 11$, kvotienten er 11.

2) Vi viser, at ‘der findes et helt tal q, således at $D = d \cdot q$’ er ækvivalent med ‘$\frac{D}{d}$ er et naturligt tal’.

‘Der findes et helt tal q, således at $D = d \cdot q$’, ses ved division med d på hver side af lighedstegnet at være ensbetydende med, at der findes et helt tal q, således at $\frac{D}{d} = q$, hvilket er det samme som, at $\frac{D}{d}$ er et naturligt tal, da D og d begge er naturlige tal. De to udsagn er således ækvivalente.

En alternativ definition af ‘går op i’ ville som følge heraf blive: Hvis d og D er naturlige tal, så siger vi, at d går op i D, hvis $\frac{D}{d}$ er et naturligt tal.

Bemærk, at vi i definition 3 har benyttet udtrykket ‘et helt tal’ hvor vi mere præcist kunne have skrevet ‘et naturligt tal’, idet hele tal også omfatter negative tal. Det væsentlige her er imidlertid, at tallet skal være helt i modsætning til et bruddent tal (en brøk).

Øvelse 19
Det vises, at hvis d, D og n er naturlige tal, så gælder: Hvis $d \mid D$, så $d \mid (n \cdot D)$
$d \mid D$ betyder ifølge definition 3, at der findes et helt tal q, således at $D = d \cdot q$. Ganger vi her med n på hver side af lighedstegnet, får vi $n \cdot D = n \cdot d \cdot q$ eller $n \cdot D = d \cdot (n \cdot q)$, hvoraf det ses, at $n \cdot D$ ved division med q giver den hele kvotient $n \cdot q$.

Det havde været lidt lettere at benytte den alternative definition fra øvelse 17, hvor det så drejer sig om at vise, at $\frac{n \cdot D}{d}$ er et helt tal, fordi $\frac{D}{d}$ er det, men det er indlysende, da tallet er n gange større.

Opgave 3
1) Fordi en forkortelig brøk $\frac{m}{n}$ altid kan forkortes til en uforkortelig brøk $\frac{p}{q}$.

Bemærk, at $\frac{m}{n} = \frac{p}{q}$, så der er i talmæssig forstand ikke tale om en anden brøk. Derfor kan vi antage, at den brøk, vi begynder med, er den uforkortelige brøk $\frac{p}{q}$.

2) Første del af beviset bygger på, at $\frac{p^2}{q^2}$ skal være en uforkortelig brøk, og det bliver den kun, fordi $\frac{p}{q}$ er det.

15. Geometri: Fænomener og beviser

Opgave 1
1) Vi antager, at l er parallel med m, og m er parallel med n. Vi skal vise, at l er parallel med n. Det gør vi med et indirekte bevis.

Vi antager, at l skærer n i et punkt P. Gennem punktet P har vi nu ifølge det forudsatte to linjer, der er parallelle med m, men ifølge aksiom 3 kan der igennem punktet P kun trækkes en linje parallel med m. Altså er l lig med n. Vi har hermed vist, at enten er l parallel med n eller også er l lig med n, og så er de også parallelle (definition 6).

2) Det er ikke noget bevis inden for den Euklidiske ramme, fordi det ikke bygger på definitioner, aksiomer og sætninger, vi allerede har vist, men derimod bygger på noget, vi af anden vej ved om parallelle linjer.

Øvelse 3
1) Tegn en skitse af firkanten for at følge argumentet.

Da $\angle B + \angle C = 180°$ skæres linjen BC på en sådan måde, at ensliggende vinkler er lige store, idet vinklen, der dannes af BC's forlængelse ud over C og CD er $80°$. Det følger nu af følgesætning 2b, at AB og CD er parallelle. Dette er definitionen på, at $ABCD$ er et trapez.

2) Det fremgår af eksempel 1 i kapitel 16 s. 648, at de vinkler, som betegnes BAC og ACD er lige store. Derfor er DC parallel med AB ifølge sætning 2b.

Opgave 5
Den omvendte sætning må hedde: Hvis en trekant ABC har $\angle A = \angle B$, så er trekanten ligebenet, altså $AC = CB$.

Bevis (tegn selv):
Vi tegner vinkelhalveringslinjen CM i trekanten, de to trekanter CAM og CMB er kongruente, fordi de har to vinkler og dermed også den tredje vinkel parvis lige store, og de har CM fælles. Altså er $AC = CB$, hvilket skulle vises.

Alternativ:
Hvis man ikke er helt fortrolig med at bruge kongruenssætninger, kan man klare sig med et spejlingsargument, hvor man spejler trekanten i midtnormalen til AB. Men her har man i første omfang det problem, at man ikke kan vide, om midtnormalen går gennem C. Det kan man dog ret hurtigt argumentere for.

Øvelse 8
1) Lad de to trekanter hedde ABC og $A'B'C'$ med $\angle A = \angle A'$, $AB = A'B'$ og $AC = A'C'$. Vi flytter trekanten ABC hen, så AC dækker $A'C'$ og $\angle A$ falder i $\angle A'$, så vil siden AB falde ud ad siden $A'B'$, og da de er lige store, vil B falde i B', derfor vil også BC falde i $B'C'$, og trekant ABC falder lige akkurat i $A'B'C'$. Trekanterne er derfor pr. definition kongruente.

2) Det ville have været hurtigere, fordi man med reference til K3 umiddelbart kan se, at trekant *ACM* er kongruent med trekant *BCM*.

Øvelse 9
Vi definerer at:

- '$\angle A$ er større end $\angle B$' er ensbetydende med, at $\angle B$ kan placeres inde i $\angle A$,

- '$\angle A$ er mindre end $\angle B$' er ensbetydende med, at $\angle A$ kan placeres inde i $\angle B$,

- '$\angle A$ er en stump vinkel' er ensbetydende med, at der kan placeres en ret vinkel inde i $\angle A$,

- '$\angle A$ er en spids vinkel' er ensbetydende med, at $\angle A$ kan placeres inde i en ret vinkel.

Opgave 10
Bredden af hver løbebane er 12 m : 10 = 1,2 m. Det er kun i de halvcirklede former i hver ende af banen, at turen bliver længere ved at komme i et ydre løbespor. Ved at flytte 1,2 m ud bliver radius i de halvcirklede former 1,2 m længere. Den samlede forlængelse af en omgang bliver derfor $2 \cdot \pi \cdot 1{,}2\,\text{m}$, idet omkredsen af en cirkel er $2 \cdot \pi \cdot r$, her 7,54 m. Så banerne skal forskydes 7,54 m i forhold til hinanden.

Øvelse 12
1) Ved at tegne jordens cirkel færdig skønnes, at jorden har dobbelt så stor diameter som månen. Månens diameter bliver herefter halvdelen af jordens eller 20.000 stadier ifølge Eratosthenes tal.

2) Idet månen er $\frac{1}{2}$ grad, skal der 720 måner til at nå hele vejen rundt i månebanen, som derfor har en længde på 720 · 20.000 stadier.

3) Vi kan nu ud fra formlen "omkredsen = $2 \cdot \pi \cdot r$" beregne radius i månebanen som
$r = \frac{\text{omk}}{2 \cdot \pi} = \frac{720 \cdot 20.000}{2 \cdot \pi} =$ ca. 2.292.000 stadier.

Øvelse 12 fortsat

4) Afstanden til solen fandt han således til $20 \cdot \dfrac{720 \cdot 20.000}{2 \cdot \pi}$ = stadier. Da der går 720 sole til at dække hele solbanen om jorden, får solen en udstrækning/diameter på $20 \cdot \dfrac{720 \cdot 20.000}{2 \cdot \pi} \cdot \dfrac{2 \cdot \pi}{720}$, der forkortes til $20 \cdot 20.000 = 400.000$ stadier.

Undersøgelse 4

Her giver vi kun hint til undersøgelsens del 2), 3) og 5):

2) Omkredsen af Reuleaux-trekanten findes ved at betragte omkredsen som en halvcirkel med radius 10, idet hver af cirkelbuerne i Reuleaux-trekanten svarer til en vinkel på $60°$.

For at finde arealet betragtes Reuleaux-trekanten som bestående af en indre ligesidet trekant ABC plus tre cirkelafsnit svarende til $60°$. Vi har behandlet, hvordan man finder arealet af disse former i kapitel 2.

3) Hvis 5-kanten tegnes og omkredsen måles nøjagtig vil man opdage, at den har samme omkreds som cirklen og Reuleaux-trekanten. Det har noget at gøre med, at vinkelsummen i en stjerne altid er $180°$ jf. undersøgelse 1 i dette kapitel.

5) Hvis du har konstrueret en Reuleaux-firkant, er du den første matematiker, der har gjort det! Før vi siger tillykke, foreslår vi at tjekke resultatet … Faktisk har andre bevist, at det er umuligt at gøre det.

16. Problemløsnings- og ræsonnementskompetence

Opgave 4
Da gitterpunkterne på trekantpapir fire og fire sidder sammen i romber, og hele papiret faktisk kan ses som en projektion af kvadratisk papir og altså også et sømbræt, må der gælde akkurat samme formel som i Picks teorem bortset eventuelt fra en faktor, der skyldes forskelligt areal af den grundlæggende enhed. Men hvis vi siger, at den grundlæggende rombe i trekantpapiret får areal lig 1, så gælder Picks teorem uændret også for trekantpapir.

I det hele taget er Picks teorem invariant over for parallelprojektion, så der vil – bortset fra en konstant faktor, der har at gøre med arealenheden – gælde et Picks teorem på alt papir, der bygges op af kongruente parallelogrammer.

Øvelse 3
Det er $A_T = ½R_T + I_T - 1$, som vi søger at bevise.

Kaldes antallet af søm som elastikken mellem P og Q rører ved for c, så kan vi bestemme antallet af randsøm på vores trekant som $a + b + c + 1$. Vi har altså $R_T = a + b + c + 1$. Da vi i rektanglet havde $R = 2a + 2b$, og vi nu har $2R_T = 2a + 2b + 2c + 2$, får vi $R = 2R_T - 2c - 2$.

Antallet af de indre søm I i rektanglet er (det dobbelte af de indre søm i trekanten T)+ c, altså
$I = 2I_T + c$.

Endelig har vi, at arealet af trekanten er det halve af rektanglets areal $A_T = ½A$ eller $A = 2A_T$.

Vi sætter nu de tre formler, vi har fundet, $R = 2R_T - 2c - 2$, $I = 2I_T + c$ og $A = 2A_T$ ind i Picks formel for rektanglet, som vi jo har bevist i sætning 1: $A = ½R + I - 1$.

Nu håber vi blot, at Picks formel for trekanten kommer ud i den anden ende. Vi prøver:

$A = ½R + I - 1$ bliver ved de tre indsættelser ensbetydende med

$2A_T = ½(2R_T - 2c - 2) + (2I_T + c) - 1$ som er ensbetydende med

$2A_T = (R_T - c - 1) + (2I_T + c) - 1$ som er ensbetydende med

$2A_T = R_T + 2I_T - 2$ eller

$A_T = ½R_T + I_T - 1$, hvilket skulle bevises.

Opgave 7
Tegn selv en vilkårlig skæv trekant T på et sømbræt og omskriv den tæt med et akseparallelt rektangel R, hvilket i den generelle situation betyder, at R kommer til at bestå af T samt tre akseparallelle retvinklede trekanter 1, 2 og 3.

Picks teorem for rektanglet har vi bevist, og det lyder $A = ½R + I – 1$.

Men vi har også vist Picks teorem for de tre akseparallelle trekanter 1, 2 og 3, og de hedder med indlysende notation:

$A_1 = ½R_1 + I_1 – 1$

$A_2 = ½R_2 + I_2 – 1$

$A_3 = ½R_3 + I_3 – 1$

Vi ønsker så på snedig vis herfra at argumentere for den formel, vi er ved at bevise:

$A_T = ½R_T + I_T – 1$

Ved betragtning af vores figur ser vi, at hele arealet A består af de fire trekanters areal, altså at

$A = A_1 + A_2 + A_3 + A_T$ eller $A_T = A – (A_1 + A_2 + A_3)$.

Erstatter vi nu med højresiderne I alle vores formler får vi:

$A_T = A – (A_1 + A_2 + A_3) = (½R + I – 1) – (½R_1 + I_1 – 1 + ½R_2 + I_2 – 1 + ½R_3 + I_3 – 1) =$

$½(R – R_1 – R_2 – R_3) + (I – I_1 – I_2 – I_3) + 2$

Vi har altså ved ren talmanipulation uden megen geometrisk omtanke nu opnået at vise

(1) $A_T = ½(R – R_1 – R_2 – R_3) + (I – I_1 – I_2 – I_3) + 2$

Vi må så håbe, at vi ved nogle geometriske overvejelser kan nå frem til, at højresiden bliver $½R_T + I_T – 1$.

Her bliver der ligesom i øvelse 3 nogle mellemregninger, der skal holde øje med de indre søm i rektanglet, som kommer til at blive randsøm på trekant T's sider. Lad os sige, at der er hhv. c_1, c_2 og c_3 sådanne indre rektangelsøm, som ligger på de indre sider af hhv. trekant 1, 2 og 3 (hvilke tre sider tilsammen også bliver T's sider).

Så bliver det samlede antal indre søm I i rektanglet lig summen af de indre søm i de fire trekanter, der udgør rektanglet plus $c_1 + c_2 + c_3$, formelt opskrevet $I = I_T + I_1 + I_2 + I_3 + c_1 + c_2 + c_3$, eller mere interessant for os:

$I – I_1 – I_2 – I_3 = I_T + c_1 + c_2 + c_3$, der også kan skrives $I – I_1 – I_2 – I_3 = I_T + R_T – 3$, idet $R_T = c_1 + c_2 + c_3 + 3$. Vi har altså:

Opgave 7 fortsat

(2) $I - I_1 - I_2 - I_3 = I_T + R_T - 3$

Det ligger lige for at skulle indsætte (2) i (1), men først vil vi lige gøre det samme med randsømmene. Hvis vi lægger alle randsømmene i de tre trekanter 1, 2 og 3 sammen, får vi langt flere randsøm end i rektanglet, idet vi får sømmene $c_1 + c_2 + c_3$ med. Hvis vi fintæller, hvor tit vi får hjørnesømmene i trekant T med, får vi faktisk

$R_1 + R_2 + R_3 = R + c_1 + c_2 + c_3 + 3 = R + R_T$, idet $R_T = c_1 + c_2 + c_3 + 3$.

Dette betyder, at

(3) $(R - R_1 - R_2 - R_3) = - R_T$.

Nu er alt klart til, at vi kan indsætte (2) og (3) i (1), hvilket giver:

$A_T = - \frac{1}{2}R_T + I_T + R_T - 3 + 2$, hvilket er ensbetydende med

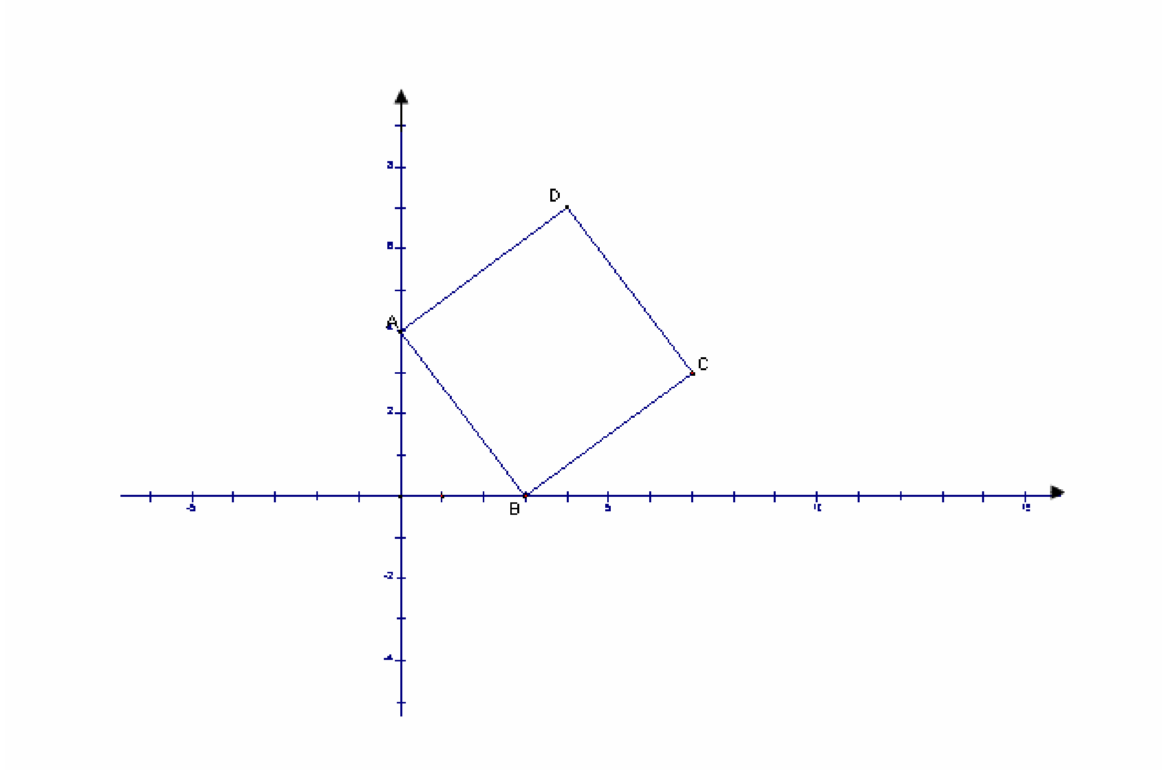
$A_T = \frac{1}{2} R_T + I_T - 1$, hvilket skulle vises.

Øvelse 7
L + L = L, L + U = U, U + U = L, U + L = U

L · L = L, L · U = L, U · U = U, U · L = L.

17. Indledning til analytisk geometri

Øvelse 3

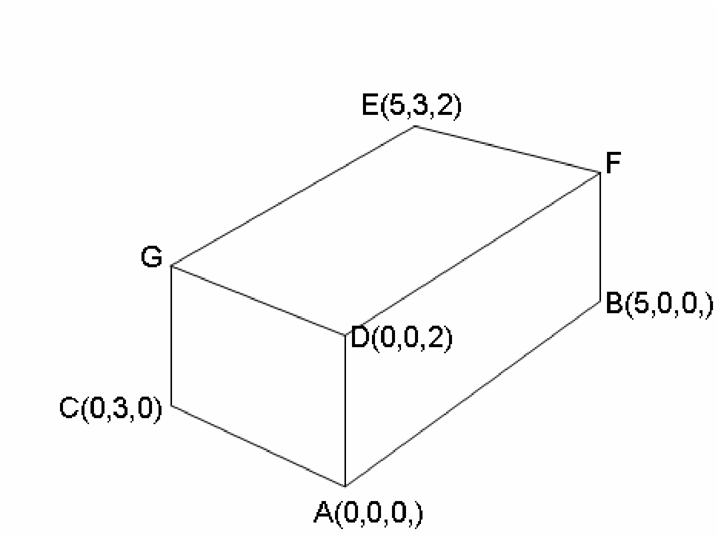


D = (4,7).

Sidelængden = $\sqrt{(0-4)^2+(3-0)^2}=5$.

Diagonallængden = $\sqrt{(4-3)^2+(0-7)^2}=\sqrt{50}=5\sqrt{2}$.

Øvelse 7



De punkter, der ligger længst fra hinanden er diagonalt modsatte punkter, fx *A* og *E*.

F har koordinaterne (5,0,2)

G har koordinaterne (0,3,2)

Punktet, som ikke kan ses på tegningen, har koordinaterne (5,3,0)

2) Kassens rumfang er $(5 \cdot 3 \cdot 2)\ \mathrm{m}^3 = 30\ \mathrm{m}^3$.

3) Her anvendes midtpunktsformlen fra øvelse 5. Koordinaterne til kassens midtpunkt *M* bliver således: $M = \left(\frac{5+0}{2}, \frac{3+0}{2}, \frac{2+0}{2}\right) = \left(\frac{5}{2}, \frac{3}{2}, 1\right)$

Øvelse 9

For at finde koordinatudtrykket for vektoren $\overrightarrow{AB}$ skal vi ifølge teorien øverst på side 677 parallelforskyde $\overrightarrow{AB}$, så startpunktet for $\overrightarrow{AB}$ falder i (0,0). Dette kan man gøre ved at trække x_1 fra førstekoordinaten og y_1 fra andenkoordinaten i de to punkter $A(x_1, y_1)$ og $B(x_2, y_2)$. Herved får det nye endepunkt B' koordinaterne $(x_2 - x_1, y_2 - y_1)$. Det betyder, at $\overrightarrow{AB}$ har koordinaterne $(x_2 - x_1, y_2 - y_1)$.

Øvelse 12
1) Der er en fejl i opgaven, kun én tur bringer os hjem, og det er tur A. Hvis i tur C den sidste vektor rettes til (-400,-400) så bringer den os også hjem.

2) Hvis man adderer alle *x*-koordinaterne og får 0 som resultat samt alle *y*-koordinaterne og ligeledes får 0 som resultatet, så fører turen tilbage til udgangspunktet. Det er, fordi vektorer adderes ved at addere *x*-koordinaterne for sig og *y*-koordinaterne for sig. Fx
$(3,6)+(-7,4)=(3+(-7),6+4)=(-4,10)$

Øvelse 14
Efter 20 minutter er han nået $\frac{1}{3}$ ud af vektoren (3,4), dvs. nået ud til begyndelsespunktet plus $\frac{1}{3}$ af vektoren (3,4): $\left(\frac{1}{3}\cdot 3,\frac{1}{3}\cdot 4\right)=\left(1,\frac{4}{3}\right)$, altså til $(-5,-2)+\left(1,\frac{4}{3}\right)=\left(-4,-\frac{2}{3}\right)$.

Efter en halv time er han kommet til begyndelsespunktet plus det halve af vektoren (3,4): $\left(\frac{1}{2}\cdot 3,\frac{1}{2}\cdot 4\right)=\left(\frac{3}{2},2\right)$, altså til $(-5,-2)+\left(\frac{3}{2},2\right)=\left(-\frac{7}{2},0\right)$.

Efter to timer er han kommet dobbelt så langt, altså begyndelsespunktet plus 2 gange vektoren (3,4): $(2\cdot 3,2\cdot 4)=(6,8)$, altså til $(-5,-2)+(6,8)=(1,6)$

Generelt kan vi se, at til tiden *t* er han nået til begyndelsespunktet plus vektoren (3*t*, 4*t*), således at svar nr. 2 er det korrekte.

Øvelse 19
Man kan som antydet overføre hele tankegangen fra beviset for sætning 1. Men vi vælger at vise en alternativ måde, hvor vi i stedet for udnytter resultatet i sætning 1, der siger, at den rette linjes ligning hedder $y = ax + b$ for passende valg af a og b. Indsættes heri de to givne punkter fås:

(1) $y_1 = ax_1 + b$
(2) $y_2 = ax_2 + b$

Trækkes (1) fra (2) fås $y_2 = a(x_2 - x_1)$, hvoraf findes det ønskede resultat $a = \frac{y_2 - y_1}{x_2 - x_1}$.

Øvelse 21
1) Vi tegner en cirkel med centrum i (0,0) og radius 5. Det ses let, at vi her har følgende otte punkter på cirkelperiferien: (5,0); (3,4); (0,5); (-3,4); (-5,0); (-3,-4); (0,-5); (3,-4).

Cirklen skulle imidlertid have centrum i (-3,-2), derfor forskydes cirklen med centrum i (0,0) med vektoren (-3,-2). Herved får vi de otte punkter på cirkelperiferien koordinaterne:
(2,-2); (0,2); (-3,3); (-6,2); (-8,-2); (-6,-6); (-3,-7); (0,-6).

2) Mængden af løsninger til $(x-3)^2+(y-2)^2<25$ kan karakteriseres som mængden af punkter (x,y), der befinder sig inden for cirkelperiferien i denne cirkel.

Øvelse 22
1) Afstanden er lig $\sqrt{(12-0)^2+(3-0)^2+(4-0)^2}=13$.

2) Afstanden er lig $\sqrt{\left(12+\sqrt{2}-\sqrt{2}\right)^2+\left(3+\sqrt{3}-\sqrt{3}\right)^2+\left(4+\sqrt{5}-\sqrt{5}\right)^2}=13$.

3) Afstanden er lig $\sqrt{(4-1)^2+(5-2)^2+(6-3)^2}=3\sqrt{3}\approx 5,2$.

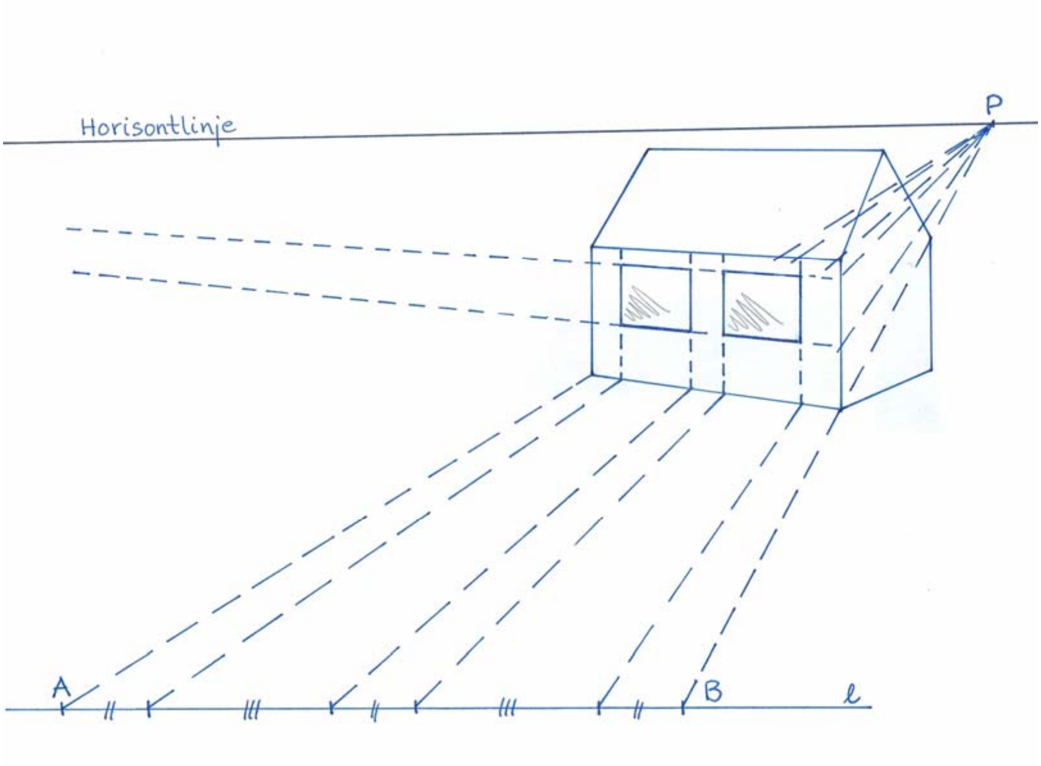
4) Afstanden er lig $\sqrt{(3,4-(-12,4))^2+(14,7-(-4,5))^2+(-3,7-17,2)^2}=\sqrt{1055,09}\approx 32,5$.

Opgave 6
Vi definerer en kugleflade med centrum i A og radius R som mængden af punkter, der har afstanden R til A. Et punkt P (x,y,z) ligger på kuglefladen, hvis afstanden AP er lig med R.

Udnyttes afstandsformlen i sætning 4, ser vi, at betingelsen bliver $\sqrt{(x-a)^2+(y-b)^2+(z-c)^2}=R$ eller ved kvadrering på begge sider af lighedstegnet lig med $(x-a)^2+(y-b)^2+(z-c)^2=R^2$, hvilket skulle vises.

18. Tegning af tredimensionale objekter

Øvelse 8



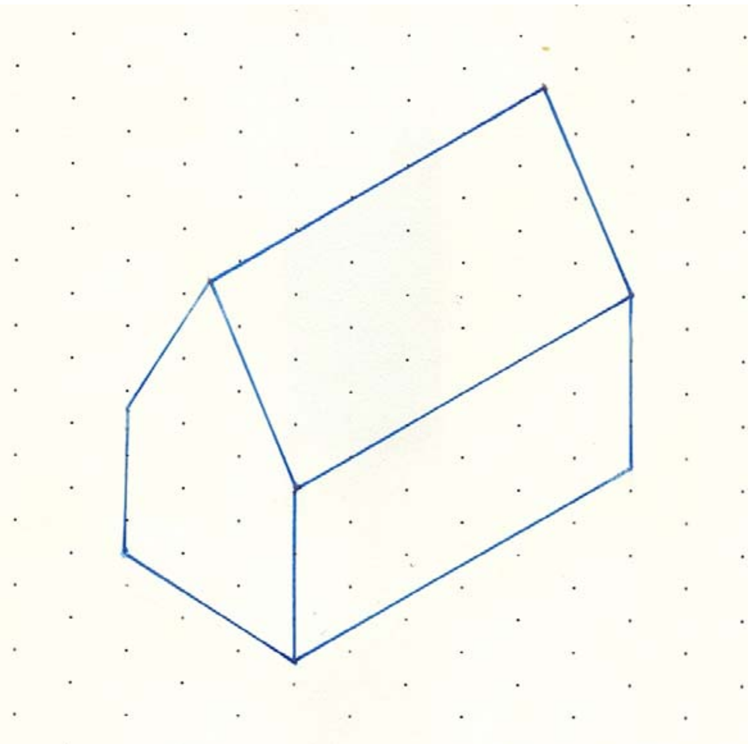
Huset skitseres så godt man kan ud fra de retningslinjer, vi har givet indtil nu, hvorefter selve øvelsen går ud på at anvende ‘anden teknik’ s. 707 til at tegne vinduer i facaden.

Først tegner man en vandret linje *l* nederst på tegneplanen, derefter vælger man et tilfældigt punkt *P* på horisontlinjen og der tegnes linjer gennem de nederste, forreste hushjørner til *l*, hvor skærings-punkterne kaldes *A* og *B*.

Størrelse og antal af vinduer vælges frit og målene afsættes i forhold til længden *AB*. Den samlede længde af vinduerne subtraheres fra *AB*, og differensen fordeles ligeligt til mellemrummene. Længden af vinduer og mellemrum afsættes på *l*. Herfra tegnes linjer til *P*, hvor disse skærer husets facade (ved jorden) tegnes lodrette linjer. Nu er vinduerne fordelt i en perspektivisk rigtig afstand. Højden af vinduerne fastlægges, og der tegnes linjer til venstre forsvindingspunkt på horisontlinjen.

Opgave 5

Den isometriske tegning:

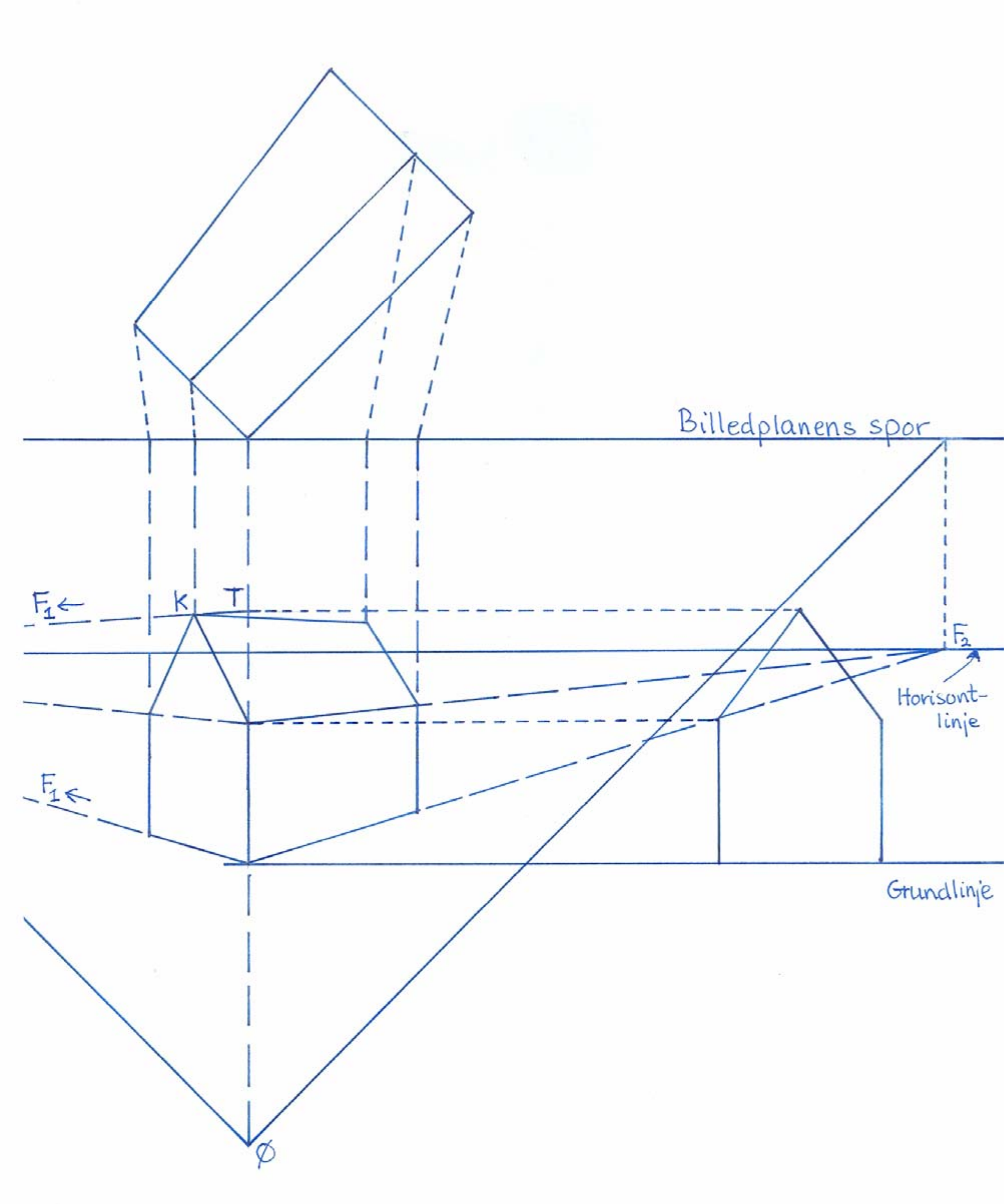


Perspektivtegningen:

Man begynder på samme måde som i *Første fase* s. 710. Dog har vi undladt at tegne synslinjerne helt ned til øjet, men stoppet ved billedplanens spor for at undgå for mange linjer. Se tegningen på næste side.

Herefter fortsætter vi som i *Anden fase* og vælger grundlinje og horisontlinje. Vi har tegnet husets gavl ind, for at kunne føre højderne ind på tegningen (vi ved ikke noget om husets og tagets højde, så det har vi selv valgt). Når kippen, (punktet K) den øverste spids på gavlen, skal fastlægges, tegnes en linje fra tagets top ind til den lodrette linje ud fra øjet (punktet T), så ved vi, hvor højt taget skulle have været, hvis det lå med kippen op mod billedplanen. Men da det ikke er tilfældet, skal man tegne en linje til forsvindingspunktet F_1 til venstre, og hvor denne skærer den lodrette linje fra billedplanens spor, der er ført ned fra synslinjen til tagrygningen (markeret med en linje på grundplanen af huset), har vi endelig fået fastlagt kippen (K).

Opgave 5 fortsat



19. Statistik

Opgave 8

Hvis segregationsraten er $\frac{1}{2}$, så er frekvensen, eller det relative antal, kun halvt så stort i den givne gruppe, som den er i helheden. Hvorimod den er dobbelt så stor, hvis segregationsraten er 2. Da sådanne frekvenser er positive rationale tal, kan forholdet mellem disse to tal blive til enhver positiv brøk.

Segregationsraten $\frac{1}{2}$ i Hørsholm-eksemplet kan generelt udtrykkes som $\frac{p}{q}$, og vi ser på, hvordan denne fremkommer i en kommune med 6.000 indbyggere. Hvis p ud af de 6.000 indbyggere i kommunen er efterkommere, og antallet af efterkommere blandt Danmarks 6.000.000 indbyggere er $1000 \cdot q$, udregnes segregationsraten til:

$$\frac{\frac{p}{6000}}{\frac{1000 \cdot q}{6.000.000}} = \frac{\frac{p}{6000}}{\frac{q}{6.000}} = \frac{p}{q}$$

Da segregationsraten har at gøre med frekvenser, dvs. brøkdele, skulle man tro, at man nemt kunne udregne segregationsraten for frafald i erhvervsuddannelser (eud) for efterkommere af indvandrere. Men problemet er, at der mangler den tilsvarende frekvens for hele Danmarks befolkning. For 'danskere' i denne statistik omfatter ikke indvandrere. Derfor kan segregationsraten ikke udregnes. Hvis der kun var ganske få efterkommere i Danmark, ville segregationsraten være cirka $\frac{58}{32,2}$. Hvis derimod næsten hele Danmarks befolkning var efterkommere, så ville frafaldet for indvandrere (58%) ligge meget tæt på frafaldet på landsplan. Derfor ville segregationsraten være tæt på $\frac{58}{58} = 1$. Der ville altså ikke være en segregation over for indvandrere. Det ville der derimod være over for danskere, idet segregationsraten for frafald i eud for danskere i så fald ville være cirka $\frac{32,2}{58}$ eller godt 0,5.

Heraf lærer vi bl.a., at segregationsraten aldrig vil afvige ret meget fra 1 for den store majoritet i et samfund, og det er da også netop minoriteternes særlige situation, segregationsraten er opfundet til at beskrive.

Opgave 10

1) Det er vel først og fremmest, fordi det nu er sådan, indekset er defineret. Og hvorfor er det så defineret på netop den måde? Antag, at man i stedet for numerisk afvigelse havde regnet med absolut afvigelse. Så ville summen af afvigelserne blive 0. Hvilket kommer af, at hvis der er for mange på en skole, må der tilsvarende blive for få på en anden, så alle afvigelserne vil ophæve hinanden.

Opgave 10 fortsat
2) Jo, man kunne vel godt have brugt summen uden at halvere den.

Der er imidlertid en uskreven norm for noget der kaldes et indeks, at det kunne være rart at det lå mellem 0 og 1 eller m.a.o. mellem 0% og 100% – sommetider, når indeks kan være negativt, prøver man at få det til at ligge mellem -1 og 1.

Så lad os se på en situation med total forskelsbehandling, total segregation: fx en kommune med to skoler med 50 lærere på hver, hvor alle de mandlige lærere er samlet på skole 1 og alle kvindelige på skole 2. I den situation bliver skemaet:

Skole	Mandlige lærere	Lærere i alt	Andels af kommunens mandlige lærere på skolen	Andel af alle kommunens lærere på skolen	Difference
1	50	50	$\frac{50}{50}=1$	$\frac{50}{100}=\frac{1}{2}$	$\frac{1}{2}$
2	0	50	$\frac{0}{50}=0$	$\frac{50}{100}=\frac{1}{2}$	$\frac{1}{2}$

Lægger vi differencerne sammen fås $\frac{1}{2}+\frac{1}{2}=1$, og det ville egentlig være et godt svar for total segregation. Så – på nuværende stade i vores undersøgelse – forekommer det lidt gådefuldt, hvorfor Gorard foreslår, at der skal deles med 2.

Vi må nok prøve med skoler af ulige størrelse, fx en kæmpeskole med 1000 lærere, der alle er kvinder, og så en lille skole på 10 lærere, der alle er mænd.

Så ville skemaet se således ud:

Dobbeltskæv kommune

Skole	Mandlige lærere	Lærere i alt	Andels af kommunens mandlige lærere på skolen	Andel af alle kommunens lærere på skolen	Difference
1	0	1000	$\frac{0}{1000}=0$	$\frac{1000}{1010}=0,9901$	0,99
2	10	10	$\frac{10}{10}=1$	$\frac{10}{1010}=0,01$	0,99

Hvis vi blot lagde differenserne sammen, ville vi i dette eksempel få et segregationsindeks på 1,98, og vi ville faktisk kunne få det op på 2, hvis vi gjorde endnu større forskel på skolestørrelserne. Så ved at dele med 2 opnår man sikkert, at indekset ligger mellem 0 og 1, hvilket kunne være meget passende for et sådant indeks. Vi har ikke bevist, at det er således, men vi er kommet på sporet af det.

Ansats til et bevis

Skole	Mandlige lærere (m)	Lærere i alt (L)	Andels af kommunens mandlige lærere på skolen	Andel af alle kommunens lærere på skolen	Difference
1	x	y	$\frac{x}{m}$	$\frac{y}{L}$	$\left\lvert \frac{x}{m} - \frac{y}{L} \right\rvert$
2	$m-x$	$L-y$	$\frac{m-x}{m} = 1 - \frac{x}{m}$	$\frac{L-y}{L} = 1 - \frac{y}{L}$	$\left\lvert \frac{x}{m} - \frac{y}{L} \right\rvert$

Vi ser her, at der i det simple tilfælde, hvor der kun er to rækker i tabellen (to skoler), vil komme samme slutresultat (den numeriske difference) i slutkolonnen. Når vi derfor dividere resultatet med 2, bliver segregationsindekset $|\frac{x}{m} - \frac{y}{L}|$. Da de to brøker, der indgår, er frekvenser, ligger de mellem 0 og 1, og derfor må deres differens også ligge mellem 0 og 1. Hvis indekset er lig med 0, så betyder det, at skole 1 har samme brøkdel af kommunens mandlige lærere $\frac{x}{m}$ som den i det hele taget har andel af kommunens lærere. Og så følger det ret let, at det samme må være tilfældet for den anden skole. Vi ser altså, at indekset gør det, vi forventer af et sådant indeks i det tilfælde, at det er lig med 0.

Men hvornår er det så lig med eller tæt på 1? Det kan det kun være, hvis den ene af de to brøker, der indgår i $\frac{x}{m} - \frac{y}{L}$, er lig med 0, og den anden er lig med 1 eller tæt på. Den ekstreme situation, hvor $\frac{y}{L} = 0$ eller 1, er ikke mulig, idet alle lærerne da vil være på den ene skole. For at få indekset tæt på 1, skal vi åbenbart have næsten alle lærere på den ene skole og meget få på den anden. Lad os sige, at $\frac{y}{L} = 0$ på den første skole. For at få indeks tæt på 1 skal vi samtidig have $\frac{x}{m}$ tæt på 1, og det skal være på den skole, hvor der er y lærere, altså meget få lærere. Vi skal med andre ord have en ren mandsdomineret 'lille skole'. Det eneste mulige eksempel er altså det, som vi ovenfor kaldte en 'dobbeltskæv kommune'. Det vil sige, at dette indeks faktisk er en smule mere anvanceret og fintmålende, end man lige umiddelbart skulle tro.

3) Besvarelsen her må være subjektiv. Men vi kan da lige overveje, om det er rimeligt, at segrationsindekset kun er $\frac{1}{2}$ i vores første eksempel med to skoler på i alt 100 lærere, mens det er tæt på 1 i eksemplet med de to skoler med 1010 lærere. Dette er ikke noget matematisk spørgsmål, men et spørgsmål om man ville opleve større forskelsbehandling rettet mod mændene i det sidste tilfælde end i det første. For hvis indekset er vokset til det dobbelte, burde det afspejle en reel ændring til det værre i mændenes oplevede segregations-situation.

Opgave 13
Ifølge vores definition ovenfor betyder odds 2 det samme som 2:1. Ud af tre tilfælde vil den betragtede hændelse således indtræffe i 2 af tilfældene, men ikke i det tredje (i det lange løb altså, noget vi skal se mere på i sandsynlighedsregning i næste kapitel).

Så odds 1 betyder 1:1, altså at den betragtede hændelse vil indtræffe i 1 af 2 tilfælde. Det er den situation, der også omtales 'fifty-fifty', 50% chance for og 50% imod.

Hvis odds for, at jeg klarer det næste maratonløb i København, er $\frac{1}{2}$, betyder det $\frac{1}{2}$:1 eller 1:2, idet man altid kan forlænge odds. Jeg vil altså klare det i 1 ud af 3 tilfælde, hvordan man så ellers når frem til det resultat på forhånd.

Opgave 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.

Opgave 16
Hvis vi af det foregående har fået den vane at udregne odds ratio af en tabel som denne:

a	*b*
c	*d*

ved regneudtrykket $\dfrac{\frac{a}{c}}{\frac{b}{d}}$

så gør det ikke noget, at man kommer til at spejlvende tabellen til:

a	*c*
b	*d*

med tilhørende regneudtryk $\dfrac{\frac{a}{b}}{\frac{c}{d}}$

For begge regneudtryk giver: $\dfrac{ad}{bc}$

Det giver en slags krydsprodukt, idet man ganger de faktorer, der står i modsatte hjørner. Krydsproduktforholdet kommer af, at man sætter de to produkter i forhold til hinanden, som i eksemplet $ad : bc$. Husk dog, at man stadig kan få det reciprokke udtryk, hvis man alene bytter om på rækkerne eller på kolonnerne. Derfor skal man lige holde i hovedet, hvad tendensen er i undersøgelsen, så man ikke på grund af regneteknik kommer til at argumentere for det modsatte af de faktiske forhold.

Opgave 20
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 multiplikation 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 3
Omkreds $= 4 \cdot \text{sidelængde}$, altså $f(x) = 4 \cdot x$ udtrykt som abstrakt funktion. Så i dette tilfælde vil observationssættet have en tendenslinje (ret linje), der matcher præcist til punkterne.

$\text{Sidelængde} = \sqrt{\text{arealet}}$, altså $f(x) = \sqrt{x} = x^{\frac{1}{2}}$ udtrykt som abstrakt funktion. Så også i dette tilfælde vil observationssættet have en tendenslinje (en parabelarm), der matcher præcist til punkterne.

20. Sandsynligheder og sandsynlighedsregning

Øvelse 1
Man ser på, om personen ville indgå et væddemål, der gav 80 kroner tilbage på en indsats på 1 krone, hvis Rifbjerg faktisk fik Nobelprisen inden for de næste fire år, men var uvillig til det, hvis han kun fik 79 kroner.

Hvis der er tale om en rationelt tænkende person (hvilket altid er udgangspunktet i sådanne teorier), så betyder det, at personen mener, at det lige akkurat kan løbe rundt med en gevinst på 80 kroner. Altså må vedkommende være rede til at tabe 79 spil og vinde kun 1 ud af 80 spil, idet den samlede udgift så ville blive $80 \cdot 1$ kroner – der betales også for det vundne spil – og den samlede indtægt ville blive $1 \cdot 80$ kroner.

Personen ville altså tillægge hændelsen “Rifbjerg får Nobelprisen inden for de næste 4 år” sandsynligheden $\frac{1}{80}$.

Øvelse 4
1) Kombinatorisk set er der to farver {R, S}, hvilket vil sige fire ordnede farvekombinationer i to kortspil {(R, S), (R, R), (S, R), (S, S)}. Kun en af disse kombinationer giver begge røde, nemlig udfaldet (R,R). Sandsynligheden er altså $\frac{1}{4}$.

2) kombinatorisk set er der fire kulører {Ruder, Klør, Spar, Hjerter}, hvilket vil sige 16 ordnede farvekombinationer i to kortspil {(R, R), (R, K), (R, S), (R, K), (K, R), (K, K) …(H, H)}. Kun en af disse kombinationer giver begge klør, (K, K). Sandsynligheden er altså $\frac{1}{16}$.

3) I forlængelse af 1) ses, at dette sker i to ud af de fire muligheder (R,R) og (S,S). Sandsynligheden er altså $\frac{2}{4} = \frac{1}{2}$.

4) I forlængelse af 2) ses, at dette sker i fire ud af 16 muligheder: (R, R), (K, K), (S, S) og (H, H). Sandsynligheden er altså $\frac{4}{16} = \frac{1}{4}$.

5) I forlængelse af 3) ses, at de har forskellig farve i de øvrige tilfælde, 2 ud af 4. Sandsynligheden er altså igen $\frac{2}{4} = \frac{1}{2}$.

6) I forlængelse af 4) ses, at dette sker i 12 ud af 16 muligheder. Sandsynligheden er altså $\frac{12}{16} = \frac{3}{4}$.

Øvelse 5
Vi tillader os allerede her i opgavebesvarelsen at bruge den præcise sprogbrug, som indføres i Ypsilon på side 777: Udfaldsrum og $P(H)$ som forkortelsen for 'sandsynligheden for hændelsen H'.

En dansk mønt viser Plat og Krone: P og K.

Vi opfatter kastet med tre mønter som et ordnet sæt, hvor PPK er forskelligt fra KPP, idet vi antager, at vi så får en symmetrisk situation, hvor alle udfald har samme sandsynlighed:

Udfaldsrummet bliver så: {KKK, KKP, KPK, KPP, PKK, PKP, PPK, PPP}

Da der er 8 udfald, får de hver sandsynligheden $\frac{1}{8}$, og det bliver nemt af svare på de stillede spørgsmål ved simpel optælling.

P(3 krone) = P(K,K,K) = $\frac{1}{8}$

P(netop 1 krone) = P({KPP, PKP, PPK}) = $\frac{3}{8}$

P(netop 2 krone) = P({KKP, KPK, PKK}) = $\frac{3}{8}$

P(ingen krone) = P(alle plat) = P(PPP) = $\frac{1}{8}$

Øvelse 6
Dette er en øvelse i at læse og forstå definitionen af et sandsynlighedsfelt (U,P). Vi skal blot tjekke, at kravene 1., 2. og 3. til et sandsynlighedsfelt er opfyldt for kast med en 'matematisk' terning.

1) Da sandsynligheden for en hændelse H findes ved at optælle antal udfald i H, hvilket jo må ligge mellem 0 og 6, og så dividere dette tal med 6. Svaret må så være en brøk mellem 0 og 1 som krævet.

2) $P(U)$=1, for hændelsen U består af alle seks udfald, hvorfor $P(U) = \frac{6}{6} = 1$. Tilsvarende består Ø slet ikke af nogen udfald, så $P(Ø) = \frac{0}{6} = 0$.

Øvelse 6 fortsat

3) $P(A \cup B) = P(A) + P(B)$.

Hvis A og B er disjunkte delmængder af udfaldsrummet $\{1,2,3,4,5,6\}$, så bliver $P(A) = \dfrac{\text{antal elementer i } A}{6}$ og $P(B) = \dfrac{\text{antal elementer i } B}{6}$.

Men da antallet af elementer i $A \cup B$ er lig $(\text{antal elementer i } A) + (\text{antal elementer i } B)$, så fås:

$$P(A \cup B) = \frac{\text{antallet af elementer i } A \cup B}{6} = \frac{\text{antallet af elementer i } A + \text{antallet af elementer i } B}{6} =$$
$$\frac{\text{antallet af elementer i } A}{6} + \frac{\text{antallet af elementer i } B}{6} = P(A) + P(B).$$

Man kunne måske også sige, at 3) er indlysende da sandsynlighed bestemmes ved optælling og addition, men vi har altså valgt at skrive en detaljeret besvarelse.

Opgave 12

$X(i, j) = i \cdot j$.

Hvis X skal give 3, må enten i eller j være lig med 3, da 3 er et primtal, så den del af udfaldrummet på 36 elementer, som giver $X = 3$, er $\{(1,3), (3,1)\}$. Så $P(X = 3) = \dfrac{2}{36}$.

$X = 4$ kan derimod optræde på lidt flere måder, da 4 kan faktoriseres som $2 \cdot 2$. Mulighederne er her $\{(1,4), (2,2), (4,1)\}$, så $P(X = 4) = \dfrac{3}{36}$.

Tal med flere faktorer optræder med større sandsynlighed. Således er $P(X = 12) = \dfrac{4}{36}$, fordi de brugbare udfald her er $\{(2,6), (3,4), (4,3), (6,2)\}$.

Den samlede sandsynlighedsfordeling for den stokastiske variabel $X(i, j) = i \cdot j$
bliver:

x	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Vi kontrollerer lige, at summen giver 36 i nederste række, og det gør den.

Opgave 14

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.

Opgave 16
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 12
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 16
Ved fire kast med en mønt kan vi bruge figur 3 over fire kast med en terning til at holde styr på tankerne og beregningerne, idet vi erstatter S med K og I med P. Men vi skal huske, at sandsynligheden nu er $\frac{1}{2}$ på hver eneste gren.

$P(0 \text{ Krone}) = P(\text{fire Plat}) = P(\text{PPPP}) = P(\text{at følge grenen til højre}) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16}$.

$P(1 \text{ Krone}) = P(\text{at følge en af de fire grene der indeholder netop 1 Krone}) = 4\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{4}{16} = \frac{1}{4}$.

$P(2 \text{ Krone}) = P(\text{at følge en af de seks grene der indeholder netop 2 krone}) = 6\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{6}{16} = \frac{3}{8}$.

$P(3 \text{ Krone}) = P(\text{at følge en af de fire grene der indeholder netop 3 Krone}) = 4\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{4}{16} = \frac{1}{4}$.

$P(\text{fire Krone}) = P(\text{KKKK}) = P(\text{at følge grenen til venstre}) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16}$.

Hvis man bliver usikker i en sådan beregning, så kan man altid vende tilbage til metoden med at skrive hele udfaldsrummet op med alle 16 udfald: KKKK, KKKP, KKPK, KKPP, …, PPPK, PPPP.

Øvelse 21
Situationen dækkes af binomialfordelingen med $n = 8$, og $p = 0{,}10$.

$P(\text{netop et produkt er med fejl}) = \text{K}(8,1) \cdot 0{,}10^1 \cdot 0{,}90^7 = 8 \cdot 0{,}10^1 \cdot 0{,}90^7$.

$P(\text{netop to produkter er med fejl}) = \text{K}(8,2) \cdot 0{,}10^2 \cdot 0{,}90^6 = 28 \cdot 0{,}10^2 \cdot 0{,}90^6$.

Øvelse 21
Situationen dækkes af binomialfordelingen. I modelleringssprog vil man sige, at den dækkes af binomialmodellen.

$P(120 \le \text{antal K} \le 220) = P(\text{antal K} \le 220) - P(\text{antal K} \le 119)$

$P(\text{antal K} \le 220)$ bestemmes ved 'tabelopslag' i et regneark. I Excel indsætter vi blot binomialfordelingen i et felt Binomialfordeling(220;400;0,5;1) og får svaret 0,979885.

Tilsvarende findes $P(\text{antal K} \le 119)$ = Binomialfordeling(119;400;0,5;1) = 1,62E – 16 = 0.
Sandsynligheden for at havne i det ønskede interval er altså ca. 98%.

Opgave 18

1) P(fem ens) = 6 $\cdot$ P(fem enere) = $6\cdot\left(\frac{1}{6}\right)^5=\left(\frac{1}{6}\right)^4\approx 0{,}0008$.

2) P(fem forskellige) = P(første viser hvad som helst) $\cdot$ P(den anden viser en af de andre 5 udfald) $\cdot$ P(den tredje viser et af de andre 4 udfald) $\cdot$ P(... 3 udfald) $\cdot$ P(... 2 udfald)$=1\cdot\frac{5}{6}\cdot\frac{4}{6}\cdot\frac{3}{6}\cdot\frac{2}{6}\approx 0{,}09$.

3) P(mindst et par) kan måske opfattes på flere måder, men vi opfatter det som det modsatte af, at alle er forskellige, fordi det faktisk kan tælle som et par i Yatzy, og vi får derfor ved hjælp af formlen for modsat hændelse: P(mindst et par) = 1 – P(alle er forskellige) = 1 – 0,09 = 0,91.

4) Når man kaster en gang er sandsynligheden for ikke at få et par 1 – 0,91= 0,09, som vi også så i delopgave 2). Sandsynligheden for ikke at få et par tre gange i træk er ifølge multiplikations-princippet $0{,}09^3 = 0{,}00073$.

Den modsatte hændelse, hvor man får et par mindst en gang har så sandsynligheder 1 – 0,00073 = 0,9993 ≈ 99%, hvilket er tæt på en sikker hændelse.

5) Dette tilfælde dækkes af binomialmodellen P(netop 2 seksere) = $\mathrm{K}(5{,}2)\cdot\left(\frac{1}{6}\right)^2\cdot\left(\frac{5}{6}\right)^3=$ $10\cdot\left(\frac{1}{6}\right)^2\cdot\left(\frac{5}{6}\right)^3\approx 0{,}16$.

6) Sandsynligheden for fire ens må være 6 gange så stor som sandsynligheden for fire seksere, vi bruger igen binomialmodellen.

P(fire ens) = 6 $\cdot P$(fire seksere) $=6\cdot\mathrm{K}(5,4)\cdot\left(\frac{1}{6}\right)^4\cdot\left(\frac{5}{6}\right)^1=6\cdot 5\cdot\left(\frac{1}{6}\right)^4\cdot\left(\frac{5}{6}\right)^1=30\cdot\left(\frac{1}{6}\right)^4\cdot\left(\frac{5}{6}\right)^1\approx 0{,}019$.

7) P(tre ens) = 6 $\cdot P$(tre seksere) = 6 $\cdot$ K(5,3) $\cdot\left(\frac{1}{6}\right)^3\cdot\left(\frac{5}{6}\right)^2=3\cdot 10\cdot\left(\frac{1}{6}\right)^3\cdot\left(\frac{5}{6}\right)^2\approx 0{,}096$.

8) Når opgaven her er formuleret lidt blødere, skyldes det, at det slet ikke er så nemt at sikre sig, at man regner rigtigt og/eller bruger de rigtige principper. For at være sikre vælger vi at arbejde i hele det ‘store’ udfaldsrum for fem terninger, hvor der er i alt $6\cdot 6\cdot 6\cdot 6\cdot 6=6^5$ udfald.

Men lad os først beslutte, at vi vil have et ægte fuldt hus og altså ikke vil medregne 5 ens. Et fuldt hus ser altså således ud (A,B,B,A,B), hvor A og B er forskellig øjenvisninger, A er dem, der er to ens af, og B dem der er tre ens af. Men det er klart, at øjentallet A ikke behøver at falde netop på første- og fjerdepladsen.

Så vi kan bruge denne indledende sondering af opgaven til at opdele konstruktionen af et fuldt hus i tre delprocesser:

Opgave 18 fortsat

1A. Først vælger vi på hvilke to pladser, de to ens skal være – dvs. A's placering.

1B. Så vælger vi de pladser, B skal være på, men her er faktisk ikke noget valg, for det skal være de resterede tre pladser.

2. Så vælger vi det øjental, som A skal vise.

3. Endelig vælger vi det øjental, som B skal vise.

Bemærk, at disse valg (med bemærkningen ved 1B indregnet) er uafhængige af hinanden. Så vi kan optælle på hvor mange måder, vi kan konstruere et fuldt hus ved at benytte kombinatorikkens multiplikationsprincip (som vi arbejder mere med i ε-bogen).

1A. Kan gøres på K(5,2) måder = 10 måder.

2. Kan gøres på 6 måder.

3. Kan gøres på 5 måder (for vi accepterer ikke 5 ens).

Det samlede antal måder, vi kan konstruere et fuldt hus på er altså $10 \cdot 6 \cdot 5 = 300$, og sandsynligheden for et fuldt hus bliver så 300 divideret med antallet af udfald i det 'store' udfaldsrum, som var 6^5, dvs. $\frac{300}{6^5} \approx 0{,}039$.

Hvis man prøver at udregne sandsynligheden for to par på lignende måde, skal man passe på med, at de to par A og B kan skifte plads i et kast som (A,A,B,B,C), gør man ikke det, får man et resultat, som er dobbelt så stort, som praktisk erfaring i Yatzy viser, noget der ikke sker i delspørgsmål 8. Det siger sig selv, at hvis praksis konsekvent afviger fra teorien, så må man se, om man ikke er kommet til at lave nogle forkerte beregningsforudsætninger i teorien. Tilsvarende kan man bruge praksis eller simuleringer af Yatsy til at kontrollere, om de teoretiske beregninger ovenfor virker korrekte. Men det er nok lettest at underkaste dem en logisk analyse og se, om der er argumenteret godt for de principper, der bygges på undervejs.

21. Stikprøver, Estimationsteori

Opgave 3
Opgaven er ret frit stillet, men her er en vinkel, man kunne tage med i en besvarelse.

Hvis man vil gøre sig et skøn over, hvor andelen af røde kugler som oftest falder, så kunne man prøve at lokalisere en klump med 80% af sandsynligheden i midten af fordelingen. Vi kunne altså ønske at finde grænserne for de 10% stikprøver med færrest røde og de 10% stikprøver med flest røde.

De 10 % færreste må findes ud fra en tabelstump af binomialfordelingen b(100; 0,18). Vi kan ved at se på figur 2 i bogen skønne os frem til, hvilke dele af tabellen vi behøver at tage med. Tabellen laves som beskrevet i afsnittet om binomialfordelingen s. 794-95 fx i Excel ved hjælp af den indbyggede binomialfordelingsfunktion. Vi viser en stump af tabellen her:

	b(100; 0,18)
k	$P(X \leq k)$
7	0,001
8	0,004
9	0,009
10	0,020
11	0,039
12	0,071
13	0,118
14	0,182
.	.
.	.
.	.
21	0,820
22	0,878
23	0,920
24	0,950
25	0,971
26	0,983
27	0,991
28	0,995

Af tabellen aflæser vi, at $P(X \leq 13) = 0{,}118$, hvilket vil sige, at ca. 12% af udfaldene har færre end 14 røde.

I den anden ende kunne vi vælge $P(X \leq 23) = 0{,}920$, hvilket vil sige, at 8 % af udfaldene har mere end 23 røde. Da 8% + 12% = 20% har vi hermed fået afgrænset en ret sandsynlig klump i midten i intervallet [14;23], hvor vi vil havne i 80% af stikprøverne.

Opgave 3 fortsat

Omsat til den faktiske undersøgelse af danske elever, hvor man mener, at 18% i 9. klasse er dårlige læsere. Så betyder det, at man i en skole med et hundrede 9. klasseselever ofte (80% af tilfældene) vil finde, at der på en sådan skole er mellem 14 og 23 dårlige læsere, hvis vel at mærke vi kan have tillid til undersøgelsen med de 18%, og hvis skolen ikke afviger fra landsresultatet. Omvendt kan man næppe sige, at skolen afviger fra situationen på landsplan, blot fordi der er 25 dårlige læsere blandt skolens 9. klasseselever. For vores valg af en klump på 80% i midten var jo et tilfældigt skud, og det er faktisk ret sandsynligt, at en skole falder udenfor, fordi vi netop har lavet det så ca. 20% falder udenfor.

Opgave 4

Vi skal finde grænserne, hvor vi kan skære $2\frac{1}{2}\%$ af i hver side af binomialfordelingen b(1000; 0,359). Vi ser på den kumulerede fordeling nederst på s. 809 og ser, at den nedre grænse må ligge omkring 330, mens den øvre (ved $97\frac{1}{2}\%$) må ligge omkring 390. Dvs. at ud fra dette skøn ligger Vilstrups resultat på 35,9% i 95%-konfidensintervallet.

For at få en mere præcis vurdering laves en tabel over b(1000;35,9) på regneark – evt. kan man begrænse tabellen til intervallet [325;393], da vi allerede ved, at det er her omkring, vi skal lede.

k	$P(X = k)$	$P(X \leq k)$
325	0,0021	0,0132
326	0,0024	0,0156
327	0,0028	0,0184
328	0,0032	0,0217
329	0,0037	0,0254
330	0,0042	0,0296
331	0,0048	0,0343
332	0,0054	0,0397
.	.	.
.	.	.
.	.	.
386	0,0054	0,9646
387	0,0048	0,9694
388	0,0042	0,9736
389	0,0037	0,9773
390	0,0033	0,9806
391	0,0029	0,9835
392	0,0025	0,9860
393	0,0022	0,9882

Af tabellen fremgår det, at sandsynligheden for, at højst 328 personer stemmer på Socialdemokratiet, er 0,0217, mens 329 personer bringer os op på 0,0254. Altså tilhører 329 konfidensintervallet.

Ligeledes fremgår det, at sandsynligheden for, at højst 389 personer stemmer på Socialdemokratiet, er 0,9773, mens 390 bringer os op på 0,9806. Altså afslutter 389 konfidensintervallet.

Vi ser nu, at 95%-konfidensintervallet går fra og med 329 til og med 389. Det virkelige valgresultat blev på 34,9%, eller det viste sig, at 349 personer ud af 1000 stemte på Scialdemokratiet. Dette resultat blev fint fanget af undersøgelsens 95%-konfidensinterval.

Dog må man understrege, at Vilstrup og andre bør angive deres valgresultater med de nødvendige forbehold, fx 95%-konfidensinterval. I så fald ville avisoverskrifterne nok ikke være så bastante.

Øvelse 7
Man må her forestille sig, at man ikke har adgang til at lave den præcise udregning baseret på binomialfordelingen i et regneark og blot ønsker sig et hurtigt overblik ud fra tommelfingereglen på s. 810.

Vi regner med, at AIM spurgte 1000 vælgere, hvoraf 324 svarede ja til, at de ville stemme på Socialdemokratiet, svarende til $p = \frac{x}{n} = \frac{324}{1000} = 0,324$.

Benytter vi nu formlen for 95% konfidensinterval får vi:

$\left[0,324 - 1,96 \cdot \sqrt{\frac{0,324 \cdot (1 - 0,324)}{1000}};\ 0,324 + 1,96 \cdot \sqrt{\frac{0,324 \cdot (1 - 0,324)}{1000}} \right]$, der ved udregning giver [0,295; 0,353].

Det ser igen ud til, at meningsmålerne her fangede det korrekte resultat 0,349 i deres konfidens-interval.

Nu kunne man sætte spørgsmålstegn ved, om en sådan tommelfingerregel giver et pålideligt svar. Men det gør den – i al fald i dette tilfælde – som det fremgår af sammenligningen med tabeludskrift fra den relevante binomialfordelingstabel b(1000;0,324):

k	*P(X = k)*	*P(X ≤ k)*
291	0,0022	0,0135
292	0,0026	0,0160
293	0,0030	0,0190
294	0,0034	0,0224
295	0,0039	0,0264
296	0,0045	0,0308
297	0,0051	0,0359
298	0,0058	0,0417
.	.	.
.	.	.
.	.	.
350	0,0058	0,9627
351	0,0051	0,9678
352	0,0045	0,9723
353	0,0040	0,9763
354	0,0035	0,9798
355	0,0030	0,9828
356	0,0026	0,9854
357	0,0023	0,9877

Øvelse 8
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 fordobles. Dette gælder helt generelt og derfor også i det konkrete taleksempel i opgaven.

Øvelse 10
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.

22. Stokastikkens didaktik

Ingen løsningsforslag.