

1. årgang
nr. 1 / 2002

KEMI MATEMATIK FYSIK *i*

PERSPEKTIV

Statistik - viden eller tilfældighed

Statistikker laves ofte for at kunne opdage afvigelser fra normale forhold. Når man støder på hændelser som for eksempel tilsyneladende usædvanligt mange dødsfald på et plejehjem eller mange varme vintre, kan sandsynlighedsregningen hjælpe til at finde ud af, om der kan være tale om tilfældige udsving. Selvom sådanne beregninger ikke giver egentlige beviser, er de ofte grundlag for meget vigtige beslutninger og konklusioner. Et specielt aspekt ved statistiske metoder er, at et talmateriale ofte kan behandles korrekt på flere måder med varierende konklusioner som resultat.

DNA-test - hvor sikker er metoden ?

En klassisk DNA-test går kort fortalt ud på at bestemme længderne af udvalgte DNA-sekvenser på kromosomerne.

Typisk bestemmes 8 sådanne fragmentlængder, og disse 8 tal udgør en persons DNA-profil. Hver af fragmentlængderne varierer i et vist omfang fra person til person, og DNA-profilen giver således en karakteristik af personen.

Ved opklaringen af mord og voldtægtssager sammenlignes DNA-profilen fra et spor, fx en blodplet, en sædplet eller et hår fra gerningsstedet, med den DNA-profil, man finder hos eventuelle mistænkte.

Der er imidlertid en vis måleusikkerhed ved bestemmelsen af en fragmentlængde, idet to målinger af samme DNA-sekvens ikke altid giver samme resultat. For at kunne bruge metoden i kriminalsager er man derfor nødt til at vælge, hvor meget to fragmentlængder må afvige fra hinanden, for at man vil sige, at der er overensstemmelse mellem fragmentlængderne. Valget af denne maksimale afvigelse, som vi vil kalde acceptgrænsen, har betydning for sandsynligheden for at lave hver af de to typer fejlkonklusioner, som man kan komme til at lave:

- **at frikende den skyldige**
- **at anklage en uskyldig**

En relativ lav acceptgrænse vil give en relativ stor sandsynlighed for at frikende den skyldige, men en relativ lille sandsynlighed for at anklage en uskyldig. Omvendt med en relativ høj acceptgrænse.

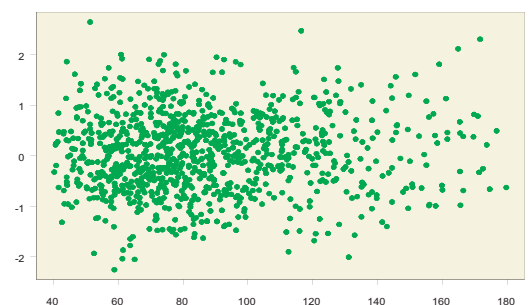
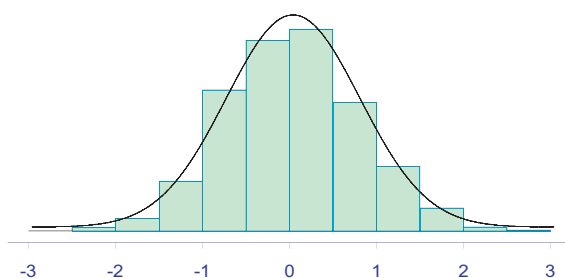
Da metoden skulle tages i brug i Danmark i begyndelsen af 90'erne, blev der indsamlet to store

statistiske materialer, som blev brugt til at beregne sandsynlighederne for at lave fejlkonklusioner. Det ene skulle belyse måleusikkerheden og det andet hvor meget fragmentlængderne varierede fra person til person.

For at vurdere måleusikkerheden blev 1.000 DNA-sekvenser hver målt to gange. Fragmentlængderne blev angivet i mm og kunne ligge mellem 40 mm og 180 mm. Differenserne mellem to sammenhørende fragmentlængder fra samme person havde en middelværdi på 0 og en spredning på 0,77.

Normalfordelingen med en middelværdi på 0 og en spredning på 0,77 giver en god beskrivelse af fordelingen af differenserne, hvilket ses ved at sammenligne normalfordelingskurven med et histogram over differenserne. Desuden var der ikke tegn på fx større måleusikkerhed på lange fragmenter. Dette har den praktiske betydning, at det er muligt at vælge en acceptgrænse, der er uafhængig af fragmentlængden.

Man kan nu beregne konsekvensen af at vælge en acceptgrænse på fx 3 mm. At spor og mistænkt har samme DNA-profil betyder så, at alle 8 par af sammenhørende fragmentlængder ikke afviger mere end 3 mm. Ved brug af normalfordelingen får man en sandsynlighed på 0,9999 for overensstemmelse mellem to sammenhørende fragmentlængder fra samme person. En skyldig udelukkes, hvis mindst én af otte fragmentlængder er forskellig fra sporets tilsvarende fragmentlængde, og sandsynligheden for dette er lig med $1 - 0,9999^8$, som er cirka 0,0008. Det vil sige, at 8 ud af 10.000 skyldige ville blive frikendt ved en acceptgrænse på 3 mm, hvis beslutningen udelukkende var baseret på DNA-testen.



For hvert par af målinger (L_1, L_2) af fragmentlængder svarende til samme DNA-sekvens afsættes punktet $((L_1 + L_2)/2, L_1 - L_2)$

Ved at tage hensyn til, hvor meget fragmentlængderne varierer fra person til person, kan man vise, at sandsynligheden for at en uskyldig har samme DNA-profil som sporet, er mindre end 0,00001, eller sagt på en anden måde: ud af 100.000 tilfældigt valgte personer vil der højst være én, der har samme DNA-profil som sporet. Når man skal vurdere størrelsen af denne sandsynlighed, kan man i hvert fald sige, at hvis man tænkte, at DNA-profiler for samtlige voksne danske mænd blev sammenlignet med sporet, ville man godt kunne finde adskillige der passede; men hvis man kun leder blandt en gruppe mistænkte, fx dem der bor eller har været i nærheden af gerningsstedet, vil det være meget usandsynligt at finde andre end den skyldige med samme DNA-profil som sporet.

I forbindelse med de første retssager, hvor en DNA-test indgik som bevismateriale, var der stor uenighed, både i og uden for retssalen, om, hvorvidt man skulle lægge vægt på denne test. Modstanden skyldtes naturligvis især risikoen for at komme til at dømme en uskyldig. Det blev dog hurtigt ret almindeligt at bruge DNA-analyserne, når man samtidig havde andre indicier. Der var desuden en række sager, hvor DNA-

profiler kunne udelukke en række mistænkte og her var der ikke så stor betænkelighed ved at benytte testene. Medvirkende til at omvende en del skeptikere har sikkert også været, at man ved at undersøge gamle sager ved hjælp af DNA-analyser opdagede flere tilfælde, hvor den dømte sandsynligvis var uskyldig, og både i Danmark og udlandet genoptages nu op til 30 år gamle sager.

Første gang, en DNA-test indgik som det eneste tekniske bevis, var i forbindelse med en sag i 1998, hvor den 10-årige Susan var blevet voldtaget og myrdet. En 51-årig mand, der boede i den ejendom, hvor liget blev fundet, havde i første omgang nægtet at lade sig teste. Det viste sig imidlertid at han havde samme DNA-profil som spor på liget, og han blev dømt.

Der sker hele tiden tekniske forbedringer af DNA-analyserne, således at sandsynligheden for at dømme en uskyldig bliver stadig mindre, og det diskuteres stort set ikke længere om man skal bruge DNA-analyser, men derimod om man skal opbygge et register med så mange DNA-profiler som muligt. •



Tegning: Lisbeth Andersen

DNA-analyser blev første gang anvendt, da to engelske teenage-piger var blevet voldtaget og myrdet i den engelske landsby Enderby i 1983 og 1986. Spor på ofrene havde samme DNA-profil. For at finde gerningsmanden blev der taget DNA-prøver på 3.000 mænd fra egnen, men ingen af prøverne svarede til sporene. Ved en tilfældighed opdagede man imidlertid, at to af de afgivne prøver var identiske, og det viste sig, at en mand havde afgivet to prøver; den ene gang var han mødt op i stedet for en ven, der ikke havde tid til at møde op. Da man endelig fik bestemt denne vens DNA-profil passede den med sporene.

Massemord - eller moderne

Danske plejehjem er sjældent i mediernes søgelys, og plejehjemsassistenter kan som regel færdes i fred for nysgerrige fotografer. Men i oktober 1997 blev der vendt op og ned på hverdagen for plejehjemmet Plejebo på Vesterbro i København, og i særdeleshed for plejehjemsassistent D, der havde ansvaret for en af Plejebos afdelinger. Hun blev anholdt og sigtet for 22 tilfælde af drab på Plejebos patienter. Aviserne fortalte om århundredets drabssag, og deres spisesedler skrev med katastrofetyper om Dødens engel og D for Død.

Sagen var imidlertid ikke så oplagt, for D var ikke pågrebet med en kanyler i hånden og en plejehjmsgang fuld af lig. De mange dødsfald havde fundet sted over en toårig periode, og ingen af dødsfaldene virkede i sig selv mistækelige. Der var ingen beviser, der direkte forbandt D med dødsfaldene. Der var håndgribelige tegn på mindre uregelmæssigheder på Plejebo: penge, der var forsvundet fra af-dødes bankkonti, medicin, der blev tilordnet uden lægetilsyn. Men argumentet bag drabsanklagen var af statistisk natur: I løbet af D's ansættelse (1994-96) var der 44 dødsfald på hendes afdeling, mod 16 dødsfald på naboafdelingen. Begge afdelinger havde 15 værelser, og en gennemgang af Plejebos journaler viste at beboerne på de to afdelinger havde sammenlignelige helbredsforhold.

Denne forskel i dødelighed lyder umiddelbart voldsom. På den anden side er det nemt for forsvaret at så tvivl om lødigheden af argumentet. Lad os forestille os en række plejehjem, der i alt væsentligt er ens, og hvor personalet ikke slår beboerne ihjel. Ren og skær tilfældighed vil betyde, at der er en vis variation i antallet af dødsfald på disse plejehjem, og derfor er der et plejehjem, der tilsyneladende har en højere dødelighed end de andre. Dette plejehjem er ikke farligt at bo på, næste år har det sikkert færre dødsfald end gennemsnittet. Men lige netop i den betragtede periode kommer det til at fremstå som et uhyggeligt sted.

Et simulationseksperiment

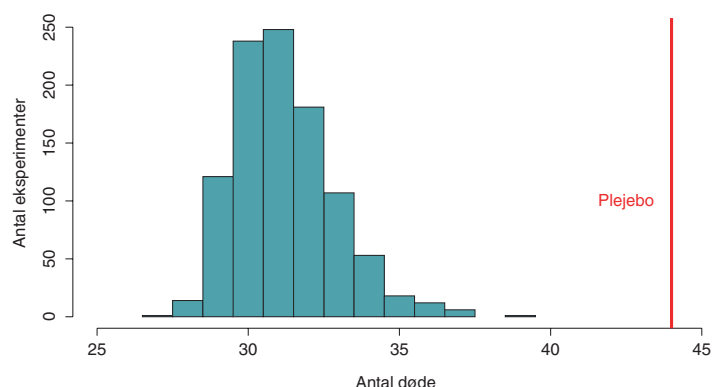
Hvis forsvarets argument om tilfældighed skal tilbagevises, må man vide noget om hvor store udsving i dødelighed, der kan forekomme af sig selv. Det vil vi undersøge ved simulation.

I Danmark er der ca. 30.000 plejehjemspladser. Lad os for argumentet skyld lade som om disse plejehjemspladser er fordelt på 2.000 lige store afdelinger af samme størrelse som D's afdeling på Plejebo. Lad os endvidere forestille os, at der på alle disse afdelinger er en vis konstant sandsynlighed p for at der sker et dødsfald på en given dag. Forudsætter man,

at der højst sker et dødsfald per dag, vil det forventede antal døde i løbet af en periode på $n = 730$ dage være $n \cdot p$. Hvis man går ud fra at naboafdelingen med 16 dødsfald svarer til det normale, vil man estimere p til at være $16/730$.

I et eksperiment har vi simuleret antallet af døde på hvert af 2.000 plejehjem i løbet af $n = 730$ dage, med $p = 16/730$. Herefter har vi udvalgt plejehjemmet med flest døde. Resultatet af 1.000 gentagelser af

FIGUR 1



Opsummering af 1.000 simulationseksperimenter. I hvert eksperiment følges 2.000 plejehjemsafdelinger i $n = 730$ dage. Hvert plejehjem oplever et dødsfald på en given dag med sandsynlighed $p = 16/730$. Resultatet af eksperimentet er antallet af døde på det plejehjem, hvor der er sket flest dødsfald.

dette eksperiment er optegnet i figur 1.

Resultatet varierer fra eksperiment til eksperiment, men kommer aldrig i nærheden af 44. Et enkelt eksperiment resulterede i 39 døde på det hårdst ramte plejehjem, men i langt størstedelen af eksperimenterne har det hårdst ramte plejehjem haft mellem 29 og 33 dødsfald.

Fortolkningen af dette eksperiment er, at under den givne model for dødsfald på plejehjem, så ville der i hele Danmark ikke ved en tilfældighed opstå en eneste plejehjemsafdeling, der er så hårdt ramt som D's afdeling, om vi så ventede 1.000 perioder af to år! Enten må der være et eller andet rivende galt på Plejebo - eller også er modellen bag eksperimentet urealistisk.

Kritik af eksperimentet

Forsvaret har adskillige muligheder for at angribe simulationseksperimentet. Det nemmeste at tage fat på, er det benyttede p . Dette p var et estimat, baseret på en antagelse om at naboafdelingen var typisk. I virkeligheden er det sikkert den store forskel på de to afdelinger, der er sprunget myndighederne i øjnene, snarere end det er antallet af dødsfald på D's afdeling. I så fald er sandheden om naboetagen nærmere, at der er usædvanligt få dødsfald. I så fald har vi benyttet et for lille p i simuleringseksperimentet.

plejehjemsdrift

Vi kan komme denne kritik i møde ved at ændre lidt på eksperimentet. Vi forestiller os de 2.000 plejehjemsafdelinger knyttet sammen to og to i virtuelle plejehjem. Nu opsøger vi det plejehjem, hvor der observeres den største forskel på de to afdelinger. Som resultat af eksperimentet afrapporterer vi denne maksimale forskel.

Dette eksperiment er gennemført 1.000 gange for $p = 27/730$, en sandsynlighed, der gør begge Plejebo-

har gjort mere end at lufte vores manglende evne til at tage et standpunkt: vi har præciseret at nøglen til svaret ligger i en grundig undersøgelse af dødsfaldsmønsteret, ikke blot på Plejebo, men på alle landets plejehjem. Og dermed har vi peget på en vej til at opnå den ønskede viden.

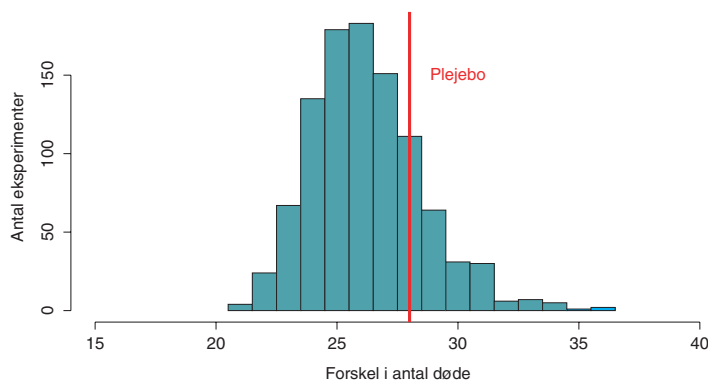
Plejebosagens forløb

Hvis vi vender tilbage til forløbet af Plejebosagen, så er det klart at statistiske overvejelser ikke i sig selv kan være tilstrækkelige til en domsfældelse. Det ville kræve håndfaste forbindelser mellem D og dødsfaldene.

I første omgang støttede politiet sig til Københavns stads-læge, der havde gennemgået journalerne for de mange dødsfald, og på den baggrund mente, at D havde ydet særdeles aktiv dødshjælp. Stads-lægens rapport blev i sagens forløb kritiseret voldsomt. Retslægerådet måtte svare på over 200 kritiske spørgsmål til rapporten (primært fra forsvaret). På baggrund af disse svar, opgav anklage-myndigheden at føre sagen ved retten, og D stod tilbage som frikendt.

Den eneste retlige konsekvens af århundredets drabssag blev, at den privatpraktiserende læge, der var knyttet til Plejebo, fik en bøde på nogle få tusinde kroner for ikke at have overholdt sin tilsynspligt. •

FIGUR 2



Opsummering af 1.000 simulationseksperimenter. I hvert eksperiment følges 1.000 plejehjem med 2 afdelinger i $n = 730$ dage. Hvert plejehjem oplever et dødsfald på en given dag med sandsynlighed $p = 27/730$. Resultatet af eksperimentet er forskellen på antallet af døde på de to afdelinger, for det plejehjem hvor denne forskel er størst.

afdelinger temmelig usædvanlige - D's afdeling dog mere end naboafdelingen. Resultaterne er afbildet på figur 2.

Nu fremstår Plejebosituationen slet ikke som urimelig - faktisk er der i mere end 25 procent af eksperimenterne en større forskel på afdelingerne for det værst ramte plejehjem, end der er på Plejebo. Fortolkningen af eksperimentet er at en situation som Plejebo må forventes at opstå på mindst et dansk plejehjem i løbet af fire observationsperioder af to år.

Denne konklusion er i forsvarets favør. Men eksperimentet er baseret på en værdi af p , der gør begge Plejeboafdelinger atypiske. Næste skridt må være at skaffe et bedre funderet bud på p ved at udnytte information om andre plejehjem. En sådan analyse bliver indviklet, for virkelighedens plejehjem vil ikke have samme p ! Store plejehjem vil alt andet lige opleve flere dødsfald end små plejehjem, og man må derfor finde en måde at korrigere p for størrelsen af plejehjemmene.

Så vores undersøgelse om, hvorvidt Plejebosituationen kan forklares ved tilfældighed, munder ud i den karakteristiske statistiker-konklusion: Det kan vi ikke sige noget om på det foreliggende grundlag. Men vi

I en sag i Wisconsin i USA i 1980'erne, døde tre børn i dagpleje hos dagplejemoderen Sandra Pankow i løbet af en fem-årig periode, alle med dødsårsagen vuggedød. Vuggedød er en rest-årsag, der gives til spædbørnsdødsfald, hvor man ikke kan finde en mere oplagt dødsårsag. På det tidspunkt døde ca. 2 promille af alle spædbørn i den vestlige verden med denne diagnose inden deres første fødselsdag - i dag er vuggedød næsten forsvundet som begreb.

Dagplejemoderen blev anklaget for at have myrdet børnene, og en vigtig del af anklagerens argumentation var et regnestykke, hvor alle amerikanske spædbørn blev delt ud på virtuelle dagplejemødre med hver 20 børn, det samme antal børn som Sandra Pankow havde haft. Konklusionen på regnestykket var, at man i hele USA ville se en dagplejemor med tre vuggedøde spædbørn en gang i løbet af 600 år.

I Pankow-sagen var der udover det statistiske også visse medicinske indicier. Det sidste dødsfald blev obduceret, og ligene af de to tidligere tilfælde blev gravet op og undersøgt.

Sarah Pankow blev dømt for at have kvalt de tre børn.

Misvisende sammenligninger

Vi forestiller os, at en eller anden hårdnakket påstår, at føddernes størrelse har indflydelse på unges uddannelsesforløb. For at tilbagevise dette åbenlyse vrøvl spørger vi alle elever i en matematikerklasse og alle elever i en sproglig klasse om deres skonummer, og finder at matematikerne som helhed faktisk har større fødder end de sproglige.

For at sikre at resultatet ikke bare skyldes en tilfældighed, tager vi 9 andre matematikerklasser og sammenligner med 9 andre sproglige klasser, og hver gang finder vi, at matematikerne faktisk har større fødder.

Dette forsøg tyder åbenbart på en sammenhæng.

En anden måde at undersøge, om der er en sammenhæng mellem skonummer og uddannelsesforløb, vil være at sammenligne skonummer for sproglige drenge og matematikerdrenge hver for sig. Derefter gøre det samme for pigerne, og som forventet finde, at der ikke er nogen forskel. Derefter kan vi så konkludere, at skonummer ikke spiller nogen rolle for (denne del af) de unges uddannelsesforløb.

Den ovenfor beskrevne situation er forholdsvis let gennemskuelig. Alle ved, at køn i høj grad er bestemmende for uddannelsesvalg. Alle ved, at køn og skonummer hænger sammen, men ingen tror, at passende doser væksthormon til den danske ungdom vil afhjælpe den svigtende tilgang til de naturvidenskabelige uddannelser, selv om det ville medføre, at de unge fik større fødder. Skonummer har nemlig i sig selv ingen virkning på uddannelsesvalg.

Når det alligevel kommer til at se sådan ud i undersøgelsen ovenfor, er det fordi drenge og piger ikke er jævnt fordelt mellem sproglig og matematisk linie. Mens der er nogenlunde lige mange drenge og piger i matematikerklasserne, så er kun en fjerdedel af eleverne i de sproglige klasser drenge. Der er altså en overvægt af piger i de sproglige klasser, og derfor har de sproglige også i gennemsnit mindre fødder, når man ikke tager hensyn til køn.

Selv om skonummerets betydning for uddannelsesvalget er et spøgefuldt eksempel, så møder man ofte den slags fejlslutninger, hvor der konstateres betydningsfulde sammenhænge mellem to faktorer, og hvor sammenhængen i virkeligheden skyldes påvirkning fra uobserverede størrelser.

Rundt om på siden er vist en række eksempler på dette formuleret som fiktive avisnotitser. De strækker

sig fra det let gennemskuelige (kræft og hoftebrud) til det knapt så indlysende (alder og arbejdsløshed). Præsenteret sådan her kan de virke overdrevne. Men det er kun til man får øjnene op for, hvad man bliver udsat for i virkeligheden.

I eksemplet med hoftebrud og kræft har de to grupper, man sammenligner en helt forskellig alderssammensætning. Og da de udfald, man studerer: kræft og hoftebrud, begge er noget der oftest forekommer hos ældre mennesker, vil der være en tilsyneladende sammenhæng. Man får imidlertid ikke færre kræfttilfælde ved at forhindre hoftebrud, sådan som bandagist NN foreslår, for der er ingen årsags-sammenhæng mellem de to betragtede størrelser.

I hospitalseksemplet kan man også godt se, hvad det drejer sig om, nemlig at patienter på intensivafdelingen er dårligere end patienterne på de ortopædkirurgiske afdelinger. Så selv om intensivafdelingen var fremragende og behandlingen på ortopædafdelingen var elendig, ville man stadig have større dødelighed på den intensive afdeling. Og indlæggelse af alle patienter på ortopædkirurgisk ville næppe fremme overlevelsen i amtet. Eksemplet med de arbejdsløse er lidt sværere at gennemskue, men man kan se fx i Statistisk Årbog, at den aktuelle arbejdsløshed er højest blandt ufaglærte. Der er af gode grunde langt flere ufaglærte blandt de 16-18-årige i arbejdsstyrken end blandt de 30-35-årige. Hvis man meningsfyldt vil udtale sig om overarbejds-løshed i en bestemt aldersgruppe, må der bl.a. kontrolleres for, hvordan gruppens uddannelsesniveau er.

Det er vigtigt at inddrage relevante baggrundsvariable, hvis man vil undersøge sammenhængen mellem to variable. Spørgsmålet er så, hvordan man inddrager baggrundsvariable i sin analyse. Og hvad er relevante baggrundsvariable.

I nogen tilfælde kan almindelig sund fornuft fortælle en, at en (tænkt) baggrundsvariabel ikke kan influere undersøgelsen. Skonummer, hvis man skal undersøge uddannelsesvalg og køn, og fregner, hvis man skal sammenligne to hospitalsbehandlinger. Men i enhver undersøgelse er der mange mulige baggrundsvariable og man kan ikke altid forudse, hvad man kan se bort fra. •

Nedlægger intensiv afdeling.

Amtssygehuset i XX har besluttet at nedlægge den intensive afdeling, idet en undersøgelse har vist, at dødeligheden på denne afdeling ligger langt over hospitalets gennemsnit. I fremtiden vil alle alvorligt syge patienter blive indlagt på ortopædkirurgisk afdeling, hvor man har den højeste overlevelsesprocent. Hospitalsdirektøren tror, at den nye organisering vil blive en væsentlig sundhedsmæssig gevinst for amtets borgere.

Hoftebrud medfører forhøjet kræftrisiko

Ved en undersøgelse på Rigshospitalet har det vist sig, at patienter indlagt for brud på lårhalsknoglen har en 10 gange så høj risiko for at få kræft inden for en periode på 5 år, som patienter indlagt med henblik på en meniskoperation. Lægerne undersøger nu, hvorfor en forskellig traumelokalisering giver en så markant forskel i risikoen for at få kræft. På baggrund af undersøgelsen påpeger bandagist NN, at systematisk brug af firmaets nyudviklede underbukser med stødabsorberende indlæg til beskyttelse af hoftepartiet vil kunne reducere den alt for høje danske kræftrisiko.

De helt unge har foruroligende stor arbejdsløshed

Ved en opgørelse foretaget af arbejdsformidlingen i København har det vist sig, at arbejdsløsheden blandt de 16-18-årige er mere end dobbelt så høj som arbejdsløsheden i aldersgruppen fra 30-35. Dette karakteriseres fra alle sider som foruroligende.

Matematikere				Sproglige			
Køn	Skonummer		Pct med	Køn	Skonummer		Pct med
	≤ 40	> 40	små sko		≤ 40	> 40	små sko
Drenge	15	75	16,7%	Drenge	5	25	16,7%
Piger	50	30	62,5%	Piger	50	30	62,5%
Alle	65	105	38,2%	Alle	55	55	50,0%

Tabel 1: Sproglige og matematiske drenge har her den samme fordeling af skonummer, og det samme er også tilfældet for pigerne. Ser man bort fra køn, har de sproglige imidlertid i gennemsnit de mindste fødder, men det skyldes blot, at gruppen har relativt flest piger.

Behandling 1				Behandling 2			
Patient	Overlevelse		Pct der	Patient	Overlevelse		Pct der
tilstand	ja	nej	overlever	tilstand	ja	nej	overlever
Dårlig	20	80	20%	Dårlig	5	20	20%
God	70	30	70%	God	70	30	70%
Alle	90	110	45%	Alle	75	50	60%

Tabel 2: To behandlinger har her samme overlevelse for sammenlignelige patientgrupper, mens behandling 2 ser ud til at virke bedst, hvis ikke man tager hensyn til, hvor raske patienterne er. Hvis den ene gruppes patienter er raskere end den andens, kan to ligeværdige behandlinger se ud til at give forskellige resultater. I virkeligheden vil tal i sådanne tabeller have stokastisk usikkerhed, så ens behandlinger fx ikke giver samme overlevelseshesprocenter, men kun næsten samme overlevelseshesprocenter, men det ser vi bort fra i denne sammenhæng.

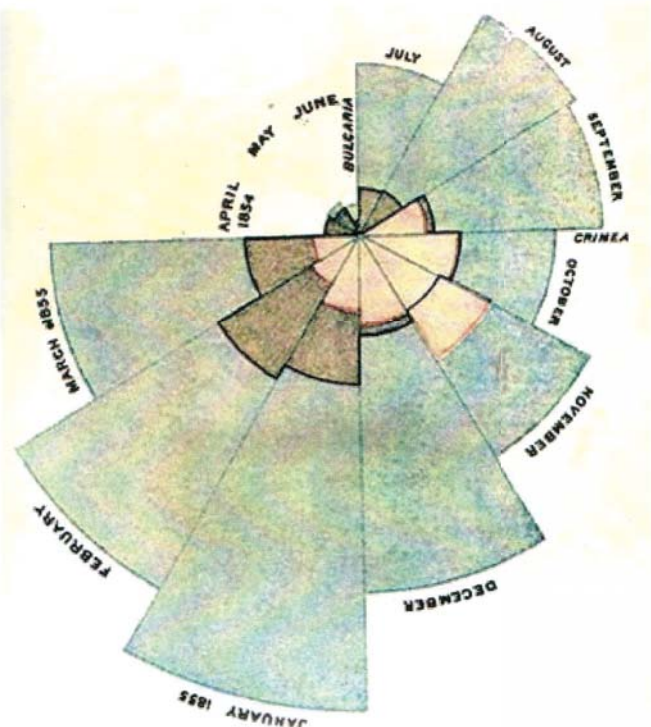
FLORENCE NIGHTINGALE

Florence Nightingale (1820-1910) er mest kendt for sit arbejde som sygeplejerske under Krimkrigen og som grundlægger af moderne sygepleje; kun de færreste véd, at hun også var en foregangskvinde inden for deskriptiv statistik, blandt andet ved at være en af de første, der benyttede grafiske metoder til at illustrere statistisk materiale.

Som ung var hun blevet interesseret i matematik, men måtte kæmpe med forældrene for at få lov til at dyrke denne interesse, da det ikke blev anset for nyttigt for en giftefærdig kvinde at kunne matematik. Heller ikke hendes ønske om at blive sygeplejerske var det, hendes velhavende forældre havde forestillet sig. På den tid var det næsten kun fattige piger, der blev sygeplejersker og disse fik ikke nogen speciel uddannelse.

Under Krimkrigen (1854-1856), hvor hun blev kendt som "the Lady with the Lamp", en mild engel, der opofrende trøstede og helbredte syge og sårede, indsamlede hun datamateriale om antal døde og dødsårsager. Hun bearbejdede dette materiale statistisk og fandt på en ny type diagram, der kunne illustrere frekvensen af forskellige dødsårsager måned for måned.

En del af hendes diagram er vist nedenfor. Det minder om det, nogle kalder et lagkagediagram. Hvert stykke af kagen repræsenterer en måned og arealet af stykket er propotionalt med antallet af dødsfald den pågældende måned. De rødlige områder svarer til de dødsfald, der skyldtes sår efter kamp, de blågrønne områder yderst svarer til dødsfald på grund af kolera eller tyfus, og resten er dødsfald af andre årsager. Florence Nightingale brugte diagrammet til at overbevise hærledelsen om, at bedre sanitære forhold ville kunne reducere antallet af dødsfald betydeligt. •



PERSPEKTIV

Udgivet af Fysikforlaget med støtte fra Undervisningsministeriets tips/lottomidler og af Birch & Krogboe Fonden

Redaktion: Niels Elbrønd Hansen
Layout: Mette Qvistorff

Produktionsgruppe: Aksel Bertelsen, Ernst Hansen, Inge Henningsen og Poul Jungersen (fagredaktør)

www.perspektiv.gymfag.dk

Tryk: Budolfi Tryk, Aalborg

Oplag: 12.000

ISSN 1602-5059



9 771602 505002