Masteroppgave

Begreper i sannsynlighet og statistikk på første trinn i videregående skole

En studie av elevers respons på oppgaver om utvalgte begreper

Av

Lars Aga

Masteroppgaven er gjennomført som et ledd i utdanningen ved Universitetet i Agder og er godkjent som sådan. Denne godkjenningen innebærer ikke at universitetet innestår for de metoder som er anvendt og de konklusjoner som er trukket.

Veiledere: Kirsten Bjørkestøl og Trygve Breiteig

Universitetet i Agder, Kristiansand

20.mai 2008

Forord

Dette er en masteroppgave i matematikkdidaktikk. Den er skrevet studieåret 2007/2008 ved Universitetet i Agder.

I dette forordet vil jeg takke alle de som har hjulpet meg i arbeidet med denne oppgaven. Først vil jeg rette en takk til mine to veiledere, Kirsten Bjørkestøl og Trygve Breiteig for godt og positivt samarbeid gjennom hele oppgaveprosessen. Jeg har hatt fått mange gode tips og ideer underveis som jeg har satt pris på.

Jeg vil også takke de tre skolene der jeg gjennomførte den diagnostiske testen. En spesiell takk til elevene og lærerne som deltok.

I tillegg vil jeg takke mine medstudenter for diskusjoner og støtte under begge mine år ved Høgskolen/Universitetet i Agder. En spesiell takk til min medstudent Magne Revheim for godt samarbeid og god hjelp gjennom alle årene vi har studert sammen.

Tilslutt vil jeg takke familien min som har støttet meg gjennom hele studieperioden min.

Kristiansand, 19.mai 2008

Lars Aga

Sammendrag

Denne masteroppgaven handler om begrepsforståelsen i emnet statistikk og sannsynlighet til elever i den videregående skolen. Det er valgt ut noen begreper som elevene er blitt testet i. Dette er gjort ved hjelp av en diagnostisk test. 141 elever fordelt på tre forkjellige skoler og åtte klasser har gjennomført testen. Alle elevene gikk på det første trinnet i den videregående skolen.

Oppgavene er diagnostiske, noen er hentet fra tidligere lignende undersøkelser, mens andre er laget selv. Testen ble gjennomført på de tre skolene høsten 2007. Denne oppgaven går ut på å analysere svarene fra denne testen.

Forskningsspørsmålene som ble stilt var:

- Hvilke misoppfatninger innenfor noen utvalgte begreper innen statistikk og sannsynlighet finnes det blant elevgrupper i den videregående skolen og hvor utbredt er disse?
- Hvilke begreper har elevene god kontroll på?
- Hva kan gjøres for å forbygge disse misoppfatningene?
- Hva kan gjøres for å bygge robuste begreper innen emnet?

Resultatene fra testen ble analysert i lys av tidligere forskning. Noen resultater fra enkeltoppgaver ble både sammenlignet med resultater fra tidligere undersøkelser og resultater fra lignende oppgaver i testen. Det ble også kort sett på forskjeller mellom forskjellige elevgrupper som gutter og jenter, og elevene som hadde valgt 1T- og 1P-modulen.

Konklusjonen ble at elevene har problemer med begreper som de store talls lov, representativitet og tilfeldighet. Spesielt de store talls lov er det begrep det virker som mange elever ikke har forstått. Kun 1 av 5 elever svarte riktig på begge oppgavene som tok for seg dette begrepet.

Andre begreper, som middelverdi og forhold, er to begreper som elevene har en god begrepsforståelse for. Spesielt guttene har en god forståelse av begrepet middelverdi, hele 9 av 10 gutter svarte riktige på oppgavene som handlet om dette.

For forebygge disse misoppfatningene blant elever blir det foreslått noen grep. Både samfunnet generelt og skolen har ansvar i dette spørsmålet. Samfunnet må lære seg å bruke begreper og hjelpemidler som grafer og diagrammer på en riktig måte, mens skolen må legge til rette for en god læring.

For å bygge robuste begreper blir det poengtert noen undervisningsmetoder som kan brukes. Blant annet kan arbeide med begrepene på en eksperimentell måte. Man presenterer en egnet oppgave og elevene gjetter svaret. Deretter tester man ut før man presenterer resultatene ved hjelp av grafer, diagrammer og lignende.

English summary

This master thesis is about students' conceptual understanding in statistics and probability. Some concepts have been chosen which the students have been tested in. The aid for this work has been a diagnostic test. The students taking this test were all at the first year of the upper secondary level.

The tasks in the test were all diagnostically. Some have been found in other similar tests, while others have been made especially for this purpose. The tests were carried through in October and November 2008. The object for this thesis is to analyze the results from this test.

The research questions were as follows:

- What misconceptions of some selected stochastic concepts are there amongst students or student groups in the upper secondary level and how widespread are they?
- What concepts are well understood?
- What can be done to change these misconceptions
- What can be done to build a sturdy foundation for concepts in stochastic?

The results from the test were analyzed in light of earlier research and theory. Some results were compared with results from other tests while other results were compared with results from similar tasks in this test. In some tasks the results were also compared between different student groups like boys and girls for instance.

The results showed that students don't fully understand the law of large numbers and chance. Students also use judgmental heuristics as representativeness. Especially misconceptions about the law of large numbers were found. Only 1 of 5 students answered correctly on the two tasks which involved this.

Other findings were that conceptions like mean and ratio were well understood. Especially the boys scored well on the tasks about mean. 9 out of 10 boys got the right answer on those two tasks.

To prevent these misconceptions some points are mentioned. The responsibility of this lies both with the school and the society in general. The society must use concepts and tools like graphs and diagrams in the correct way, while the school must arrange education in the best possible way.

In order to build a sturdy foundation for the understanding of concepts in stochastic there are mentioned some teaching methods. For instance give the students experimental tasks. First the teacher presents the task and the students guess the result. Then the students simulate the task and present the result with the assistance of graphs, diagrams or some other suitable tool.

Innholdsfortegnelse

Forord ... 2
Sammendrag ... 3
English summary ... 4
1 Innledning ... 7
1.1 Valg av tema ... 7
1.2 Problemstilling ... 7
1.3 Disposisjon ... 8
1.4 Glimt fra statistikk og sannsynlighets historie ... 8
1.5 Statistikk og sannsynlighet i samfunnet ... 10
1.6 Statistikk og sannsynlighet i skolen ... 12
2 Teori ... 16
2.1 Generelt om statistikk og sannsynlighet ... 16
2.2 Tidligere lignende undersøkelser ... 17
2.3 Matematiske begreper innen statistikk og sannsynlighet ... 18
2.3.1 Forholdsbegrepet ... 18
2.3.2 Representativitet ... 19
2.3.3 Tilfeldighet ... 20
2.3.4 De store talls lov ... 20
2.3.5 Middelverdi ... 21
2.3.6 Betinget sannsynlighet og uavhengighet ... 21
2.3.7 Konjunksjonsloven ... 23
2.3.8 Tilgjengelighet ... 23
2.4 Hvordan kan vi endre misoppfatninger? ... 24
3 Metode ... 25
3.1 Kvantitativ og kvalitativ metode ... 25
3.2 Diagnostisk test ... 25
3.3 Intervju ... 27
3.4 Skolene ... 28
3.5 Hjelpemidler ... 28
3.6 Validitet ... 29
4 Analyse ... 30
4.1 Oppgave 1 ... 30
4.1.1 Oppgave 1a ... 30
4.1.2 Oppgave 1b ... 31
4.1.3 Oppgave 1c ... 32
4.1.4 Oppsummering av oppgave 1 ... 33
4.2 Oppgave 2 ... 34
4.3 Oppgave 3 ... 36
4.4 Oppgave 4 ... 37
4.5 Oppgave 5 ... 39
4.6 Oppgave 6 ... 40
4.6.1 Oppgave 6a ... 40
4.6.2 Oppgave 6b ... 41
4.7 Oppgave 7 ... 41
4.7.1 Oppgave 7a ... 41
4.7.2 Oppgave 7b ... 42
4.8 Oppgave 8 ... 44
4.9 Oppgave 9 ... 46
4.9.1 Oppgave 9a ... 46

4.9.2 Oppgave 9b 47
4.9.3 Oppgave 9c 48
4.10 Oppgave 10 49
4.10.1 Oppgave 10a 49
4.10.2 Oppgave 10b 50
4.11 Oppgave 11 51
4.11.1 Oppgave 11a 52
4.11.2 Oppgave 11b 53
4.12 Oppgave 12 54
4.12.1 Oppgave 12a 54
4.12.2 Oppgave 12b 55
4.13 Oppgave 13 56
4.13.1 Oppgave 13a 57
4.13.2 Oppgave 13b 58
4.14 Oppgave 14 58
4.15 Sammenligning av testresultat blant ulike oppgaver 59
4.15.1 Oppgave 2 mot oppgave 3 60
4.15.2 Oppgave 4 mot oppgave 5 61
4.15.3 Oppgave 6a mot oppgave 6b 62
4.16 Sammenligning av testresultat 64
4.16.1 Mellom kjønnene 64
4.16.2 Mellom blandede, teoretiske og praktiske klasser 67
4.16.3 Mellom hvordan de selv mener de presterer i matematikkfaget 70
4.16.4 Mellom hvordan de liker matematikkfaget 71
5 Diskusjon 73
5.1 Forholdsbegrepet 73
5.2 Representativitet og tilfeldighet 74
5.3 De store talls lov 76
5.4 Middelverdi 77
5.5 Konjunksjonsloven 77
5.6 Betinget sannsynlighet 78
5.7 Forventning 78
6 Konklusjon 80
7 Pedagogiske implikasjoner og videre forskning 82
8 Refleksjoner over eget arbeid 84
Referanseliste 85
Oversikt over vedlegg 87

1 Innledning

1.1 Valg av tema

Temaet for denne oppgaven er begrepsforståelsen i emnet statistikk og sannsynlighet hos elever i den videregående skole. En av grunnene til at jeg har valgt dette temaet er at jeg alltid har hatt sannsynlighet som mitt favorittemne innen matematikkfaget. En annen grunn er at jeg jobbet med dette temaet i fjor, i en annen oppgave i samband med masterstudiet, og jeg likte veldig godt å arbeide med det.

Som nevnt skal jeg undersøke begrepsforståelsen hos elever. Jeg har valgt ut noen begreper jeg vil forske nærmere på, blant annet på bakgrunn av hva tidligere forskning har vist at elever sliter med eller begreper der tidligere forskning har vist andre interessante funn. De utvalgte begrepene vil jeg komme nærmere tilbake til i kapittel to.

Det faktum at emnet statistikk og sannsynlighet er såpass nytt innen matematikkfaget i skolen, gjør at det også er noe begrenset forskningsmateriale innen dette feltet. På det internasjonale plan er det gjort en del forskning på emnet begrepsforståelse hos elever innen statistikk og sannsynlighet tidligere. I følge Shaughnessy (1992) er størsteparten av bidragsyterne enten kognitive psykologer eller europeiske matematikk- og statistikkundervisere. Noen av de viktigste bidragsyterne han nevner er blant annet Falk, Green, Kapadia, Kahnemann og Tversky, og Konold.

I Norge har det blant annet vært noen masterstudenter som har skrevet om statistikk og sannsynlighet tidligere. Blant disse kan jeg nevne Per Sigurd Hundeland, Hege Therese Syvertsen og Wenche Rolandsen, som alle holdt til i Agder, men deres oppgaver tok for seg andre temaer innen statistikk og sannsynlighet enn de jeg tar opp. Jeg vil også nevne at masterstudenten Thomas Thorsen jobber med en oppgave om sannsynlighet ved universitetet i Oslo på samme tid som jeg arbeider med denne.

Telemarksforskning har hatt et prosjekt gående som de kaller " KIM – Kvalitet i matematikkundervisningen". En vesentlig del av prosjektet er at man bruker diagnostiske oppgaver som et middel for å kartlegge eventuelle feilforståelser av sentrale begreper hos elever. Heftene som ble laget i samband med disse testene markedsføres i dag av Utdanningsdirektoratet. Flere sentrale emner i matematikkundervisningen i grunnskolen og i den videregående opplæringen har vært testet ut, som blant annet tallregning, algebra og geometri. Sannsynlighet stod også på planen, men ble av ulike grunner ikke fullført. Oppgaven og forskningen min kan dermed bidra til å fylle ut et tomrom i forskningsmiljøet i Norge.

1.2 Problemstilling

Jeg vil i løpet av denne oppgaven prøve å besvare følgende spørsmål:

- Hvilke misoppfatninger innenfor noen utvalgte begreper innen statistikk og sannsynlighet finnes det blant elevgrupper i den videregående skolen og hvor utbredt er disse?
- Hvilke begreper har elevene god kontroll på?
- Hva kan gjøres for å forbygge disse misoppfatningene?
- Hva kan gjøres for å bygge robuste begreper innen emnet?

I løpet av oppgaven vil jeg i hovedsak se på alle elevene samlet, men jeg vil også se om jeg kan finne noen interessante resultater blant ulike elevgrupper. De gruppene jeg da vil se på er kjønn, 1T- og 1P-elever, og grupper av hvordan elevene selv vurderer at de presterer og trives i matematikkfaget.

Statistikk og sannsynlighet er et emne som er blitt mer og mer viktig innen matematikkfaget i skolen. Ser vi for eksempel på læreplanverket fra 1997 (KUF, 1996) kommer statistikk og sannsynlighet inn under målområdet "behandling av data". Dette målområdet er kun nevnt på mellom- og ungdomstrinnet, mens det er utelatt på småskoletrinnet. Dersom vi ser på den seneste læreplanen, Kunnskapsløftet (Utdanningsdirektoratet, i.d.), har statistikk og sannsynlighet blitt et eget hovedområde som starter allerede på småskoletrinnet og er med på alle trinn opp til og med den videregående opplæringen. Dette gjør at emnet er svært aktuelt, og det vil være nyttig å forske på elevenes forståelse for begreper innen emnet for å forebygge eventuelle misoppfatninger og gjøre undervisningen så god og effektiv som mulig. Det er viktig at elevene får best mulig utbytte av undervisningen, og dette gjelder selvsagt for alle fag og ikke bare matematikkfaget.

1.3 Disposisjon

Masteroppgaven er inndelt i åtte kapitler. Først kommer innledningen, der jeg kort presenterer oppgaven og problemstillingen min, samt kommer kort inn på statistikk og sannsynlighetshistorie og emnets plass i både dagens samfunn og dagens skole. I neste kapittel vil jeg gi et resymé av utvalgte deler av tidligere forskning og annen relevant litteratur innen læring av statistikk og sannsynlighet. Jeg vil komme litt inn på generell litteratur i tillegg til mer spesifikk litteratur som retter seg mot akkurat det jeg forsker på. I kapittel 3 skal jeg gjennomgå og utdype grundigere de ulike metodene jeg bruker gjennom arbeidet med denne masteroppgaven. I tillegg vil jeg prøve å begrunne valgene av disse metodene. En vesentlig del av empirien finner jeg ved å gjennomføre en diagnostisk test.

Kapittel 4 er det jeg vil kalle mitt hovedkapittel. Her vil jeg legge frem resultatene og analysene fra den tidligere nevnte diagnostiske testen av elevers begrepsoppfattelser i faget. Videre vil jeg i kapittel 5 komme med en oppsummering og diskusjon av noen av de resultatene jeg finner mest interessante. I kapittel 6 vil jeg komme med en konklusjon på hva arbeidet mitt har ledet frem til før jeg i kapittel 7 vil komme med noen forslag som har som mål å føre til en forbedring av undervisningen i faget. I dette kapittelet vil jeg også komme med eventuelle forslag til videre forskning innen emnet. Tilslutt vil jeg gi noen egne refleksjoner over arbeidet mitt. Hvert kapittel vil ha en egen liten innledning der jeg kort presenterer innholdet i kapittelet.

1.4 Glimt fra statistikk og sannsynlighets historie

Ifølge Kapadia og Borovcnik (1991) kan man finne spor av arbeid med sannsynlighet tilbake til de gamle indiske, babylonske og egyptiske kulturer. Det tidligste kjente objektet som er blitt brukt i sjansespill er fra cirka 3500 år f.Kr.. Dette objektet var et bein fra hælen til en sau, og ble kalt astralgus. Disse tror man at man lagde terninger av som sannsynligvis ble brukt til spill. De var spesielt populære blant romerske soldater som brukte de til veddemålsspill. Dette var selvsagt primitive terninger som ikke var spesielt nøyaktige, men man har også funnet overraskende gode terninger fra cirka år 3000 f. Kr.. Disse terningene stammer fra Babylon, er laget av keramikk og er, utrolig nok, bortimot perfekte symmetriske terninger.

Selv om man spilte sjansespill og hasardspill så langt tilbake som 3500 år f.Kr. var det ikke før på 1600-tallet at man virkelig begynte å regne på sannsynlighet, og fortsatt var det i hovedsak spill og gambling man konsentrerte seg om. Hundeland (1996) skriver at på dette tidspunktet fikk matematisk forskning en ny giv i Europa etter å ha ligget nesten brakk siden oldtiden.

I følge Kapadia og Borovcnik (1991) var Gerolamo Cardano (1501-1576) og Galileo Galilei (1564-1642) to av de første vitenskapsmennene som virkelig satte seg inn i emnet sannsynlighet, men det er brevvekslingen mellom Blaise Pascal (1623-1662) og Pierre de Fermat (1601-1665) som regnes som den "offisielle" starten på sannsynlighetsregning (Hundeland, 1996). Pascal og Fermat utforsket mange spennende problemer innen emnet, blant annet et kjent delingsproblem innen spill- og gamblingsverdenen på den tiden: Spillet hadde to deltagere som på forhånd hadde avgjort hvor mange spilleomganger man skulle spille. For eksempel førstemann til å vinne fem spilleomganger vant en premiepott, som forøvrig også ble bestemt på forhånd. Problemet oppstod når spillet ble avbrutt før man hadde kåret en vinner. Hvordan skulle de nå dele potten rettferdig? Her antok man at sannsynligheten for å vinne hver av de gjenstående spillomgangene var lik for hver spiller. Pascal og Fermat brevvekslet hyppig med hverandre, og det var tilslutt Pascal som generaliserte svaret på dette problemet:

> *"Pascal generaliserte dette resultatet og sa at dersom A mangler r spill for å vinne og B mangler s spill, k er antall spill B vinner og n er det totale antall spill som gjenstår, da blir andelene fordelt slik:*
>
> *A fikk en andel på* $\frac{\sum_{k=0}^{s-1}\binom{n}{k}}{2^n}$*, mens B fikk resten. Dette beviste han ved induksjon"* (Syvertsen, 1999, s.12-13).

Nederlenderen Christian Huygens (1629-1695) fikk høre om disse problemene, og han ble med en gang inspirert til å utforske dem grundigere. Han fremstilte etter hvert ideene på en systematisk måte med generaliseringer av løsninger på ulike problemer. Dette gjorde han i avhandlingen "Van rekening in spelen van geluck" som han gav ut i 1660. Det var også Huygens som virkelig gjorde begrepet forventning kjent og som satte sannsynlighet "på kartet" (Kapadia & Borovcnik, 1991, s.32).

I 1713 ble boken "Ars Conjectandi", som for øvrig betyr "kunsten å gjette", utgitt. Denne var skrevet av Jakob Bernoulli og her ble blant annet begrepet utvalg med og uten tilbakelegging brukt for første gang. I siste del av denne boken beviser Bernoulli det som sannsynligvis er hans mest betydningsfulle resultat. Dette var et teorem som han kalte sin "gylne setning", men i dag kjenner vi det bedre som "de store talls lov".

Hundeland (1996) skriver at motivet for utviklingen av sannsynlighetsregning i begynnelsen var å vurdere vinnersjanser i forskjellige spill og veddemål. Senere ble motivet snudd til praktiske anvendelser da forsikringsbransjen kom inn i bildet. Hundeland mener at sannsynlighetsregning har vært noe forskjellig fra andre emner innen matematikken på dette planet, da disse ikke har hatt anvendelse som det viktigste motivet for utviklingen.

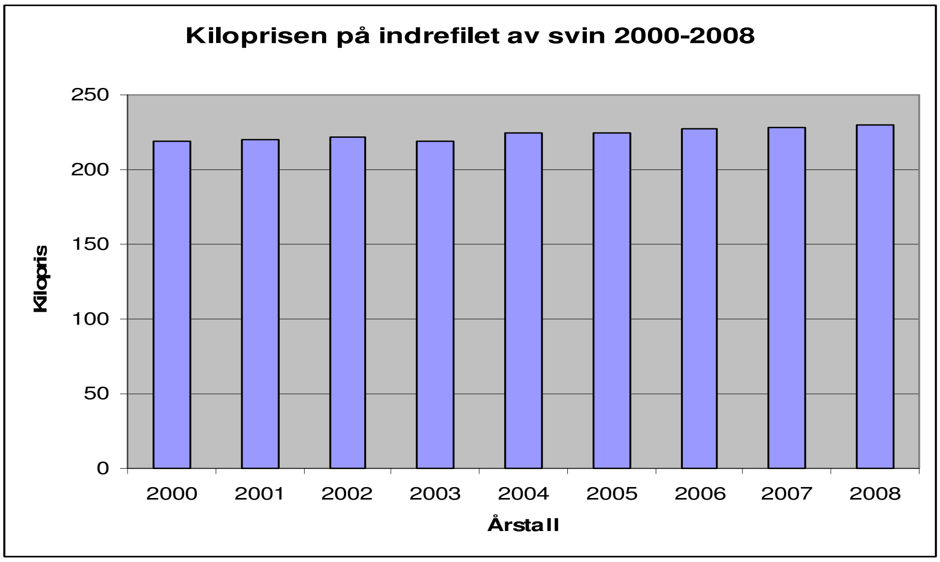
1.5 Statistikk og sannsynlighet i samfunnet

Statistikk og sannsynlighet er noe som vi alle kjenner igjen fra hverdagslivet. Man kan finne elementer fra dette i hverdagslige hendelser som pengespill, aksjemarked, værmeldinger og forsikringsavtaler. Begreper som usikkerhet, sjanse og tilfeldighet fra statistikk og sannsynlighet blir hyppig brukt både i dagligtalen vår og i media. Det faktum at emnet og disse begrepene har fått en mye viktigere plass i samfunnet, gjør at faget også etter hvert har fått en viktigere rolle i skolen.

Problemet med at disse begrepene blir så mye brukt er at vi får en egen "privat" forståelse. Etter hvert er det da stor fare for at man benytter begrepene med ulikt innhold slik at utsagn kan tolkes på ulike måter alt etter hvem som står for tolkningen. Dermed er det meget viktig at disse eventuelle misforståelsene blir rettet opp. Arbeidet i klasserommet blir her avgjørende.

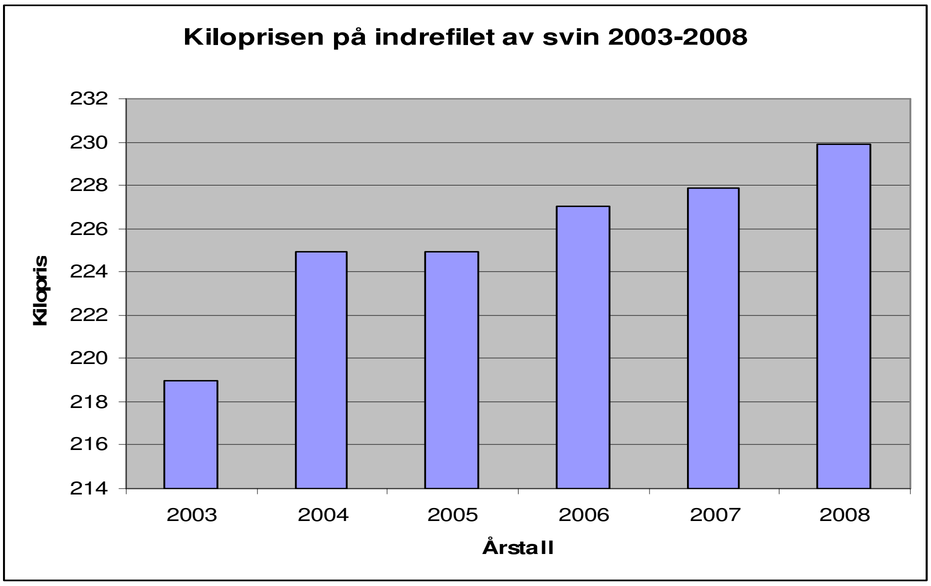
Vi ser så godt som hver eneste dag at det blir brukt grafer og meningsmålinger i forskjellige media som aviser og i fjernsynsprogrammer. Det de fleste muligens kjenner aller best til er alle søylediagrammene som florerer på tv-skjermene våre når det nærmer seg et viktig valg her i landet. Men grafer og meningsmålinger kan også være svært misvisende. Jeg vil nå presentere et eksempel som er laget selv. Alle tall, grafer og opplysninger er fiktive.

La oss si at et ukeblad har hatt en undersøkelse om hvordan kiloprisen på indrefiletkjøtt fra svin har endret seg her i landet siden tusenårsskiftet. De samler inn data og lager en artikkel der de blant annet presenterer denne grafen:



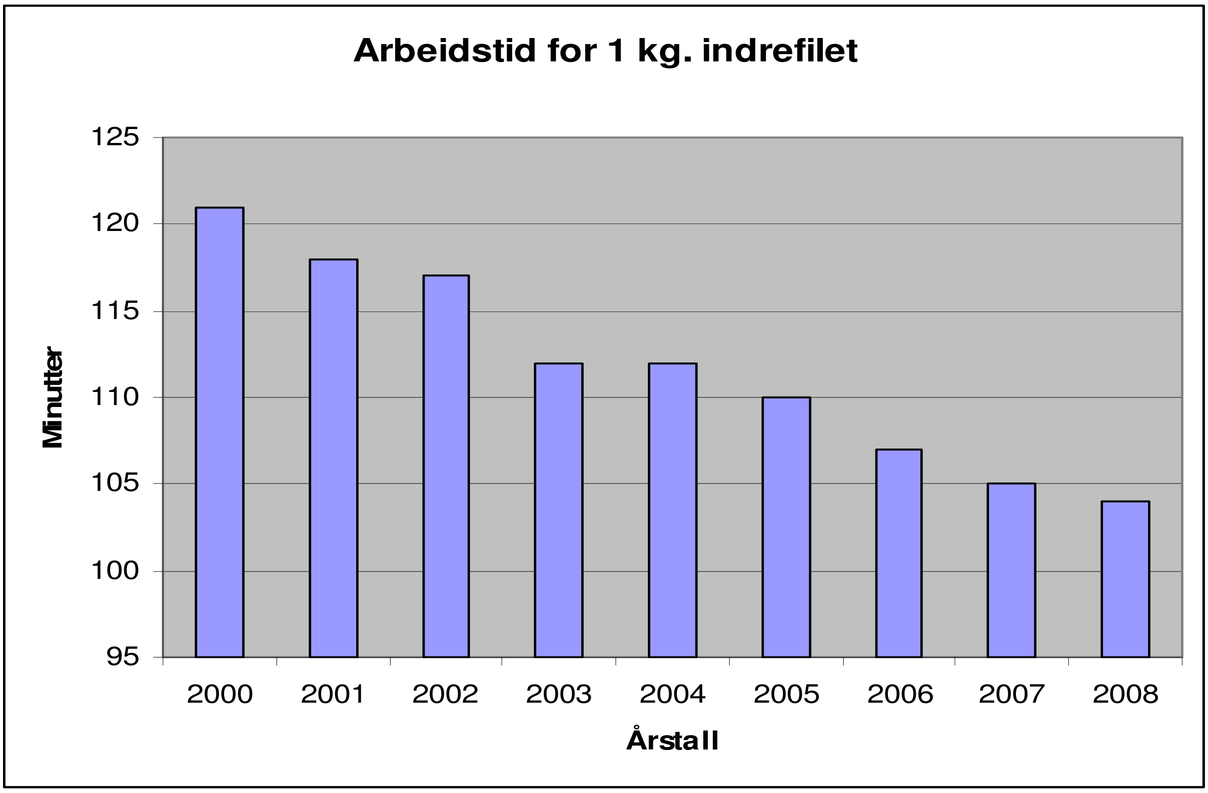
Figur 1.5.1: Eksempel på graf fra en tenkt prisundersøkelse.

Som vi ser er det relativt små endringer i kiloprisene både fra år til år og sett under ett. En avisredaksjon blir oppmerksom på artikkelen og får tillatelse til å bruke resultatene i en egen artikkel. De forandrer litt på grafen og bruker i stedet denne:



Figur 1.5.2: Eksempel på misbruk av graf fra en tenkt prisundersøkelse.

Avisen lager en sak om hvordan kiloprisen har steget så mye de siste fem årene. For å lage mer blest om saken og påvirke leserne endrer de på grafens akser, blant annet kutter de ut de tre første årene for å få en tilsynelatende mer dramatisk stigning. Avisen får mange tilbakemeldinger, mange lesere er forbannet for at prisen på indrefilet er blitt så dyr de siste årene. Etter hvert kommer en kyndig statistiker med et leserinnlegg der han blant annet tar opp kroneverdiens endring i disse årene som er lagt frem i grafene. Han har også laget en graf der han presenterer hvor lenge en gjennomsnittlig norsk arbeider må jobbe for å ha råd til 1 kg indrefilet:



Figur 1.5.3: Arbeidstid for å ha råd til 1 kg indrefilet.

Som vi ser av grafen er ikke kiloprisen på indrefilet av svin gått opp når vi ser prisen i forhold til kroneverdi og den gjennomsnittlige norske timelønnen. Kiloprisen er faktisk tvert imot på det laveste nivået siden tusenårsskiftet.

Dette er bare et tenkt eksempel på hvordan man i media kan bruke og misbruke statistikk for å gagne egne seertall, lyttere eller lesere. Selv om dette kun var et eksempel så viser det at det er nødvendig i det moderne samfunn med en dyp forståelse av statistikk, slik at man kan se bak statistiske argumenter.

1.6 Statistikk og sannsynlighet i skolen

I en studie av DeBeres (1988) i USA fant man blant annet ut at kun 2 % av elever ved videregående skoler som skulle videre på college, hadde hatt et kurs i statistikk og sannsynlighet, mens minst 90 % av elevene hadde hatt et kurs i algebra. Dette sier litt om emnet statistikk og sannsynlighet sin plass i skolen. Det er først i de seneste tiårene at statistikk og sannsynlighet har kommet inn i skolens pensum, både internasjonalt og nasjonalt. Det var blant annet ikke før sent på 1950-tallet at man fikk opprettet et hovedfagsstudium i statistikk ved Universitetet i Oslo.

I vårt nye læreplanverk som kom i 2006 kan vi se at statistikk og sannsynlighet nå har en meget viktig rolle innen matematikkfaget. Statistikk og sannsynlighet er et av tre fagområder, i lag med geometri og tall og algebra, som er med som hovedområde på alle forskjellige årstrinn:

Årssteg	**Hovudområde**				
1.–4.	Tal	Geometri	Måling	**Statistikk**	
5.–7.	Tal og algebra	Geometri	Måling	**Statistikk og sannsyn (bm.: sannsynlighet)**	
8.–10.	Tal og algebra	Geometri	Måling	**Statistikk, sannsyn og kombinatorikk**	Funksjonar
Vg1T	Tal og algebra	Geometri		**Sannsyn**	Funksjonar
Vg1P	Tal og algebra	Geometri	Økonomi	**Sannsyn**	Funksjonar
Vg2T	Geometri	**Kombinatorikk og sannsyn**	Kultur og modellering		
Vg2P	Tal og algebra i praksis	**Statistikk**	Modellering		

Figur 1.6.1: Oversikt over hovedområder i matematikkfaget.

Vi kan se i figur 1.6.1 at elevene skal starte med blant annet statistikk allerede i løpet av de første årene de er på skolen. Etter det andre året står det under kompetansemål for statistikk, sannsynlighet og kombinatorikk:

> *"Mål for opplæringa er at eleven skal kunne*
> - *samle, sortere, notere og illustrere enkle data med teljestrekar, tabellar og søylediagram"* (Utdanningsdirektoratet, i.d.).

Dette viser at elevene skal få en myk start der de skal leke seg litt med tabeller og søylediagrammer som inneholder enkle data. Dette kommer også til syne ved at man ikke har den store progresjonen frem mot endt fjerde år:

> *"Mål for opplæringa er at eleven skal kunne*
> - *Samle, sortere, notere og illustrere data med teljestrekar, tabellar og søylediagram, og kommentere illustrasjonane "* (Utdanningsdirektoratet, i.d.).

Som vi ser har nå "enkle data" fra forrige kompetansemål blitt til kun "data". Det vil i praksis bety at de nå skal kunne jobbe med litt mer avanserte tabeller og søylediagrammer. I tillegg til denne forandringen står det at elevene nå skal kunne kommentere illustrasjonene. Dette vil si at elevene skal være i stand til å gi enkle kommentarer på de tabellene og diagrammene som de arbeider med. Oppsummert kan vi da si at på barnetrinnet skal norske elever bli kjent med noen arbeidsformer innen statistikk.

På mellomtrinnet er det en stor utvikling i faget. Etter det syvende året på skolen står det under kompetansemål for statistikk, sannsynlighet og kombinatorikk:

> *"Mål for opplæringa er at eleven skal kunne*
> - *Planleggje og samle inn data i samband med observasjonar, spørjeundersøkingar og eksperiment*
> - *Representere data i tabellar og diagram som er framstilte digitalt og manuelt, og lese, tolke og vurdere kor nyttige dei er*
> - *Finne median, typetal og gjennomsnitt av enkle datasett og vurdere dei i høve til kvarandre*
> - *Vurdere sjansar i daglegdagse samanhengar, spel og eksperiment og berekne sannsyn i enkle situasjonar"* (Utdanningsdirektoratet, i.d.).

Her ser vi at det har vært en stor progresjon i fra barnetrinnet. Elevene skal nå blant annet kunne lese, tolke og representere tabeller og diagrammer og kunne finne median, typetal og gjennomsnitt av enkle datasett. I tillegg kan vi se at de nå så vidt skal ha startet med sannsynlighetsundervisning. Mellomtrinnet er dermed en meget viktig del av opplæringen innen statistikk og sannsynlighet da det er her grunnlaget blir lagt.

Undervisningen på ungdomstrinnet bygger selvsagt på det elevene lærte på mellomtrinnet. Progresjonen er, som på mellomtrinnet, relativt stor. Dette ser vi også under kompetansemål for statistikk, sannsynlighet og kombinatorikk etter 10.klasse:

> *"Mål for opplæringa er at eleven skal kunne*
> - *Gjennomføre undersøkingar og bruke databasar til å søkje etter og analysere statistiske data og vise kjeldekritikk*

- *Ordne og gruppere data, finne og drøfte median, typetal, gjennomsnitt og variasjonsbreidd, og presentere data med og utan digitale verktøy*
- *Finne sannsyn gjennom eksperimentering, simulering og berekning i daglegdagse samanhengar og spell* (sic!)
- *Beskrive utfallsrom og uttrykkje sannsyn som brøk, prosent og desimaltal*
- *Vise med døme og finne dei moglege løysingane på enkle kombinatoriske problem"* (Utdanningsdirektoratet, i.d.).

Som vi ser skal elevene nå blant annet kunne gjennomføre undersøkinger, analysere statistiske datamateriale og regne ut sannsynligheten i dagligdagse sammenhenger og spill. De skal også kunne uttrykke sannsynlighet som både brøk, prosent og desimaltall. På ungdomstrinnet kommer elevene også i kontakt med enkle kombinatoriske problem for første gang.

I den videregående opplæringen blir elevene delt. De får velge mellom to moduler, T som står for teoretisk og P som står for praktisk. Her skal T-modulen være den mest teoretiske, og dermed blir de svakeste elevene rådet til å ta P-modulen. Disse to modulene går over de to første årene i den videregående skolen. Kompetansemålene for statistikk, sannsynlighet og kombinatorikk i vg1T er som følger:

"Mål for opplæringa er at eleven skal kunne

- *Formulere, eksperimentere med og drøfte enkle uniforme og ikkje-uniforme sannsynsmodellar*
- *Berekne sannsyn ved hjelp av systematiske oppstillingar, og bruke addisjonssetninga og produktsetninga*
- *Bruke omgrepa uavhengnad (bm.: uavhengighet) og vilkårsbunde (bm.: betinget) sannsyn i enkle situasjonar*
- *Lage bionomiske sannsynsmodellar ut frå praktiske døme, og berekne binomisk sannsyn ved hjelp av formlar og digitale hjelpemiddel"* (Utdanningsdirektoratet, i.d.).

Kompetansemålene for vg1P er:

"Mål for opplæringa er at eleven skal kunne

- *Lage døme og simuleringar av tilfeldige hendingar og gjere greie for omgrepet sannsyn*
- *Berekne sannsyn ved å telje opp alle gunstige og alle moglege utfall frå tabellar og ved å systematisere oppteljingar og bruke addisjonssetninga og produktsetninga i praktiske samanhengar"* (Utdanningsdirektoratet, i.d.).

Som vi ser er det kun sannsynlighet elevene skal ha i matematikkundervisningen det første året på videregående skole, noe vi også kan se i figur 1.6.1. Vi ser også som tidligere nevnt at pensumet er klart mest krevende i T-modulen. Det andre året er det ikke lenger kun sannsynlighet elevene skal ha om. Etter vg2T står det under kompetansemål:

"Mål for opplæringa er at eleven skal kunne

- *Gjere greie for omgrepa uavhengnad (bm.: uavhengighet) og vilkårsbunde (bm.: betinget) sannsyn og bruke Bayes' setning på to hendingar*
- *Berekne sannsyn ved ordna utval med og utan tilbakelegging, og ved uordna utval utan tilbakelegging*
- *Rekne med binomisk og hypergeometrisk sannsyn"* (Utdanningsdirektoratet, i.d.).

Etter vg2P står det:

"Mål for opplæringa er at eleven skal kunne

- *Planleggje, gjennomføre og vurdere statistiske undersøkingar*
- *Berekne kumulativ frekvens og finne og drøfte sentralmål og spreiingsmål*
- *Representere data i tabellar og diagram og drøfte ulike dataframstillingar og kva inntrykk dei kan gje*
- *Gruppere data og berekne sentralmål for eit gruppert datamateriale"* (Utdanningsdirektoratet, i.d.).

Som vi både fra disse kompetansemålene og fra figur 1.6.1 er de to modulene svært forskjellige dette året. Mens man kun konsentrerer seg om statistikk i P-modulen, er hovedområdene kombinatorikk og sannsynlighet i T-modulen.

Hvorfor har emnet statistikk og sannsynlighet fått en så stor rolle innen matematikkfaget i skolen i de senere tiår? En grunn til dette er blant annet teknologiens fremskritt i samfunnet. Dagens teknologi gjør at både aviser og fjernsynskanaler lett kan bruke grafiske fremstillinger av grafer og meningsmålinger, og med bruken av disse følger også bruken av ord og uttrykk fra emnet. Dette gjelder ikke bare i mediene, men også i samfunnet generelt blir begreper fra statistikk og sannsynlighet brukt mer. Det er blitt nødvendig å kunne kalkulere og regne på usikkerhet, risiko og konsekvenser. Hacking (referert i Falk & Konold, 1992) har forklart hvordan sannsynlighet påvirker verden i dag: *"Today our vision of the world is permeated by probability, while in 1800 it was not. Probability is the great philosophical success story of the period"*. Hacking mener altså her at verden i dag er gjennomtrengt av sannsynlighet, og at dette er noe som har kommet i løpet av de siste 200 årene.

Denne endringen av bruken av statistikk og sannsynlighet i mediene og samfunnet generelt fører til at barn og unge plukker opp hvordan begreper, grafer og lignende brukes, og de tar dermed til seg kunnskap. Utviklingen av teknologi fører også til at hverdagen blir annerledes for oss lærere. Vi kan for eksempel simulere tusenvis av myntkast, terningkast eller andre sannsynlighetsundersøkelser ved hjelp av et tastetrykk på en datamaskin. Dette sparer tid og elevene får en unik mulighet til å se ved konkrete eksempler hvordan utfall kan forandre seg ved mange simuleringer.

2 Teori

I dette kapittelet vil jeg komme inn på relevant litteratur og tidligere forskning som er gjort innen undervisning og læring av statistikk og sannsynlighet.

På 1970- og 80-tallet var det mange forskere, som for eksempel Fischbein, Kahnemann, Tversky, Shaughnessy, Bergman og Truran med flere, som studerte misforståelser og fordommer innen stokastisk tenkning. Spørsmålene som har dominert arbeidene innen undervisning av statistikk og sannsynlighet på 90-tallet og i de senere år har derimot vært utvikling av stokastisk tenkning og bruken av teknologiske hjelpemidler i undervisningen (Henry, 2003). Forskere har fokusert på å prøve å forstå hvordan barn kan tilegne seg virkelig og teoretisk kunnskap innen statistikk og sannsynlighet og hvordan man på best mulig måte kan bruke dagens teknologi i skolen. I denne oppgaven skal jeg som nevnt undersøke norske elevers begrepsforståelse, og dermed blir det en del eldre litteratur i dette kapittelet da mange undersøkelser innen dette emnet ble gjort for en del år siden.

2.1 Generelt om statistikk og sannsynlighet

Først vil jeg komme litt inn på hva statistikk og sannsynlighet egentlig er. Kapadia og Borovcnik (1991) har blant annet en matematisk metafor for hva sannsynlighet er. De sier at sannsynlighet kan sees på som et måleinstrument for sjanse, på samme måte som en linjal er et måleinstrument for lengde.

Emnet statistikk og sannsynlighet er noe spesielt innen matematikkfaget. Det er ikke helt som andre emner innen faget der man har faste regler som, dersom man følger de slavisk, fører eleven frem til det riktige svaret. Mange elever opplever at deres prestasjoner i statistikk og sannsynlighet er svært forskjellig fra hva de presterer i andre emner innen matematikkfaget. Elever som er sterke i andre emner kan slite i emnet statistikk og sannsynlighet, mens elever som har gjort det svakt i andre emner kan finne statistikk og sannsynlighet som et lyspunkt i matematikkfaget. Dette har blant annet jeg erfart selv da jeg på videregående gikk fra den dårligste karakteren på en prøve om derivasjon og integrasjon til toppkarakter på sannsynlighetsprøven vi hadde noen uker senere.

Det er flere grunner til at sannsynlighetsregning er et noe annerledes emne enn de andre i matematikkfaget. Breiteig (2002) mener en av grunnene er at man ikke alltid kan etterprøve svaret. I vanlig tallregning kan man alltid sjekke svaret med en kalkulator, men dersom en person sier at det er 60 % sannsynlighet for at et fotballag skal vinne en kamp, men de ender opp med å tape, så kan vi ikke bestemt si at denne personen hadde feil.

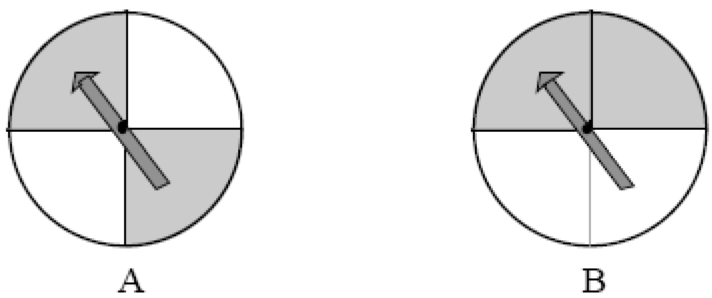
En annen grunn er ifølge Breiteig (2002) at intuisjonen vår ofte svikter innen sannsynlighetsregning. Som Kapadia og Borovcnik (1991) sier*: "People do not seem to have probabilistic intuition in the same way as they have geometric or visual intuition"* (s. 2). For eksempel vil en elev som har kastet seks kron etter hverandre, ved neste kast intuitivt ha større sannsynlighet for mynt. Et annet eksempel er fødselsdager. Dersom vi tenker oss en klasse på 30 elever, hvor stor sannsynlighet er det for at to elever i klassen har fødselsdag på samme dag? Intuitivt setter vi opp 30 elever mot 365 dager og tenker at sannsynligheten er liten siden det er over ti ganger så mange dager som elever. Dersom vi regner ut sannsynlighet får vi faktisk 0,71, betraktelig høyere enn mange ville trodd.

En tredje grunn Breiteig (2002) peker på er at emnet bygger på det han kaller en kvalitativ ”ny” tenkemåte. Forskere har ved hjelp av forskning funnet frem til tre forskjellige matematiske tenkemåter:

- Logisk tenkning
- Årsakstenkning
- Sannsynlighetstenkning

Logisk tenkning er en form for tenkning som elevene ofte bruker i matematikkfaget. Det man gjør er at man erstatter et utsagn med et likeverdig, dette igjen med et nytt, osv inntil man får et enkelt utsagn som kan tolkes (Breiteig, 2002, s. 5). Et eksempel på slik tenkning er når man løser likninger.

Årsakstenkning er når man finner relasjoner mellom årsak og virkning. Ifølge Breiteig (2002) resonnerer man ved å peke på årsaker: Dette ligger til grunn for at resultatet ble slik. Slik tenkning blir brukt i dagliglivet og spesielt i fagene naturfag og samfunnsfag. Et eksempel på at årsakstenkning tar over sannsynlighetstenkningen er følgende:



Figur 2.1.1: To ”lykkehjul”.

En elev får i oppgave å bestemme sannsynligheten for gevinst på disse to hjulene dersom man vinner når viseren stopper på hvitt felt. Eleven svarer at sannsynligheten for gevinst er 50 % på begge lykkehjulene. Deretter får eleven spørsmål om hvilket hjul han ville satset på dersom han fikk muligheten til å spille én gang. Eleven svarer da at ville valgt hjul A fordi dette hjulet har to delte hvite felt. Dette er bare et tenkt eksempel, men det er i teorien det samme eksempelet som Breiteig (2002) presenterer. Breiteig (2002) sier videre at svaret på det første spørsmålet kommer eleven frem til på bakgrunn av et resonnement på forhold, men i det andre spørsmålet inntrer et skifte. Årsakstenkningen overtar og eleven mener nå at en årsak må avgjøre hva svaret er. Da tror eleven at siden det hvite er delt i to felt på det ene hjulet må sannsynligheten være større her, selv om de hvite feltene er like store til sammen på de to hjulene. Dette er altså et eksempel på hvordan årsakstenkningen kan ta over for sannsynlighetstenkningen.

2.2 Tidligere lignende undersøkelser

Det har vært gjort noen lignende undersøkelser tidligere. Noen av funnene og resultatene deres vil jeg nå kort gå igjennom.

David Green (1982) utførte tidlig på 1980-tallet en treårig studie, der over 3000 elever i alderen 11- 16 år ble testet. Dette var, til da, den største undersøkelsen av ungdommers sannsynlighetsforståelse. Green var både interessert i hvilke begreper innen sannsynlighet som disse ungdommene besatt og på hvilket nivå de var i utviklingen i henhold til Piagets teori om nivåutvikling.

Et av de viktigste funnene til Green var blant annet at elevene hadde store problemer med flere av oppgavene som gikk på begrepet forhold. Elevene under 14 år hadde spesielt problemer med forholdsoppgavene. Dette så Green som et stort problem da ha mener at en god forståelse av forholdsbegrepet er viktig innen standard sannsynlighetsregning. Han mener også forhold er spesielt viktig i mer komplekse oppgaver.

Et annet funn var i følge Green (1982) at guttene jevnt over gjorde det bedre enn jentene. Dette gjaldt de fleste oppgavene i testen utenom på kombinatorikkoppgavene. Her gjorde begge kjønn det noenlunde likt, men med jentene et lite hakk foran.

Etter studien sin, konkluderte Green (1982) blant annet med at det var et stort behov for praktisk aktivitet fra en tidlig alder. Dette for å bygge erfaring hos elevene. I tillegg skriver Green (1982) at kun et omfattende program med klassebaserte aktiviteter kan skape bevismaterialer og erfaring nok til å få bort den villedende tenkningen som finnes hos både elever og voksne. Resultatene fra testen tydet også på at diagrammer var til god hjelp for elevenes forståelse. Green mener diagrammer bør brukes overalt hvor det er mulig.

Psykologene Daniel Kahnemann og Amos Tversky (referert i Shaughnessy & Bergman, 1993) har gjort flere undersøkelser innen læring av statistikk og sannsynlighet. En av deres teorier er at statistiske naive mennesker bruker ulike strategier når de skal estimere sannsynligheten for ulike hendelser. Blant annet har de gjort funn som viser at elever bruker strategier som Kahnemann og Tversky kaller representativitet og tilgjengelighet. Disse begrepene kommer jeg tilbake til senere. Kahnemann og Tversky (1983) har også gjort funn som viser at elever har problemer med konjunksjonsloven.

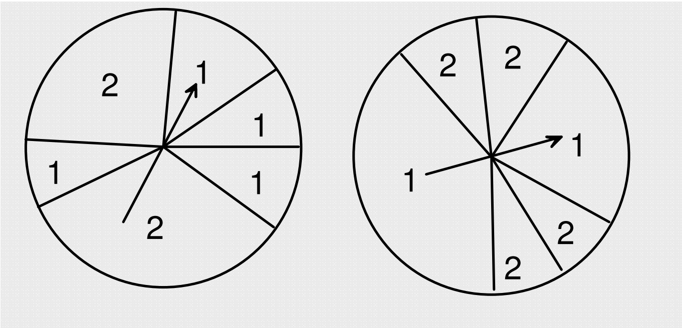
2.3 Matematiske begreper innen statistikk og sannsynlighet

I denne delen vil jeg ta for meg de ulike begrepene som jeg skal teste elevene i. Først vil jeg komme mer inn på hva de ulike begrepene er. Da vil jeg ta for meg litt generelt om begrepet og komme inn på resultater som er gjort i tidligere studier.

2.3.1 Forholdsbegrepet

Green (1983) hadde flere oppgaver som gikk på begrepet forhold i sin undersøkelse. En av disse oppgavene var at man hadde to urner med klinkekuler oppi, den ene hadde tre svarte og en hvit, den andre hadde seks svarte og to hvite. Elevene skulle så avgjøre hvilken urne som gav størst sjanse for å trekke en svart kule. Besvarelsene viste at halvparten av elevene mente at urnen med seks svarte kuler oppi gav størst sjanse. 40 % av hele elevutvalget gav begrunnelsen at i denne urnen var det flere svarte enn i den andre urnen. Dermed kunne Green fastslå at disse elevene ikke hadde brukt forholdsbegrepet som er essensielt innen sannsynlighetsregning. Green fant også ut at dette ikke bedret seg nevneverdig med alderen i hans datamaterial som gikk fra 11 til 16 år.

I en annen oppgave i Greens undersøkelse skulle elevene avgjøre hvilket "lykkehjul" som gav størst sjanse for å få tallet 1:



Figur 2.3.4: Lykkehjulsoppgave fra Greens undersøkelse.

Også på denne oppgaven svarte halvparten av elevene feil. Noen av elevene begrunnet valget av hjulet til venstre med at dette hadde flere områder med enere, mens andre begrunnet det med at dette hjulet hadde tre områder med enere etter hverandre.

Green konkluderte etter denne undersøkelsen blant annet med at elevene hadde problemer med å forstå forholdsbegrepet, som er avgjørende innen sannsynlighetsregning. Han oppdaget også at elever som hadde en god forståelse for nettopp begrepet forhold hadde betraktelig bedre forutsetninger for å gjøre det bra på noen av de mer komplekse oppgavene i testen.

Elevene i disse eksemplene ser ikke etter forhold, de leter etter en årsak for å løse oppgaven. De bruker det Breiteig (2002) kaller årsakstenkning. Elevene resonnerer seg frem til svarene ved å peke på årsaker. I tilfellene over tenker dermed elevene at der det er flest kuler eller flest områder med enere er sannsynligheten for suksess størst.

2.3.2 Representativitet

I følge Kahnemann og Tverskys teori om representativitet så estimerer folk sannsynligheten til et bestemt utfall på bakgrunn av hvordan utfallet representerer den totale populasjonen. Folk tror at selv små utvalg skal representere den fordelingen man finner i hele populasjonen. Et eksempel kan være innen myntkast. Noen tror at utfallet MKKMKM har større sannsynlighet for å forekomme enn for eksempel MMMMMK. MKKMKM virker mer representativ for den 50-50 prosentfordelingen som myntkast representerer enn rekken MMMMMK. Elevene fokuserer på at antallet av kron og mynt i de to rekkene. Folk tror også at utfallet MKKMKM er mer sannsynlig enn MMMKKK fordi den er mer representativ i forhold til den tilfeldigheten som skjer i denne myntkastprosessen. MMMKKK viser et mønster og representerer ikke det tilfeldige som myntkast er. Man kan sammenligne det med folk som spiller rulett. Noen spiller på noen bestemte tall, og dersom det er lenge siden et tall har forekommet så satser de mer på akkurat dette tallet fordi de tror at det er større sannsynlighet for at tallet nå skal ”vinne”. Dette fenomenet er noe som kalles ”the negative recency effect” eller ”the gambler’s fallacy” – spillerens mistak. Dette er også påvist hos elever (Green, 1982). Om man gjør et forsøk og resultatet av første simulering er et bestemt utfall, tror noen at dette utfallet vil ha mindre sannsynlighet i neste simulering.

Selv om representativitet kan lede til misforståelser og feilaktige estimeringer, så kan det også være et godt hjelpemiddel dersom det blir brukt riktig. Kapadia og Borovcnik (1991) poengterer blant annet at representativitet er: ”*a very statistical idea which should allow transfer of traits of samples to underlying populations”.* Selve grunnen til at vi forsøker å ta et tilfeldig utvalg av en stor populasjon er nettopp at vi håper at den er representativ for hele populasjonen. Vi ønsker å forske på det lille utvalget for å finne typiske trekk som vi også kan tillegge populasjonen. På denne måten kan vi si at representativitet er et fundamentalt hjelpemiddel innen statistikk. Dermed blir vår oppgave som matematiske undervisere å skille mellom de omstendigheter der slike strategier blir et hinder for stokastisk tenkning og der de er nyttige og ønskelige. Denne forskjellen er vi pålagt å poengtere til våre elever, fordi det er ikke det at det er noe galt med måten elevene tenker på, det er bare det at de (og vi) har en tendens til å bruke disse strategiene hinsides deres relevante område (Shaughnessy, 1993).

2.3.3 Tilfeldighet

Green (1982) testet også elevenes kunnskaper om begrepet tilfeldighet, og også på dette feltet hadde elevene til dels store problemer. Blant annet hadde de problemer med en oppgave der de skulle avgjøre hvordan de trodde de første snøflakene ville legge seg på en kvadratisk flate. Elevene slet også med en oppgave om kron- og myntkasting. Elevene fikk oppgitt to rekker med 150 kron- og myntforsøk i hver. Deretter skulle de finne ut hvilken rekke som var ”falsk” og hvilken som var ”ekte”. I den falske rekken var det ingen kron- eller myntsekvenser på mer enn tre på rad av en sort. Over halvparten av elevene svarte at den falske rekken var den ”ekte”, de trodde det skulle være et mønster eller en symmetri i de tilfeldige rekkene.

2.3.4 De store talls lov

Et annet begrep det er påvist store problemer med hos elever er de store talls lov. Det vil si at folk neglisjerer utvalgets størrelse. Folk tror at det er samme sjanse for å kaste minst syv kron på ti myntkast som å kaste minst 70 kron på 100 myntkast. I et ekstremt tilfelle vil folk tro at det er like stor sjanse for å kaste minst to kron på tre kast som det er for å kaste minst 200 kron på 300 kast.

Følgende oppgave ble brukt av Kahnemann og Tversky i en studie av studenter på collegenivå som ikke hadde studert verken statistikk eller sannsynlighet:

> Anta at sjansen for en gravid kvinne å føde en gutt, er den samme som for å føde en jente. Gjennom et tidsperspektiv på ett år, på hvilket sykehus tror du det er flest dager der minst 60 % av de nyfødte er gutter?
> a) Et stort sykehus
> b) Et lite sykehus
> c) Det er ingen forskjell

De fleste elevene valgte svaralternativet ”c) det er ingen forskjell”. Også Shaughnessy og Bergman (1993) brukte denne oppgaven i en undersøkelse, og resultatene viste at 48 av 80 studenter mente at det ikke var noen forskjell, mens resten var splittet relativt jevnt i synet om hvorvidt sannsynligheten var størst ved et stort eller lite sykehus. Shaughnessy og Bergman ba også studentene om å forklare valget sitt, og de fleste som hadde valgt alternativ c) svarte da at sannsynligheten for å få en spesiell prosentandel gutter er lik ved alle sykehus. Det studentene har neglisjert da er det vi kaller de store talls lov.

Vi kan tenke oss to konkrete eksempler, et lite sykehus og et stort sykehus. På det lille sykehuset blir det en tilfeldig valgt dag født tre babyer, mens det blir født ti babyer på det store sykehuset. For å få minst 60 % gutter ved det lille sykehuset må vi ha minst to gutter blant de tre nyfødte. Dette kan sees på som en binomisk fordeling, sannsynligheten for å føde en gutt er konstant lik 0,50. For å regne ut sannsynligheten bruker vi følgende formel:

$P(X = x) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}, x = 0,1,2\ldots, n$

I vårt tilfelle med det lille sykehuset får vi da $P(X = 2) = 0{,}375$ og $P(X = 3) = 0{,}125$. Til sammen får vi altså at sannsynligheten for å få minst 2 gutter blant de tre nyfødte er 0,50.

For å få minst 60 % gutter ved det store sykehuset må vi ha minst seks gutter blant de ti nyfødte. Ut fra formelen ovenfor får vi da: $P(X = 6) \approx 0{,}205$, $P(X = 7) \approx 0{,}117$, $P(X = 8) \approx 0{,}044$, $P(X = 9) \approx 0{,}010$ og $P(X = 10) \approx 0{,}001$. Til sammen får vi at sannsynligheten for å få minst seks gutter blant de ti nyfødte er tilnærmet lik 0,377. Dermed kan vi se at desto mindre utvalgets størrelse er, desto større sannsynlighet er det for vi får store avvik fra den relative frekvensen.

2.3.5 Middelverdi

Det har blitt påvist store misoppfatninger innen statistikk, til og med blant forskere har man funnet misoppfatninger. Et eksempel på en forskerfeil er at man skal gjenta en tidligere undersøkelse, men tror at det kan er nok med et lite utvalg siden utvalget skal være representativt uansett utvalgets størrelse. Når forskere viser slike misforståelser er det ikke rart at elever eller studenter har problemer med grunnleggende begreper, som for eksempel middelverdi. Pollatsek, Lima og Well (referert i Shaughnessy & Bergman, 1993, s.188) påviste blant annet grunnleggende misoppfatninger om gjennomsnitt når data må vektes ut fra utvalgets størrelse. De gav collegestudenter, uten noe spesiell erfaring innen statistikk, flere oppgaver av denne typen:

> Poenggjennomsnittet på en prøve til alle ungdomsskoleelever i et fylke er 400. Du gjør et tilfeldig utvalg på ti elever. Den første eleven du plukket ut hadde 250 poeng på prøven. Hvilket poenggjennomsnitt ville du nå forvente på hele utvalget på ti elever?

Mange av studentene svarte at gjennomsnittet ville ligge på 400. Disse studentene bruker kanskje representativitet som en ledesnor, da de tror at selv et så lite utvalg skal gjenspeile hele utvalgets gjennomsnitt. Det kan også være at studentene forveksler gjennomsnittet til hele populasjonen med utvalgets gjennomsnitt. Gjennomsnittet til hele populasjonen vil fremdeles være 400, mens det forventede gjennomsnittet til utvalget vil endre seg etter at vi har fått ny informasjon. Vi kan anslå det forventede gjennomsnittet til utvalget ved å vekte de ni ukjente:

$$\frac{400 \cdot 9 + 250}{10} = \frac{3850}{10} = 385.$$

Det er påvist grunnleggende misoppfatninger om gjennomsnitt når data må vektes ut fra utvalgets størrelse (Shaughnessy & Bergman, 1993).

2.3.6 Betinget sannsynlighet og uavhengighet

Betinget sannsynlighet og uavhengighet er to andre begreper som forskning har påvist at elever sliter med. Spesielt sliter elever med betinget sannsynlighet når man skal estimere sannsynligheten for et utfall gitt at noe har skjedd etterpå. Falk har et eksempel på dette, noe som senere er blitt kjent som Falks fenomen:

En urne inneholder to hvite og to svarte baller. To baller blir trukket ut i rekkefølge, uten tilbakelegging.

1) Hva er sannsynligheten for at den andre ballen er hvit, gitt at den første var hvit?
2) Hva er sannsynligheten for at den første ballen var hvit, gitt at den andre er hvit?

Elevene så raskt at svaret på 1) var 1/3 fordi etter å ha trukket en hvit ball først har man en hvit og to svarte baller igjen i urnen. Oppgave 2 slet derimot elevene betraktelig mer med. De hadde vanskelig for å skjønne, eller til og med tro, at svaret på denne oppgaven også var 1/3. De fleste tror svaret er 1/2, mens andre mener at man ikke en gang kan gjøre denne oppgaven, da de mener at resultatet ved den første uttrekningen umulig kan være avhengig av den andre. Falk mener at grunnen til at dette problemet er så vanskelig for elever å skjønne er at de ikke ser noen årsakssammenheng mellom resultatet i den andre trekningen og den første. Derimot er det en sannsynlighetssammenheng.

Et annet velkjent eksempel på denne type oppgave er det vi kaller ”Monty Hall-problemet” (Shaughnessy & Bergman, 1993). Denne oppgaven ble til gjennom et tv-show. Man hadde en deltaker som ble vist tre dører. Bak en av disse dørene skjulte det seg en stor premie, bak de to andre var det små trøstepremier. Deltakeren ble bedt om å velge seg en dør før tv-verten åpnet en av de to andre dørene. Tv-verten vet hva som er bak dørene og velger alltid en dør med trøstepremie. Deretter får deltakeren valget mellom å bytte dør eller beholde sitt opprinnelige valg. Og det er her problemet oppstår. Får deltakeren noen form for fordel ved dette valget? Hva er sannsynlighet for å få hovedpremien dersom deltakeren alltid bytter dør, og hva er den dersom deltakeren aldri bytter?

Denne oppgaven har vært mye debattert, blant annet i amerikanske Parade Magazine der debatten ble så heftig at den senere nådde fremsiden til selveste New York Times. I Parade Magazine var det Marylin von Savant, som har verdens høyeste registrerte IQ ifølge Guiness Book of Records, som forsøkte å forklare denne oppgaven etter å ha fått et leserspørsmål. Hun skrev at man har 1/3 sannsynlighet for å vinne dersom man ikke bytter, mens man har 2/3 sjanse for hovedpremien om man bytter dør. Dette fikk hun enormt mange tilbakemeldinger på, de fleste mente hun hadde feil. Mange lesere, blant annet flere med doktorgrader som jobbet ved diverse universiteter, mente at sannsynligheten var 1/2 for begge tilfellene etter at tv-verten har tatt vekk den ene døra med trøstepremie fordi da står man tilbake med to dører og to muligheter, og en av disse gir hovedpremien.

Det var von Savant som hadde riktig i denne diskusjonen, og hun forklarte til slutt problemet på denne måten: La oss si du velger dør 1, da har du 1/3 sjanse for at hovedpremien er bak denne døra, mens det er 2/3 sjanse for at hovedpremien ikke er bak denne. Så kommer tv-verten inn og gir deg et tips. Dersom hovedpremien er bak dør 2, vil tv-verten åpne dør 3. Er hovedpremien bak dør 3, åpner han dør 2. Det vil si at dersom du bytter vil du vinne hovedpremien både dersom den er bak dør 2 og bak dør 3. Dersom du ikke bytter vil du kun vinne dersom hovedpremien er bak dør 1.

Shaughnessy og Bergman (1993) har gitt denne oppgaven til elever og studenter i alt fra ungdomsskole til universitetsnivå. De fleste har trodd at så fort verten åpner en av dørene med trøstepremie bak, så stiger sjansen for å vinne hovedpremien fra 1/3 til 1/2. Shaughnessy og Bergman merket en dramatisk forbedring i elevenes og studentenes forståelse av denne oppgaven etter de hadde gjort en simulering av situasjonen. De fikk tre strategier som de skulle gjennomføre 100 ganger hver. Strategiene var 1) Hold deg til ditt opprinnelige valg av dør, 2) gjør et myntkast, og bytt dør dersom mynten lander på kron, og 3) bytt dør. Etter å ha gjennomført hver av strategiene 100 ganger innser elevene at å bytte dør etter den ene har blitt åpnet faktisk utgjør en stor forskjell, og de oppdager at det vil være lurt å bytte hver eneste gang. Suksessen som Shaughnessy og Bergman hadde ved å simulere ”Monty Hall-problemet” kan sees på som et bevis på at forslagene til Falk, Bentz og Shaughnessy om å modellere sannsynlighetsproblemer ved simuleringer kan være en god måte for å konfrontere og overkomme misforståelser (Shaughnessy & Bergman, 1993).

2.3.7 Konjunksjonsloven

I sannsynlighet har vi noe som heter konjunksjonsloven. Den går ut på at sannsynligheten for at to ulike hendelser inntreffer samtidig er mindre enn sannsynligheten for at en av hendelsene inntreffer separat fra den andre. Ett eksempel er at sannsynligheten for at en tilfeldig valgt mann er over 65 år *og* skallet, er mindre eller lik sannsynligheten for at han er over 65 år. Forskning har vist at folk ikke følger denne regelen når de estimerer sannsynligheten for hendelser.

Kahnemann og Tversky (1983) fant blant annet ut at collegestudenter uten noen spesiell erfaring fra statistikk mente at det var flere mennesker som var over 55 år *og* hadde hatt hjerteinfarkt enn folk som kun hadde hatt hjerteinfarkt. Grunnen til at folk gjør denne feilen kan ha mange forklaringer. I dette tilfellet knytter nok folk sammen de to variablene alder og hjerteinfarkt, og ser alder som en årsak til infarkt. En annen ting som kan forsterke dette er at vår egen erfaring spiller inn. Kanskje er de fleste vi kjenner som har hatt hjerteinfarkt eldre folk. Da bruker man altså tilgjengelighetsstrategien som er nevnt tidligere.

En annen årsak til at noen folk misforstår konjunksjonsloven kan være at i noen tilfeller tolker de oppgaven på feil måte. Det kan være at de blander inn betinget sannsynlighet. I eksempelet ovenfor kan det være at noen av studentene forstår oppgaven som sannsynligheten for hjerteinfarkt, gitt at personen er over 55 år.

2.3.8 Tilgjengelighet

Noen ganger bedømmer folk sannsynligheten for et utfall på bakgrunn av noen spesielle tilfeller som ligger nært til vår erindring. Et eksempel er følgende to spørsmål: 1) Hvor sannsynlig er det at et tilfeldig valgt ord har bokstaven -n som nest siste bokstav? 2) Hvor sannsynlig er det at et tilfeldig valgt ord slutter på – ene? I dette tilfellet tror mange at sannsynligheten er størst i spørsmål 2) fordi man assosierer flere ord til – ene endingen enn man gjør til ord som har n som nest siste bokstav. Dette blir kalt å bruke en tilgjengelighetsstrategi (Shaughnessy & Bergman, 1993), man estimerer sannsynligheten for et utfall ut fra noen spesielle tilfeller som man husker lett.

Shaughnessy og Bergman (1993) nevner et annet eksempel på bruken av denne strategien. Det er når folk blir bedt om å estimere komplekse kombinatoriske oppgaver. Dersom de ikke har noen erfaring innen telleteknikker vil folk tro at man kan lage flere grupper med for eksempel to av ti folk enn med åtte av ti folk. Dette fordi det kan umiddelbart tenkes at det er enklere å sette sammen en gruppe bestående av kun to mennesker, mens sannheten er at man kan danne eksakt like mange toergrupper som åttergrupper av en gruppe på ti.

2.4 Hvordan kan vi endre misoppfatninger?

Shaughnessy og Bergman (1993) poengterer at når vi underviser statistikk og sannsynlighet så starter ikke elevene med "blanke ark". De har allerede en egen forståelse av mange begreper innen statistikk og sannsynlighet basert på sin erfaring (eller mangel på). Shaughnessy og Bergman fortsetter med at vi lærere ønsker å gi studentene muligheten til å bygge en matematisk modell for statistikk og sannsynlighet som de senere kan bruke. Hvordan kan vi da gjøre dette når de ofte møter opp med potensielt misledende erfaringer og feiloppfatninger?

Det har vært gjort flere forsøk, blant annet av Batanero, Godino og Canizares og Shaughnessy, på å undervise statistikk og sannsynlighet på en slik måte at man konfronterer elevenes misoppfatninger. Shaughnessy (1993) utførte blant annet et tiukers langt kurs der man brukte eksperimenter og forsøk i undervisningen. Selv om Shaughnessy fant forbedringer i elevenes begreper etter kurset, var det fortsatt en del elever som holdt fast ved sine misoppfatninger. Del Mas og Bart (referert i Shaughnessy, 1993) gjorde blant annet funn på at noen elever må bli tvunget til å gjette resultatet av et forsøk på forhånd for deretter å sammenligne dette med det virkelige resultatet for å overkomme misoppfatninger.

3 Metode
Datainnsamlingen er i hovedsak en diagnostisk test som er gjennomført på 141 elever, fordelt på tre skoler og åtte klasser. Jeg har i tillegg intervjuet fire elever fra den ene klassen.
I dette kapittelet vil jeg ta for meg disse og andre metoder jeg har brukt gjennom arbeidet med denne oppgaven. Jeg vil også komme med en forklaring til de valgene av metoder som er gjort.

For å belyse problemstillingen min har jeg laget en test bestående av 14 oppgaver innen statistikk og sannsynlighet, i tillegg til to spørsmål der elevene blir bedt om å vurdere sine prestasjoner og sin trivsel i matematikkfaget. 141 elever fordelt på åtte ulike klasser på første steget i den videregående skolen har besvart denne testen, og de resultatene som disse gir meg er det jeg legger til grunn når jeg besvarer problemstillingen min.

3.1 Kvantitativ og kvalitativ metode
I denne masteroppgaven har jeg brukt både kvantitativ og kvalitativ metode. Data som er samlet inn består som sagt i hovedsak av tester gjennomført på 141 elever. Det er også gjort intervju med fire av elevene som deltok i denne undersøkelsen. Intervjuene inngår som en kvalitativ metode, mens jeg i resten av oppgaven har jobbet med kvantitative metoder.

Den kvalitative metoden kjennetegnes av at forskerne tar del i en naturlig setting. De tar personlig del i de situasjoner de ønsker å finne mer ut om. Dette gir forskerne en mulighet til å personlig oppleve situasjonene i detalj i det de oppstår. I den kvalitative metoden prøver forskeren å skape et tillitsforhold mellom forskeren og informanten, men dette bør man helst prøve å gjøre uten å påvirke settingen mer enn nødvendig. Et eksempel på at settingen er forandret kan være dersom man er inne i et klasserom og observerer, og elevene endrer oppførselen sin på grunn av at det er noe nytt og ukjent i klasserommet. Dette vil være svært uheldig for forskeren.

Forskning innen den kvalitative metoden vil ofte bære et personlig preg da forskningen i hovedsak går ut på å tolke og rapportere. Det vil si at man må prøve å plukke ut det man mener er det viktigste i datainnsamlingen. De vanligste hjelpemidlene innenfor kvalitativ forskning er intervju, observasjon, tekstanalyse og audio- og/eller videoopptak.

Den kvantitative forskningsmetoden kjennetegnes ved at man som regel har store mengder data å jobbe med. De vanligste hjelpemidlene for datainnsamling innen kvantitativ forskningsmetode er spørreundersøkelser, eksperimenter og offisielle statistikker.

3.2 Diagnostisk test
Spørreundersøkelser som jeg har brukt, har både positive og negative sider ved seg. På den positive siden har vi at man kan få samlet inn store mengder data fra mange folk på en relativt enkel og økonomisk og arbeidseffektiv måte. På den negative siden har vi det faktum at man ikke får muligheten til å gå så dypt inn i de ulike situasjonene man ønsker å forske på. Man må stole på de svarene som hvert individ gir om deres kunnskaper, holdninger eller erfaringer. Dette gjør at gyldigheten til den informasjonen man får gjennom en studie der man bruker spørreundersøkelser som metode er avhengig av ærligheten til deltakerne (Mertens, 2005, s. 167).

Min test inneholder 14 oppgaver om statistikk og sannsynlighet som enten er laget av meg selv i samarbeid med mine veiledere eller plukket ut fra lignende, tidligere undersøkelser. De oppgavene som er hentet fra andre undersøkelser er oppgaver der forskeren har fått interessante funn. Disse oppgavene gir meg også mulighet til å sammenligne mine resultater med resultatene deres. I tillegg til disse 14 oppgavene har jeg på slutten med to spørsmål der elevene blir bedt om å vurdere sine prestasjoner i matematikkfaget og hvordan de trives i faget. Elevene fikk 45 minutter (én skoletime) på å fullføre testen. Jeg vil nå gå igjennom hva de ulike oppgavene går ut på og begrunne utvalget. Testen kan ses i sin helhet i vedlegg 1.

Den første oppgaven er en ”urneoppgave” der man trekker ut sorte og hvite kuler fra en urne med tilbakelegging. Elevene skal avgjøre hvilken kule som er mest sannsynlig å bli trukket ut neste gang. Her blir elevenes kunnskaper om forholdsbegrepet testet. Denne type oppgave har vært vanlig å ha med i slike undersøkelser, om enn ikke i nøyaktig samme form som den jeg har. Ideene til oppgaven er hentet fra både Green (1982) og Breiteig og Brekke (1994).

Oppgave 2 er en oppgave om myntkast. Det er gjort seks myntkast og elevene får listet opp fire ulike rekkefølger av aktuelle resultat, som for eksempel MKMMKM (M=mynt, K=kron). De skal deretter avgjøre hvilken rekkefølge som er mest sannsynlig eller om alle er like sannsynlige. Ideen til oppgaven er hentet fra Green (1982) og Breiteig og Brekke (1994). Oppgave 3 er en lignende oppgave som oppgave 2, men denne går på lottorekker i stedet for myntkast. Elevene skal avgjøre hvilken av tre oppgitte lottorekker som har størst sannsynlighet for å bli trukket ut som vinnerrekke eller om alle har like stor sjanse. Disse to oppgavene valgte jeg på grunn av at jeg ville teste elevene innen Kahnemann og Tverskys teori om representativitet i tillegg til å undersøke elevens oppfatning av tilfeldighet. Ideen til denne oppgaven er hentet fra Breiteig og Brekke (1994).

Oppgave 4 og 5 er to oppgaver som begge tar for seg de store talls lov. Oppgave 4 handler om barnefødsler og er hentet fra Greens undersøkelse (1982), mens oppgave 5 handler om gjennomsnittshøyden blant rekrutter, og den er laget i samarbeid med mine veiledere. Oppgavene er med i testen for å avdekke elevenes forståelse av begrepet de store talls lov.

Oppgave 6 handler om gjennomsnitt. Elevene får vite vekten på noen egg og skal finne ut hvor mye det siste egget i pakken må veie for at gjennomsnittet skal bli som oppgitt. Dermed må elevene bruke den informasjonen de har fått og vekte denne riktig for å finne vekten på det siste egget. Oppgaven ble laget i samarbeid med mine veiledere og tester som nevnt elevenes begrep om og ferdigheter innen gjennomsnittsregning og vekting av gjennomsnitt.

Oppgave 7 er en oppgave med to deloppgaver, der den ene er hentet fra Greens undersøkelse (1982), mens ideen til den andre er hentet fra Breiteig og Brekke (1994). Oppgaven er en mer ”generell” sannsynlighetsoppgave der elevene skal svare på hvilke utfall som er mest sannsynlig å inntreffe. Det kan bli interessant å se om noen elever blander inn betinget sannsynlighet når de løser denne oppgaven. Jeg har valgt å ta denne oppgaven med fordi man tidligere har gjort noen interessante funn med lignende oppgaver.

Oppgave 8 er en oppgave som jeg har tatt med for å teste ut elevenes forståelse for konjunksjonsloven. Ideen til oppgaven er hentet fra Breiteig og Brekke (1994). Oppgaven er tatt med fordi forskere, blant annet Kahnemann og Tversky, har påvist at folk noen ganger estimerer sannsynligheten for at to ulike hendelser inntreffer som større enn sannsynligheten for at kun den ene hendelsen inntreffer.

Oppgave 9 er hentet fra Greens undersøkelse (1982) og går ut på å teste elevenes forståelse for begrepet forhold. De blir presentert med to bokser med et ulikt antall sorte og hvite kuler oppi og skal avgjøre fra hvilken boks det er størst sannsynlighet for trekke ut en hvit kule. Denne oppgaven valgte jeg å ta med i testen min fordi Green fant noen interessante resultater i sin undersøkelse med lignende oppgaver.

Oppgave 10 er hentet fra Breiteig og Brekke (1994). Den er også med for teste elevenes forståelse for forholdsbegrepet, men i denne oppgaven må også elevene bedømme vinkler og regne ut prosenter ut fra et slags kakediagram. Også denne oppgaven er tatt med fordi det er funnet resultater som tyder på at elever sliter med dette i tidligere undersøkelser.

Oppgave 11 er en "lykkehjulsoppgave". Ideen til oppgaven blant annet brukt av Green (1982), men denne versjonen er hentet fra Breiteig og Brekke (1994). Elevene skal velge mellom tre "lykkehjul". Hvert lykkehjul har forskjellig vinnersjanse og forskjellig premie. Dermed må de regne ut vinnersjansen på hvert hjul og den forventede tilbakebetalingen man får.

Oppgave 12 er en oppgave som også er brukt av Skaar og Syvertsen (2007). De brukte denne oppgaven som en introduksjon til emnet sannsynlighetsregning. Elevene får oppgitt reglene i et spill og så skal de vurdere om spillet er rettferdig. Jeg har denne med i testen blant annet fordi Skaar og Syvertsen har fått interessante resultater i bruken av den.

Oppgave 13 er veldig lik oppgave 1. Elevene får se en illustrasjon av en urne med hvite og sorte kuler oppi, men i denne oppgaven skjer uttrekningen av kuler uten tilbakelegging. Ideen til oppgaven er hentet fra Breiteig og Brekke (1994).

Oppgave 14 er en oppgave som jeg har laget i samarbeid med veilederne mine. Elevene skal vurdere om de er enige med noen forskere om hvorvidt røyking gir økt sannsynlighet for å få lungekreft eller om de er enige med en person som har opplevd at alle fire røykende besteforeldre har levd godt og lenge uten å få lungekreft. Oppgaven er laget for å teste elevenes forståelse for begrepet tilgjengelighet.

De to siste oppgavene i testen er spørsmål der elevene skal vurdere hvordan de selv mener de presterer i matematikkfaget og hvordan de liker matematikkfaget. Disse spørsmålene tok jeg med for å kunne ha muligheten til å se på resultatene på de andre oppgavene i forhold til hva de svarte her.

Testen ble dermed tilslutt på 14 oppgaver i tillegg til to spørsmål. Før jeg besøkte skolene utfordret jeg en av mine medstudenter til å gjennomføre testen. Dette gjorde jeg for å kunne avdekke eventuelle feil i testen eller for å se om det var noen oppgaver som kunne være for vanskelige, utydelige eller om det var noen oppgaver som kunne misforstås eller var uklare. I tillegg fikk jeg en bekreftelse på at den planlagte tidsbegrensningen på en skoletime kunne være passelig.

3.3 Intervju

Jeg gjorde som nevnt tidligere fire intervjuer med elever fra en av klassene fra den ene skolen jeg var innom. Disse fire elevene var valgt ut på forhånd av læreren deres. De eneste retningslinjene jeg hadde gitt læreren før utvelgelsen var at de skulle være på forskjellig matematisk nivå og i tillegg ville jeg gjerne ha elever som ikke var redde for å prate litt.

Intervjuene ble gjort like etter elevene hadde gjennomført testen slik at oppgavene og svarene de hadde gitt fortsatt skulle være friskt i minne. På forhånd hadde jeg valgt ut noen oppgaver som jeg trodde jeg kunne få noen interessante resultater på. Det negative med å gjøre intervjuene umiddelbart etter testen var at jeg ikke fikk tid til å forberede meg. Jeg fikk ikke lest igjennom det elevene hadde svart og måtte ta det meste på sparket. Dette førte til at jeg ikke fikk så mye ut av intervjuene, noe som også viste seg når jeg hørte igjennom de senere. På grunn av dette bestemte jeg meg for å ikke analysere intervjuene i denne oppgaven, men jeg velger likevel å nevne arbeidsprosessen med intervjuene da dette har vært en del av arbeidet med denne oppgaven.

3.4 Skolene

Som nevnt tidligere gjennomførte jeg testen på 141 elever på første trinnet i den videregående skolen. Disse elevene var fordelt på tre skoler og alderen på elevene var cirka 15-16 år. Utvalget av skoler skjedde ikke tilfeldig. Skolene ble valgt av praktiske årsaker som beliggenhet og tidligere kontakt. To av de tre skolene var lokalisert i industriområder på vestlandet, mens den siste var en skole i Sør-Norge. Jeg nevner kort dette selv om ikke hovedhensikten med denne oppgaven er å sammenligne skolene eller å bruke beliggenheten til skolen i analysen. Det ble tidlig tatt kontakt med skoler og alle skolene jeg kontaktet var veldig behjelpelige. Jeg sendte et brev til skolene før gjennomføring der jeg forklarte litt mer hva dette gikk ut på og hva svarene skulle brukes til (se vedlegg 2).

Skolene er kodet som skole 38, 39 og 40 videre i oppgaven. Skole 40 var en stor skole med mange elever og klasser, mens 38 og 39 var relativt små skoler. Den store skolen var på en måte "hovedskolen" min. Her ble testen gjennomført på seks klasser, tre 1T-klasser og tre 1P-klasser. På de to små skolene var jeg kun inne i en klasse på hver skole. Disse klassene var blandet med både 1P- og 1T-elever. Ingen av klassene som ble testet hadde hatt om emnet statistikk og sannsynlighet i pensum når testen ble utført. Dermed hadde alle elevene noenlunde det samme utgangspunktet. Sist gang de hadde hatt om emnet var på ungdomsskolen.

Det var som nevnt tidligere til sammen 141 elever som gjennomførte testen. Av disse var det 31 elever fra rene 1P-klasser, 69 elever fra rene 1T-klasser og 41 elever fra de to blandede klassene. Av disse 41 elevene i de blandede klassene vet jeg ikke hvilke elever eller hvor mange elever som hadde valgt T- eller P-modulen.

3.5 Hjelpemidler

I arbeidet med denne masteroppgaven har jeg brukt flere hjelpemidler, bare i arbeidet med å lage testen ble det brukt flere ulike. Noen oppgaver ble funnet i andre lignende undersøkelser, mens andre ble laget selv. I utformingen av oppgavesettet ble både programmet Paint og programmet Cabri2 Plus brukt til å lage forskjellige figurer til oppgavene. Selve testen, samt alt annet skriftlig arbeid, ble skrevet i Microsoft Word.

I prosessen med å rette testene brukte jeg programmet Microsoft Excel. Dette programmet ble for øvrig også brukt til å lage noen av grafene og histogrammene som er brukt i oppgaven. I rettingsprosessen kodet jeg elevenes svar og laget et kodeskjema som forteller hva de forskjellige kodene på hver oppgave står for (se vedlegg 3).

Programmet Minitab ble brukt til statistisk analyse av utvalgte data. Noen histogrammer og alle boksplottene i denne masteroppgaven er laget ved hjelp av dette programmet.

3.6 Validitet

Hvordan kan jeg være sikker på at de resultater jeg kommer frem til i denne oppgaven er pålitelige? Og gir resultatene fra testen meg i det hele tatt mulighet til å svare på problemstillingene? Noen av oppgavene som er med i testen min er prøvd ut tidligere og er på en måte kvalitetssikret. Videre kan jeg si at det virket som om elevene tok testen på alvor, samtidig som at alle lærere som var involvert støttet opp og bidro til at testen ble gjennomført på en skikkelig måte. Testen ble gjennomført på en av mine medstudenter på forhånd for å forsikre at elevene hadde nok tid til å gjennomføre og for å forsikre at ingen av oppgavene var uklare eller hadde upassende vanskelighetsgrad.

I selve testen har hver oppgave en eller flere linjer ledig der elevene ble bedt om å gi en forklaring på hvorfor de har svart som de har gjort. Dette er med på å ta bort tilfeldigheter som tipping, da flesteparten av oppgavene er avkrysningsoppgaver med flere valg. Retteprosessen var relativt enkel på grunn de konkrete svarene.

4 Analyse

I denne delen av oppgaven vil jeg undersøke datainnsamlingen min som et ledd i å prøve å finne svar på forskningsproblemene fra kapittel 1. Jeg vil gå igjennom svarene til elevene på de ulike oppgavene i fra den diagnostiske testen. Alle svarfrekvenser vil jeg sette inn i en tabell der det riktig svaret vil være uthevet med fet skrift. Tilslutt vil jeg sammenligne enkelte av oppgavene som tester ut noen av de samme begrepene for å se om det er noen sammenheng mellom svarene til elevene på de ulike oppgavene.

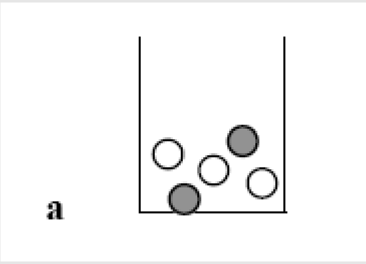
I noen av resultattabellene vil summen av svarfrekvensene være høyere enn antallet elever. Grunnen til dette er at noen elever har krysset av for flere alternativer på noen oppgaver. Dersom noen har krysset av på flere alternativer har jeg valgt å kode dette som galt selv om de skulle ha krysset av for det riktige svaret også. Det vil si at det eneste som gir riktig svar i tabellene er de elevene som kun har krysset av for det riktig alternativet.

4.1 Oppgave 1

Oppgave 1 var en urneoppgave, der man gjentatte ganger trakk, med tilbakelegging, en kule opp av en urne. Oppgaven tester elevenes forståelse for begrepet forhold, og man kan få en antydning på om elevene tror at et utfall blir mer og mer sannsynlig dersom utfallet ikke inntreffer på en stund. Oppgaven hadde tre deloppgaver som jeg nå vil gjennomgå i kronologisk rekkefølge.

4.1.1 Oppgave 1a

1. Vi trekker tilfeldig ut kuler fra en krukke. I krukka har vi både hvite og sorte kuler. Etter at vi har trukket en gang, noterer vi ned fargen og legger kula oppi krukka igjen. Vi har trukket ut noen kuler, og skal til å trekke igjen.



Vi har trukket: H S H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

Resultatene fra oppgave 1a var følgende:

Tabell 4.1.1: Resultater fra oppgave 1a

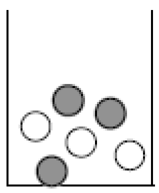
	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hvit	**14**	**14**	**58**	**15**	**101**	**71,6 %**
Sort	1	0	2	2	5	3,5 %
Like stor sjanse	9	3	9	14	35	24,8 %
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	69	31	141	100,0 %

Her ser vi at 101 av 141 elever svarte riktig, mens 40 elever svarte galt. Vi ser også at de tre teoriklassene var de som gjorde det best på denne oppgaven med over 84 % riktige svar. De tre praktiske klassene var de som hadde størst problemer med oppgaven, her var det under halvparten av elevene som svarte riktig.

De fleste som svarte galt har svart at det er like stor sjanse for begge fargene, faktisk hele 35 av 40 elever. 29 av disse 35 har begrunnet svaret sitt på en slik måte at det kan tyde på at de ikke har lest oppgaven nøyaktig nok og dermed har antatt at oppgaven var uten tilbakelegging. Et eksempel på dette kan vi blant annet lese ut fra følgende elevs begrunnelse: *"Det er like stor sjanse for begge for etter du har trukket H S H er det en hvit og en svart kule igjen og da er det like stor sjangse for å få hvit som svart"*. Av de resterende elevene som svarte feil kan det være verdt å nevne at det virker som om tre elever har ignorert forholdet mellom de sorte og hvite kulene og trodd at det var like stor sjanse så lenge det var kuler av begge farger i urnen.

4.1.2 Oppgave 1b

b



Vi har trukket: S S H S ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

Resultatene fra oppgave 1b var som følger:

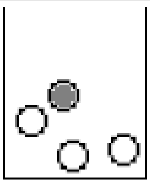
Tabell 4.1.2: Resultater fra oppgave 1b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hvit	10	3	10	18	41	29,1 %
Sort	1	1	0	1	3	2,1 %
Like stor sjanse	**13**	**13**	**58**	**10**	**94**	**66,7 %**
Ubesvart	0	0	1	2	3	2,1 %
Sum	24	17	69	31	141	100,0 %

Vi ser her at resultatene var omtrent som i oppgave 1a, men det var litt færre riktige svar. Det var 94 riktige og 44 gale svar, mens tre elever ikke hadde svart. Det var igjen de tre teoriklassene ved skole 40 som presterte det beste resultatet med 58 av 69 elever med riktig svar. De tre praktiske klassene ved skole 40 gjorde det dårligst, som i oppgave 1a, men på denne oppgaven var det under en tredjedel som svarte riktig.

Av de 44 gale svarene så var det 35 begrunnelser som kunne tyde på at elevene hadde trodd at uttrekningen var uten tilbakelegging. Av de resterende gale svarene er det interessant å se at det var bare tre elever som trodde uttrekningen ville jevne seg ut over tid og dermed svarte at hvit var mest sannsynlig fordi det var så mange sorte som allerede var blitt trukket ut. Et eksempel på dette kan vi se i denne forklaringen: *"Det er størst sjanse for at neste kule er hvit fordi dei skal bli trekt like mange gonger"*.

4.1.3 Oppgave 1c

c

Vi har trukket: H H H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

Resultatene fra oppgave 1c var som følger:

Tabell 4.1.3: Resultater fra oppgave 1c

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hvit	**14**	**13**	**55**	**10**	**92**	**65,2 %**
Sort	9	4	12	16	41	29,1 %
Like stor sjanse	1	0	1	3	5	3,5 %
Ubesvart	0	0	1	2	3	2,1 %
Sum	24	17	69	31	141	100,0 %

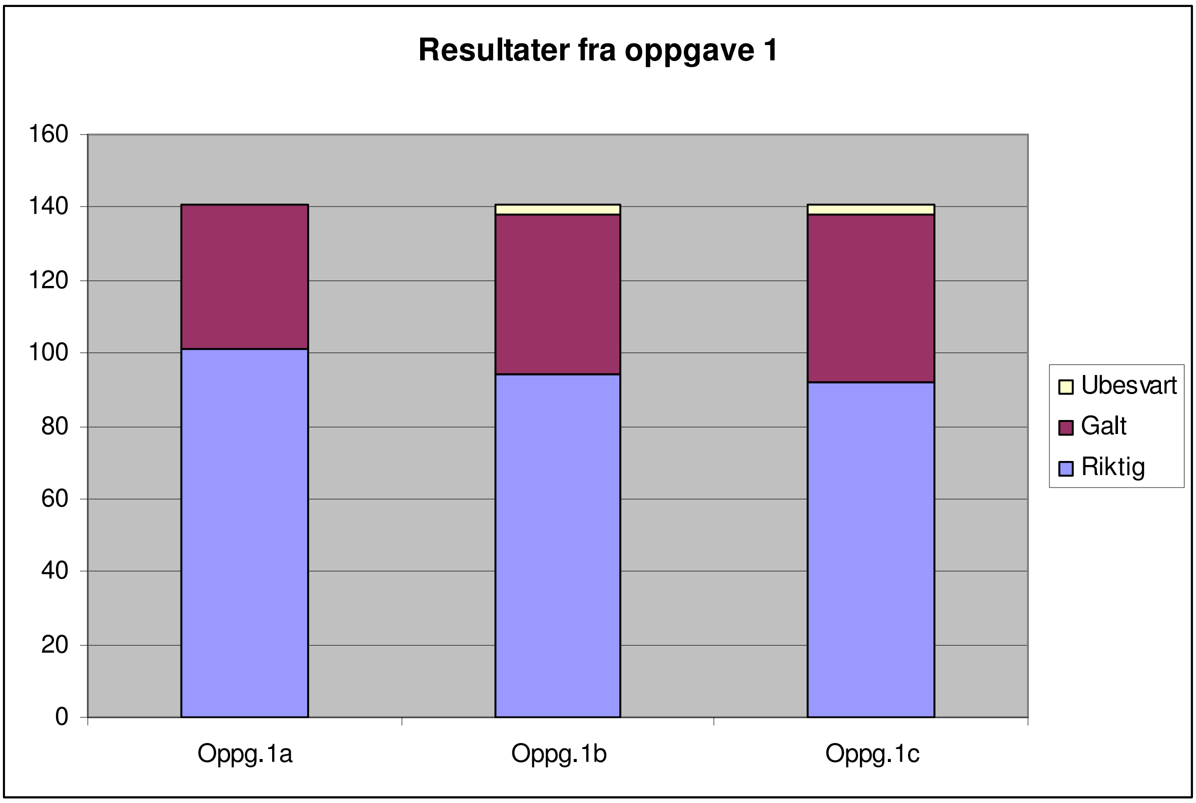
Vi ser den samme tendensen som fra de to andre deloppgavene med cirka to tredjedeler av elevene med riktig svar. Igjen er det de tre teoriklassene som kommer ut med det beste resultatet, men det kan også være verdt å merke seg at det er kun disse som har færre riktige svar enn i oppgave 1b.

Av de 46 som svarer galt virker det, ut fra begrunnelsene, som om 34 tror at oppgaven er uten tilbakelegging. Syv elever mener sort har størst sannsynlighet for å bli trukket ut fordi fordelingen vil jevne seg ut over tid. At det var denne deloppgaven som gav flest begrunnelser for at fordelingen vil jevne seg ut var ikke overraskende, da oppgaven er lagt opp til å teste akkurat denne oppfattelsen hos elevene ved at man bare har trukket ut en farge av de simulerte uttrekningene.

4.1.4 Oppsummering av oppgave 1

Oppgave 1 gav mange riktige svar, men det var uansett litt overraskende at det kun var cirka to tredjedeler som svarte riktig da oppgaven burde være relativt enkel for elever på videregående skole. Dersom vi går litt mer grundig inn på hvorfor så mange har svart galt kan det tyde på at de fleste ikke har lest oppgaven skikkelig og dermed har trodd at det ikke skulle være med tilbakelegging. Ut fra begrunnelsene til elevene virker det som om cirka tre fjerdedeler av de som har svart galt har gjort nettopp dette.

Dersom vi går inn på resultatene fra de tre deloppgavene kan vi se at det var en synkende tendens på antall riktig svar igjennom hele oppgave 1. Dette kan vi se mer tydelig i følgende graf:



Figur 4.1.4.1: Resultater fra oppgave 1.

Denne grafen viser oss at det var relativt mange riktige svar på alle de tre deloppgavene, men at det var færre riktige svar på deloppgave b enn a og at elevene scoret dårligst på deloppgave c. Vi ser også at det ikke var noen som ikke svarte på deloppgave a, mens det var noen få elever som ikke svarte på b og c.

Det kan være flere grunner til at det var en nedgang i antall riktige på oppgave 1. Det kan være en tilfeldighet. Da hver deloppgave hadde tre alternativer er det mulig at noen elever tippet når de besvarte oppgavene. En annen grunn kan være for eksempel at vanskelighetsgraden på oppgavene ble noe større, selv om jeg i utgangspunktet ville trodd at de tre deloppgavene var relativt like i vanskelighetsgrad.

4.2 Oppgave 2

Oppgave 2 var en myntoppgave som skulle teste ut elevenes forståelse av tilfeldighet. Oppgaven var som følger:

2. Du gjør seks tilfeldige myntkast. Hvilken av de følgende rekkefølgene av mynt (M) og kron (K) er mest sannsynlig eller er de like sannsynlige?

☐ MKMMKM er mest sannsynlig

☐ MMMMKK er mest sannsynlig

☐ KMKMKM er mest sannsynlig

☐ MKKMKM er mest sannsynlig

☐ Alle fire rekkefølgene er like sannsynlige

Hvorfor mener du dette?

___

De to første alternativene representerer fordelingen fire mynt og to kron der de to kronkastene er spredt eller samlet. De to siste alternativene representerer like mange kron og mynt ordnet i en rekke med mønster og en rekke uten mønster. Med tanke på resultater fra tidligere forskning skulle man på forhånd trodd at vi fikk få svar med de to første alternativene, mens det ville være en del flere elever som trodde at alternativene med tre korn og tre mynt var de mest sannsynlige da disse rekkene er mer representativ for den 50-50 prosentfordelingen som myntkast representerer. Resultatene fra denne oppgaven ble:

Tabell 4.2: Resultater fra oppgave 2

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
MKMMKM	0	2	0	1	3	1,9 %
MMMMKK	0	0	0	0	0	0,0 %
KMKMKM	4	4	24	3	35	22,3 %
MKKMKM	2	2	16	3	23	14,6 %
Like sannsynlig	**20**	**12**	**36**	**21**	**89**	**56,7 %**
Ubesvart	0	0	2	5	7	4,5 %
Sum	26	20	78	33	157	100,0 %

Vi ser her at over halvparten av elevene svarte korrekt på denne oppgaven. Det kan være interessant å merke seg at de elevene som slet mest med denne oppgaven var de som gikk i rene 1T-klasser. Vi ser også at det var nesten ingen elever som valgte alternativene med to kron og fire mynt, mens de fleste som hadde svart feil valgte ett av de to alternativene med tre kron og tre mynt.

Summen av antall svar er høyere enn antall elever, og grunnen til dette er i hovedsak at det var flere elever som både hadde krysset av på alternativet "KMKMKM" og alternativet "MKKMKM". Dette kan tyde på at disse elevene mente at så lenge det er jevnt fordelt med antall mynt og kron så er de ulike rekkene like sannsynlige. En elev som hadde gjort akkurat denne feilen hadde skrevet følgende kommentar til svaret sitt: *"Fordi det er lika stor sjangse for å få mynt som kron"*. Forklaringer av denne typen gikk igjen hos de aller fleste som hadde gjort denne feilen.

De elevene som kun har krysset av for alternativet "KMKMKM" kan man tenke seg at forventer å se et mønster av denne typen når det er snakk om utfall av et "fifty-fifty" fenomen. En av elevene som hadde krysset av for dette alternativet kom med følgende forklaring: *"Fordi mynten har bare to sider og då er det mest sannsynlig for at vi får KMKMKM"*. De fleste forklaringene på denne typen feil var av lignende karakter. Videre diskusjon av resultatene vil tas opp igjen i kapittel 5.

4.3 Oppgave 3

Oppgave 3 skulle, som oppgave 2, teste elevenes forståelse for tilfeldighet. Denne oppgaven handlet om lottorekker og var som følger:

3. Hvilken av følgende Lottorekker tror du har størst sjanse for å bli trukket ut som vinnerrekke?

☐ 1 2 3 4 5 6 7

☐ 7 13 14 19 25 31 34

☐ 3 6 9 15 27 30 33

☐ Alle har like stor sjanse

Hvorfor mener du dette?

____

Det første alternativet skulle her være en representant for det spesielle og derfor valgte jeg tallene 1-7. Det andre alternativet skulle representere den mer "vanlige", tilfeldige lottorekken med blandede, tilfeldige tall. Det tredje alternativet skulle være noe i mellom disse to, og jeg valgte da å ta en rekke med tall kun fra tregangen. Resultatene fra denne oppgaven var:

Tabell 4.3: Resultater fra oppgave 3

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
1-2-3-4-5-6-7	1	0	5	2	8	5,2 %
7-13-14-19-25-31-34	5	6	15	8	34	21,9 %
3-6-9-15-27-30-33	4	4	12	3	23	14,8 %
Like sannsynlig	**15**	**10**	**44**	**19**	**88**	**56,8 %**
Ubesvart	1	0	0	1	2	1,3 %
Sum	26	20	76	33	155	100,0 %

Vi kan se ut fra tabellen at det var over halvparten av elevene som svarte riktig, mens de fleste som svarte feil hadde krysset av på lottorekken med de blandede, tilfeldige tallene. Cirka 15 % av elevene mente at rekken med tall fra tregangen var den mest sannsynlige, mens det kun var 5 % som mente at rekken med tallene 1-7 var den rekken som hadde størst sannsynlighet for å bli trukket ut. Klassevis var det få store forskjeller på denne oppgaven.

De elevene som har svart alternativene med blandede tall eller tall fra tregangen har kanskje brukt representativitet som et hjelpemiddel for å svare på denne oppgaven. Grunnen til dette kan være at de tenker på hvilke rekker som pleier å bli trukket ut på lottotrekningen og bruker dette som en ledesnor i oppgaven. Da vil de nok ende opp med at det er mest sannsynlig at en blandet rekke blir trukket ut, da det som regel er en blandet rekke som faktisk blir trukket ut.

Ser vi på begrunnelsene til elevene støtter disse opp om teorien om at elevene har brukt representativitet når de har løst denne oppgaven. Noen av begrunnelsene fra elever som krysset av for de blandede tallene er:

- *For der er det størst variasjon mellom tallene.*
- *Fordi det er en blanding av oddetall og partal.*
- *Fordi tala blir aldri trukket ut i rekkefølge fra 1-7.*
- *For det er veldig liten sansynlighet for at 1234567 kommer etter hverandre.*
- *For de andre er så "like tall". Mest sannsynli er blandede tall som ikke har en spesiell rekke.*
- *Fordi sansynet for at me klarer å trekka tala 1, 2, 3, 4 osv er svært liten og usansynleg, skjølv om det kan henda.*

Dette er noen utvalgte begrunnelser fra både de som har valgt begge de to blandede rekkene og fra elever som har valgt én av de. Som vi ser er det to hovedgrunner til at elevene velger de to tallrekkene som representerer rekker uten et klart mønster: (1) disse representerer rekker som elevene forbinder med rekker som "pleier" å bli trukket ut i Lotto og (2) den første tallrekka (1-7) ser de på som veldig spesiell og usannsynlig og derfor velger de en eller begge de to andre rekkene.

Omtrent alle elevene som svarte riktig på denne oppgaven hadde også en god begrunnelse. De fleste skrev at det er helt tilfeldig hvilke tall som blir trukket ut og derfor har de ulike rekkene like stor sannsynlighet for å bli trukket. Det kan også nevnes at av de elevene som har svart rekken med tallene 1-7 kan det virke som de enten har tippet (ingen begrunnelse) eller at de har misforstått. Flere av disse elevene hadde begrunnet dette valget med at dette var rekken med færrest tall. Her kan vi lure på om de tenker at hvert siffer er et tall, og at de ikke registrerer at de er tosifrede tall i de to siste rekkene. Da kan det være at de tenker videre at dess færre siffer vi har å trekke ut, dess større sannsynlighet er det for å trekke ut alle sifrene.

4.4 Oppgave 4

Oppgave 4 var en oppgave som skulle teste ut elevenes forståelse for de store talls lov. Oppgaven var hentet fra Greens undersøkelse, men jeg hadde gjort noen små forandringer fra hans oppgavetekst, blant annet hadde jeg fjernet et av hans svaralternativer. Oppgaveteksten var som følger:

4. På et stort sykehus registrerer man alle barn som blir født et år. Følgende kan da hende:

A: Av de første 10 barna som blir født, er det 7 jenter.

B: Av alle de 100 nyfødte, er det 70 jenter.

☐ Hendelse A er mer sannsynlig enn hendelse B.

☐ Hendelse B er mer sannsynlig enn A.

☐ A og B er like sannsynlige.

Hvorfor mener du dette?

____

Green (1982) hadde som sagt denne oppgaven i sin undersøkelse. Resultatene han fikk var svært dårlige, og selv om han også hadde mye yngre elever med i undersøkelsen så var det faktisk de eldste elevene som endte opp med det dårligste resultatet. Dette tyder dermed på at det ikke var noen særlig forbedring med alderen. Resultatene fra denne oppgaven i min undersøkelse var:

Tabell 4.4: Resultater fra oppgave 4

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hendelse A	**5**	**7**	**24**	**12**	**48**	**34,0 %**
Hendelse B	3	3	5	4	15	10,6 %
Like sannsynlig	16	7	40	15	78	55,3 %
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	69	31	141	100,0 %

Det var som forventet mange elever som hadde problemer med denne oppgaven. Dette var som sagt tidligere ikke noen overraskelse, da dette er noe som vi også har sett i tidligere undersøkelser. Vi ser ut fra tabellen at det var cirka en tredjedel av elevene som svarte riktig på denne oppgaven, mens over halvparten av elevene svarte at de to hendelsene var like sannsynlige. Det var cirka 10 % som mente at hendelse B var mest sannsynlig. Elevene på skole 38 slet spesielt med denne oppgave. Kun 5 av 24 elever krysset av for riktig alternativ ved denne skolen, mens hele to tredjedeler krysset av for at de to hendelsene var like sannsynlige.

I forhold til resultatene fra Greens undersøkelse så gjorde disse norske elevene det klart bedre, selv om de ikke gjorde det veldig bra. Det kan være flere grunner til dette, blant annet at Greens elever var yngre, men Green påviste at forståelsen for denne oppgaven ikke bedret seg nevneverdig med alderen og derfor er det mer sannsynlig at den store forskjellen ligger i enten det faktum at Green hadde med et fjerde alternativ, ”No one can say” som hele 61 % av hans elever krysset av for eller at de norske elevene har utviklet en god forståelse for de store talls lov. Det var for øvrig kun 8 % av elevene i Greens undersøkelse som krysset av for det riktige svaret på denne oppgaven, mens det kun var 6 % av 16-åringene.

De fleste elevene som svarte at de to hendelsene var like sannsynlig har brukt begrunnelsen om at hendelse A og hendelse B gikk ut på det samme eller at forholdet var likt. Noen av begrunnelsen var:

- *Fordi* $\frac{7}{10}$ *er det samme som* $\frac{70}{100} = \frac{7}{10}$.
- *Fordi sju av ti kan vi skrive som 70 %. Og 70 % er det samme som den* (sic!) *andre tallet som er 70 av 100 = 70 %.*
- *Det er samme forholdet mellom 7 av 10 og 70 av 100.*
- $\frac{7}{10} = \frac{70}{100}$.
- *Fordi i A og B er det 70 % jenter.*

Elevene har her neglisjert de store talls lov, noe som også Shaughnessy og Bergman (1993) oppdaget i sin undersøkelse. Noen elever har skrevet at dersom vi dividerer tallene i hendelse B eller multipliserer tallene i hendelse A med ti, vil vi få akkurat det samme. Elevene tror at sannsynligheten for å få en viss prosentandel gutter eller jenter er lik uansett hvor stor populasjonen er.

Av de som har svart riktig på denne oppgaven kan det virke som om noen har brukt en årsakstenkning. Noen av begrunnelsene til elevene som har svart riktig er:

- *Fordi e har vanskli for å tro at 70 av 100 barn er jenter.*
- *Det vert fødd fleire jenter enn gutar, men 70 jenter av 100 er litt urealistisk.*
- *Det er på grunn av at det er statistisk umulig å ha 70 jenter av 100, mens 7 av 10 går bra.*
- *I utgangspunktet vil det vera ca. 50/50 gut og jente. Eg ser det som meir sannsynleg at 7 av 10 kontra 70 av 100 er jenter.*

Som vi ser er det ikke alle elevene som svarte riktig som har en begrunnelse som er helt riktig. Noen elever har som sagt brukt en årsakstenkning, mens andre har skrevet en helt korrekt begrunnelse, som *"Fordi det jevnar seg ut jo fleire ungar som vert fødd"*. Det er også en del elever som har latt være å begrunne svaret sitt på denne oppgaven. En av grunnene til dette kan nok være at de har tippet.

4.5 Oppgave 5

Oppgave 5 er i teorien svært lik oppgave 4. Denne oppgaven gikk også på de store talls lov og var som følger:

5. Gjennomsnittshøyden til alle rekrutter i Norge er et år 182 cm. Hva er da mest sannsynlig av følgende hendelser:

☐ At gjennomsnittshøyden til en liten tropp på 3 mann er over 185 cm

☐ At gjennomsnittshøyden til en tropp på 30 mann er over 185 cm

☐ Begge hendelsene er like sannsynlige

Hvorfor mener du dette?

____

Resultatene fra denne oppgaven var:

Tabell 4.5: Resultater fra oppgave 5

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
3 mann	**11**	**7**	**36**	**8**	**62**	**43,7 %**
30 mann	9	5	17	12	43	30,3 %
Like sannsynlig	4	6	15	10	35	24,6 %
Ubesvart	0	0	1	1	2	1,4 %
Sum	24	18	69	31	142	100,0 %

Som tabellen viser så var det relativt store forskjeller fra resultatene i oppgave 4. Det er blant annet interessant å se at det cirka 10 % flere elever som har svart riktig på denne oppgaven, mens under en fjerdedel av elevene har svart at de to hendelsene er like sannsynlige. I oppgave 4 krysset over halvparten av elevene av for dette alternativet. Vi ser også at det er atskillig flere elever som har svart at det er mest sannsynlig at gjennomsnittshøyden til en tropp på 30 mann er over 185 cm enn det er som har svart at hendelse B er mest sannsynlig i oppgave 4. I oppgave 4 var det cirka en av ti, mens det er cirka tre av ti i denne oppgaven. Disse relativt store forskjellene er interessante da oppgavene er såpass like.

Ut fra begrunnelsene til elevene kan det virke som om dette var en noe vanskelig oppgave. Det er relativt mange elever som ikke har skrevet noe begrunnelse eller rett og slett skrevet at de har tippet. Flere elever har brukt samme begrunnelse som de gjorde i oppgave 4. Dette gjelder både elever som har svart riktig på begge oppgavene og elever som har svart galt på begge. Jeg vil komme mer tilbake til denne oppgaven i diskusjonsdelen.

4.6 Oppgave 6

Oppgave 6 skulle teste elevenes forståelse for begrepet gjennomsnitt. Oppgaven hadde to deloppgaver:

6. Ole skal pakke fire egg som skal veie i gjennomsnitt 70 gram. Hva må det siste egget veie, dersom de tre første veier:

A: 70, 70 og 68 gram **Svar:** _____________

B: 69, 68 og 70 gram **Svar:** _____________

Forklar hvordan du tenkte:

A: ____

B: ____

4.6.1 Oppgave 6a

Resultatene fra deloppgave a var:

Tabell 4.6.1: Resultater fra oppgave 6a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
69 gram	3	2	3	4	12	8,5 %
70 gram	3	2	3	2	10	7,1 %
72 gram	**13**	**13**	**60**	**13**	**99**	**70,2 %**
Andre svar	2	0	0	4	6	4,3 %
Ubesvart	3	0	3	8	14	9,9 %
Sum	24	17	69	31	141	100,0 %

Vi ser at det var cirka 70 % av elevene som svarte riktig, mens elevene som svarte galt fordelte seg noenlunde jevnt på svarene 69 gram, 70 gram og andre svar. Vi ser også at det var cirka 10 % som valgte å ikke svare på denne oppgaven. Dette var en av oppgavene som flest elever ikke svarte på. En mulig grunn til at det var så mange som ikke svarte på oppgaven kan være at de var usikre på begrepet gjennomsnitt. Dersom vi ser på resultatene klassevis kan det være interessant å merke seg at over 25 % av elevene i de rene praktiske klassene ikke svarte på oppgaven, mens kun cirka 42 % svarte riktig.

Elevene gjorde det relativt bra på denne oppgaven. Spesielt resultatene til elevene fra teoriklassene tyder på at de har en god og robust forståelse for begrepet gjennomsnitt. Nesten 87 % av elevene i disse tre klassen hadde svart riktig. Det var flere elever som hadde svart, men ikke skrevet noe forklaring på hvordan de tenkte på denne oppgaven. Grunnen til dette kan være formuleringen i oppgaven. På de andre oppgavene står det "Hvorfor mener du dette?", mens det her står "Forklar hvordan du tenkte". Å forklare hvordan man tenker er ikke enkelt derfor kan det være at noen elever har ignorert dette spørsmålet.

4.6.2 Oppgave 6b

Resultatene fra oppgave 6b var:

Tabell 4.6.2: Resultater fra oppgave 6b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
70 gram	6	3	5	5	19	13,5 %
73 gram	**13**	**14**	**57**	**15**	**99**	**70,2 %**
Andre svar	2	0	5	2	9	6,4 %
Ubesvart	3	0	2	9	14	9,9 %
Sum	24	17	69	31	141	100,0 %

Vi ser at det var en identisk fordeling i antall elever som har svart enten riktig, galt eller ubesvart som det var i oppgave 6a. Det resultatet kommer ikke som noen overraskelse da det eneste som er endret fra oppgave 6a er et av tallene på de forhåndsveide eggene.

Det var som nevnt hele 10 % av elevene som ikke svarte på disse to oppgavene. Det var bare en oppgave i testen som hadde en høyere prosentandel av ubesvarte. En av forklaringene på hvorfor 10 % av elevene ikke har svart på oppgave 6 kan være at dette ikke er en avkrysningsoppgave med flere alternativer. Dette er en oppgave der de selv må skrive sitt forslag til svar og det kan ha ført til at noen elever ikke har svart.

4.7 Oppgave 7

Oppgave 7 var en sannsynlighetsoppgave med to deloppgaver, der deloppgave a var hentet fra Greens undersøkelse, mens deloppgave b er hentet fra Breiteigs sannsynlighetstest. Svaralternativene er noe omgjort fra hvordan de presenterte disse oppgavene.

4.7.1 Oppgave 7a

7. I en klasse er det 13 gutter og 16 jenter. Klassen har fått to gratisbilletter til en konsert. Klassen vil trekke lodd om billettene. Hver elevs navn skrives på en lapp. Lappene legges så oppi en boks, og læreren trekker ut to lapper uten å se.

a)

☐ Det er størst sjanse for at den første som trekkes ut er en gutt

☐ Det er størst sjanse for at den første som trekkes ut er ei jente

Hvorfor mener du dette?

____

Green (1982) hadde også denne oppgaven i sin undersøkelse. Resultatene han fikk var at elevene gjorde det relativt bra på denne oppgaven. Over halvparten av elevene i hans undersøkelse svarte riktig, og hele 71 % av 16-åringene. Resultatene mine fra denne oppgaven var som følger:

Tabell 4.7.1: Resultater fra oppgave 7a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Gutt	1	1	2	0	4	2,8 %
Jente	**22**	**16**	**66**	**31**	**135**	**95,7 %**
Ubesvart	1	0	1	0	2	1,4 %
Sum	24	17	69	31	141	100,0 %

Som vi ser ut i fra tabellen så var dette en relativt enkel oppgave for elevene. Hele 135 av 141 elever svarte riktig på oppgaven og kun fire elever svarte galt, mens to elever ikke svarte på oppgaven. Resultatet er mye bedre enn det elevene i Greens undersøkelse oppnådde, men igjen må elevenes alder tas i betraktning. I tillegg må det nevnes at Green hadde fire svaralternativer i stedet for to.

Begrunnelsene fra denne oppgaven viste også at dette var en oppgave som elevene hadde god kontroll på. De aller fleste hadde begrunnet svaret med at det var flere jenter enn gutter i klassen eller med at sannsynligheten for å trekke ut en jente er 16/29 som er større enn sannsynligheten for å trekke en gutt som er 13/29.

4.7.2 Oppgave 7b

b)

☐ Det er størst sjanse for at de to som trekkes ut er en gutt og ei jente

☐ Det er størst sjanse for at de to som trekkes ut er to jenter

☐ Det er størst sjanse for at de to som trekkes ut er to gutter

Hvorfor mener du dette?

____

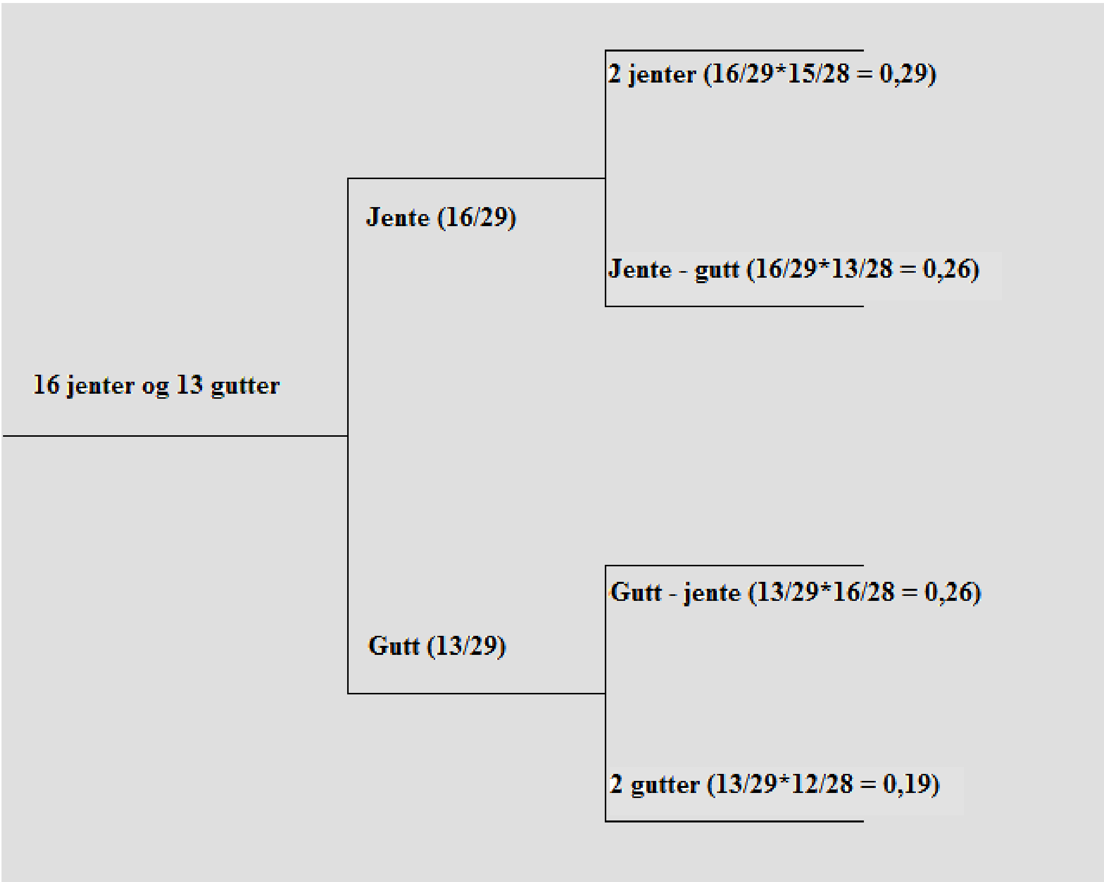
Resultatene fra oppgave 7b ble som følger:

Tabell 4.7.2: Resultater fra oppgave 7b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Gutt og jente	**5**	**1**	**15**	**7**	**28**	**19,7 %**
To jenter	17	16	52	24	109	76,8 %
To gutter	1	0	2	0	3	2,1 %
Ubesvart	1	0	1	0	2	1,4 %
Sum	24	17	70	31	142	100,0 %

Denne oppgaven viste seg, i motsetning til oppgave 7a, å være svært problematisk for elevene. Vi ser at det var under 20 % av elevene som svarte riktig, og spesielt på skole 39 hadde elevene store problemer, der kun en av 17 elever svarte riktig. Omtrent alle gale svar var at det var mest sannsynlig å trekke ut to jenter, hele 109 av 114 gale svar var av denne typen.

De fleste begrunnelsene til elevene som har svart galt er at det er flere jenter enn gutter i klassen. Ut fra dette er det vanskelig å si noe om hvordan elevene har tenkt, men jeg har satt opp to mulige feil som elevene kanskje gjør. Den første grunnen til at det var så mange som gjorde feil på denne oppgaven kan være at elevene ikke har tatt hensyn til rekkefølgen. Dermed har de tenkt at det er tre utfall, mens det er fire om vi tar hensyn til rekkefølgen, da kan gutt og jente skje på to forskjellige måter. Man kan trekke gutt først og deretter jente eller man kan trekke jente først og deretter gutt. Jeg har laget et sannsynlighetstre som skal illustrere dette:



Figur 4.7.2.1: Sannsynlighetstre til oppgave 7b.

Dette treet kan vi se på som en fasit til denne oppgaven. Først har vi hele klassen på 29 elever, 16 jenter og 13 gutter. Ved første trekning kan vi få to mulige utfall, en jente (sannsynligheten er 16/29) eller en gutt (sannsynligheten er 13/29). Dersom vi trekker en jente i den første trekningen har vi nå to nye utfall som kan inntreffe på den andre trekningen. Vi kan trekke en ny jente (sannsynligheten er 15/28) eller vi kan trekke en gutt (sannsynligheten er 13/28). Dersom vi trekker en gutt i den første trekningen har vi også to mulige utfall i den andre trekningen. Vi kan trekke en jente (sannsynligheten er 16/28) eller vi kan trekke en ny gutt (sannsynligheten er 12/28). Dermed ser vi at sannsynligheten for å trekke to jenter er 0,29, sannsynligheten for å trekke to gutter er 0,19, mens sannsynligheten for å trekke en jente og en gutt er 0,26*2 = 0,52.

En annen grunn til at elevene har hatt så store problemer med denne oppgaven kan være at elevene har blandet inn betinget sannsynlighet. Ut fra noen av begrunnelsene virker det som om noen elever misforstår oppgaven og tror at spørsmålet gjelder kun den andre trekningen, uavhengig av resultatet i den første trekningen eller eventuelt gitt at det er blitt jente i den første trekningen.

Jeg synes som sagt at det er vanskelig å si noe om begrunnelsene på denne oppgaven. De fleste begrunnelsene som er gitt kan tolkes slik at de passer innenfor begge disse feiltypene. Noen av begrunnelsene for at det er størst sannsynlighet for å trekke to jenter er:

- *Fordi det er framleis fleire jenter.*
- *Fordi det enda er flere jenter enn gutter når en jente er trukket.*
- *Fordi det er flest lapper med jentenavn på oppi boksen.*
- *For hvis det blir trekt ut en jente først, er det fremdeles flere jenter.*

Tilslutt vil jeg ta med begrunnelsen til en elev som hadde svart at det er størst sjanse for at de to som trekkes ut er en gut og en jente, altså riktig, på denne oppgaven: *"Fordi det er sånn det blir gjort på skoler."*

4.8 Oppgave 8

Oppgave 8 var en oppgave som skulle teste elevenes forståelse loven om konjunksjon. Oppgaven var i sin helhet hentet fra Breiteig…:

8. En kvinne fyller 80 år. Hun er sprek og frisk, og daglig går hun turer og svømmer.

Hvilken av følgende hendelser er mest sannsynlig:

☐ Hun blir over 85 år.

☐ Hun blir over 90 år.

☐ Hun blir over 100 år.

Hvorfor mener du dette?

____

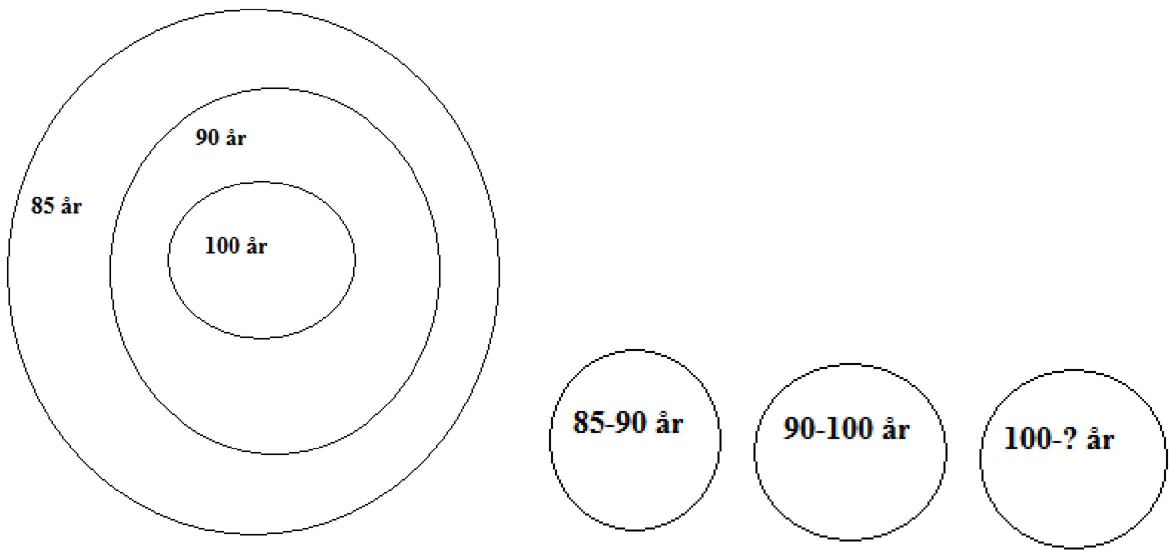
På denne oppgaven vil det være umulig å si hvor gammel denne kvinnen vil bli, men det jeg ville teste elevene i var om de forstod at svaralternativ en "Hun blir over 85 år" også inkluderer de to andre svaralternativene. Resultatene fra oppgave 8 ble som følger:

Tabell 4.8: Resultater fra oppgave 8

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Over 85 år	**8**	**4**	**29**	**5**	**46**	**30,9 %**
Over 90 år	10	12	31	23	76	51,0 %
Over 100 år	6	1	11	7	25	16,8 %
Ubesvart	1	0	1	0	2	1,3 %
Sum	25	17	72	35	149	100,0 %

Som vi ser i tabellen så svarte kun litt over 30 % av elevene riktig. Over halvparten av elevene mente at det var mest sannsynlig at hun ville bli over 90 år, mens rundt 17 % mente at det var mest sannsynlig at hun ble over 100 år. 2 elever valgte å ikke svare på oppgaven. Klassevis er det relativt små forskjeller. Skole 39 og de praktiske klassene ved skole 40 sliter litt mer en de andre.

De fleste begrunnelsene som ble brukt for de to gale svar alternativene gikk på at kvinnen holdt seg i god form og var sunn og frisk og derfor var det mest sannsynlig at hun ble over 90 eller 100 år. Noen av begrunnelsene kunne også tyde på at en del elever så på svaralternativene som intervaller. Med dette mener jeg at de trodde at det første alternativet var at kvinnen ble mellom 85 og 90 år, det andre alternativet var mellom 90 og 100 år, mens det siste var at hun ble over 100 år (se figur 4.8.1).



Figur 4.8.1: Svaralternativene i riktig form og i intervallform.

For elevene som tror det er intervaller kan svaralternativene illustreres som de tre sirklene i figuren til høyre i figur 4.8.1. Grunnen til at noen elever tror det er intervaller kan være i oppgaveteksten. De tre svaralternativene er listet opp en og en under hverandre noe som kan føre til at enkelte elever tror at det er snakk om fra 85 år til 90 år, fra 90 år til 100 år og 100 år og oppover. Ser vi på figuren til venstre i figur 4.8.1 ser vi at over 85 år også inkluderer 90 og 100 år. Det er slik løsningen var tenkt fra min side.

Resultatene fra denne oppgaven støtter det Kahnemann og Tversky (1983) fant i sin undersøkelse. De testet mer enn 3000 personer i deres forståelse for konjunksjonsloven og de fant ut at folk hadde store problemer med å forstå denne loven. Som jeg nevnte i teoridelen testet de blant annet collegestudenter uten noen spesiell erfaring innen statistikk og sannsynlighet. Disse mente blant annet at prosentandelen av folk som var over 55 år og hadde hatt hjerteattakk som større enn prosentandelen av folk som var over 55 år. På noen oppgaver var feilprosenten oppe i cirka 85 %, og da var til og med studenter som tok utdannelse innen statistikk og sannsynlighet med.

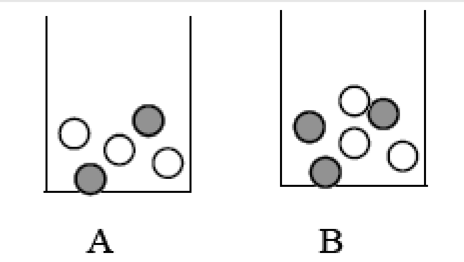
4.9 Oppgave 9

Oppgave 9 var en urneoppgave. Oppgaven skulle teste elevenes forståelse for begrepet forhold. Ideen til oppgaven er hentet fra Greens undersøkelse. Oppgaven hadde tre deloppgaver.

4.9.1 Oppgave 9a

9. På et tivoli skal Petter trekke ei kule fra en boks uten å se. Boksen inneholder et visst antall hvite og sorte kuler, ***og han får en premie dersom han trekker en hvit kule***. På forhånd får Petter velge mellom to bokser, der han får vite antallet hvite og sorte kuler i hver boks.

a) Hvilken boks er det lurest å trekke fra, A eller B, eller er det det samme?



Svar: ____________________________

Hvorfor mener du dette?

____

Resultatene fra oppgave 9a var som følger:

Tabell 4.9.1: Resultater fra oppgave 9a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Boks A	**22**	**17**	**67**	**28**	**134**	**95,0 %**
Boks B	0	0	2	1	3	2,1 %
Det samme	1	0	0	2	3	2,1 %
Ubesvart	1	0	0	0	1	0,7 %
Sum	24	17	69	31	141	100,0 %

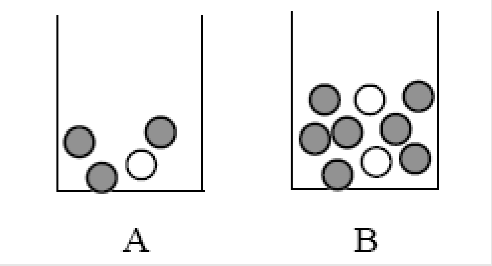
Denne oppgaven hadde elevene få problemer med. Vi kan se i tabellen at hele 95 % av elevene svarte riktig på denne oppgaven, mens kun seks elever svarte galt. Disse seks fordelte seg for øvrig likt mellom de to andre svaralternativene.

Green hadde også en lignende oppgave som dette. Han hadde forholdene 5/7 og 5/8, mens jeg altså har 3/5 og 3/6. Resultatene hans var at 67 % av elevene svarte riktig, mens 25 % svarte at det var det samme hvilken boks man trakk fra. Men Greens undersøkelse viste at resultatene bedret seg med alderen, og derfor er det mest naturlig å sammenligne med de eldste fra Greens undersøkelse. Av 16-åringene svarte 81 % riktig, mens 15 % svarte at det var det samme hvilken boks man trakk fra. Sammenligner vi med resultatene fra min test, selv om dette blir litt feil da oppgavene ikke er helt identiske, så kan det tyde på at forholdsbegrepet var noe bedre forstått av de norske elevene.

Begrunnelsene til elevene støtter teorien om at dette var en oppgave som elevene hadde god kontroll over. De fleste har begrunnet svaret sitt med enten at boks A har færre sorte kuler enn boks B eller at boks A har flere hvite enn sorte kuler. Noen elever har brukt begrunnelsen at i boks A er sannsynligheten for å trekke en hvit kule 3/5 som er større enn sannsynligheten for å trekke en hvit kule i boks B som er 3/6.

4.9.2 Oppgave 9b

b) Hvilken boks er det lurest å trekke fra, A eller B, eller er det det samme?



Svar: ____________________________

Hvorfor mener du dette?

____

Resultatene fra oppgave 9b var som følger:

Tabell 4.9.2: Resultater fra oppgave 9b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Boks A	**19**	**15**	**62**	**23**	**119**	**84,4 %**
Boks B	1	1	0	5	7	5,0 %
Det samme	3	1	7	3	14	9,9 %
Ubesvart	1	0	0	0	1	0,7 %
Sum	24	17	69	31	141	100,0 %

Vi kan se ut fra tabellen at oppgaven gikk greit for de fleste elevene. Det var nesten 85 % som svarte riktig, noe som er veldig bra, men det er likevel en nedgang på over 10 % fra oppgave

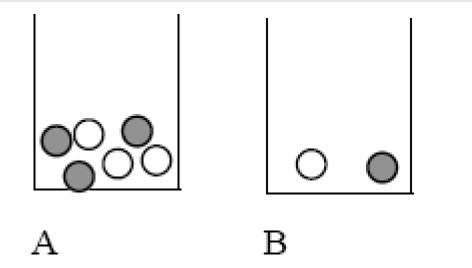
9a. Vi kan også se at av de som svarte galt svarte to tredjedeler at det var det samme hvilken boks man trakk fra.

Grunnen til at det er en nedgang i antall riktige svar på denne oppgaven i forhold til forrige kan være at oppgaven er noe vanskeligere. På deloppgave a hadde vi forholdene 3/5 og 3/6, mens vi her har forholdene 1/4 og 2/9. Dette vil si at mens vi hadde en forskjell i sannsynligheten på 0,1 i deloppgave a, er forskjellen på under 0,03 i denne oppgaven. Det faktum at det er såpass liten forskjell mellom de to forholdene kan også være grunnen til at flesteparten av de som gjør feil svarer at det er samme hvilken boks man trekker fra.

Ser vi på begrunnelsene til elevene har de fleste som har svart riktig begrunnet dette med enten at det er flere sorte kuler i boks B eller de har på en eller annen måte brukt forholdet mellom de sorte og hvite kulene. De som har brukt den siste varianten har sannsynligvis en god forståelse av forholdsbegrepet. De som har brukt den andre varianten er det litt vanskeligere å tolke noe spesielt. Det kan være at de mener at det er færre sorte kuler i forhold til hvite, men det kan også være at de mener at det er større sannsynlighet for å trekke en hvit dess færre andre kuler det er. Hadde de for eksempel fortsatt svart boks A dersom det var en eller to kuler mindre i boks B? Ut fra resultatene i deloppgave a og c, virker det som om de fleste elevene mener i forhold til de hvite kulene når de skriver begrunnelser som ”det er flere sorte kuler i boks B”.

4.9.3 Oppgave 9c

c) Hvilken boks er det lurest å trekke fra, A eller B, eller er det det samme?



Svar: ___________________________

Hvorfor mener du dette?

____

Resultatene ble som følger:

Tabell 4.9.3: Resultater fra oppgave 9c

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Boks A	0	0	1	4	5	3,5 %
Boks B	1	0	5	6	12	8,5 %
Det samme	**22**	**17**	**63**	**21**	**123**	**87,2 %**
Ubesvart	1	0	0	0	1	0,7 %
Sum	24	17	69	31	141	100,0 %

Vi ser i tabellen at det var litt flere elever som svarte riktig på denne oppgaven i forhold til deloppgave b, mens det var litt færre enn i deloppgave a. Med over 87 % riktige svar ble dette en oppgave elevene mestret relativt greit. Av de elevene som svarte galt svarte 12 av 17 at det var boks B som var den lureste boksen å trekke fra dersom man ville ha en hvit kule. Klassevis er det de rene praktiske klassene som sliter mest. Av disse elevene var det cirka to tredjedeler som svarte riktig. Elevene på skole 39 gjorde det igjen veldig bra. Som i deloppgave a svarte alle elevene riktig.

I denne oppgaven hadde vi altså det samme forholdet i de to boksene, men ulikt antall. Grunnen til at over 70 % av de som svarte feil mente at det var lurest å trekke fra boks B kan være at de tror det er lettere å trekke ut det gunstige dess færre kuler det er i boksen. Denne påstanden blir styrket av følgende elevs begrunnelse: *"Det er lettere å trekke en hvit kugle i B, da der er færre kugler i B end i A"*. Denne type begrunnelse går også igjen hos flere av elevene som mener at det er lurest å trekke fra boks B.

Begrunnelsene til elevene som har svart riktig tyder på at de har en god forståelse for forholdsbegrepet. Noen av begrunnelsene er:

- *Forholdet er likt mellom de to fargene i begge krukkene.*
- *Fordi sjansen for å trekke en hvit er 50/50 i begge boksene.*
- *Fordi forholdet mellom kulene i A og B er like stort.*
- *Det er like mange sorte som hvite kuler i begge boksene.*

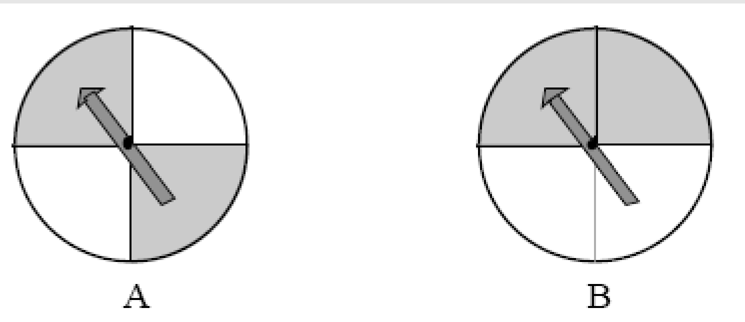
Som vi ser ut fra disse begrunnelsene viser elevene tydelig at de skjønner forholdsbegrepet, i alle fall i samband med denne oppgaven.

4.10 Oppgave 10

Oppgave 10 var en oppgave man kan kalle en "lykkehjulsoppgave" eller en oppgave om det berømte "Wheel of fortune". Oppgaven var hentet i sin helhet fra Breiteigs sannsynlighetstest, men Green hadde også noen lignende oppgaver i sin undersøkelse. Denne oppgaven har to deloppgaver.

4.10.1 Oppgave 10a

10. På et tivoli er det flere lykkehjul. En viser blir snurret rundt. Den stopper på et tilfeldig sted. ***Dersom den stopper på et hvitt felt vinner spilleren en premie.***



Hvilket hjul gir størst vinnersjanse?

☐ A

☐ B

☐ Samme vinnersjanse på begge hjul

Hvorfor mener du dette?

____

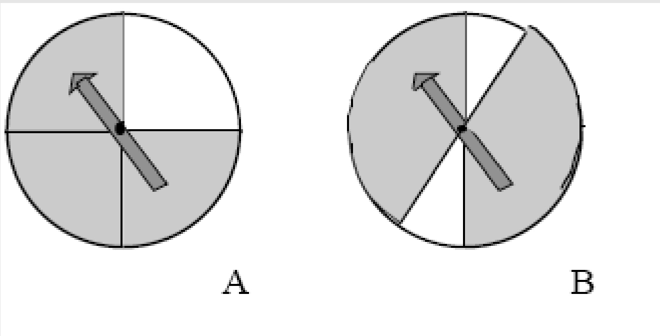
Resultatene fra oppgave 10a var:

Tabell 4.10.1: Resultater fra oppgave 10a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hjul A	4	1	3	6	14	9,9 %
Hjul B	3	2	6	6	17	12,0 %
Det samme	**17**	**14**	**60**	**20**	**111**	**78,2 %**
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	69	32	142	100,0 %

Som vi ser i tabellen så svarte nesten 80 % av elevene riktig. Det var 14 elever som mente at hjul A var det hjulet som gav størst vinnersjanse, mens 17 elever mente at hjul B gav størst vinnersjanse. Klassevis var det ikke de store forskjellene, de rene teoriklassene presterte best, mens de praktiske klassene fikk det dårligste resultatet.
De fleste elevene som mener at hjul A er det lykkehjulet som gir størst sannsynlighet for gevinst har gitt som begrunnelse at her er de hvite feltene mer spredt enn i hjul B. En elev som krysset av for hjul A skrev følgende begrunnelse: *”Det er bedre å ha det kvite litt spreidd”*. Elevene som mener at hjul B gir størst vinnersjanse har brukt akkurat samme begrunnelse, men bare motsatt. Et eksempel på dette er følgende begrunnelse fra en elev som mente hjul B gav størst vinnersjanse: *”Det er større sammenhengende vinnerareal”*.

4.10.2 Oppgave 10b



Hvilket hjul gir størst vinnersjanse?

☐ A

☐ B

☐ Samme vinnersjanse på begge hjul

Hvorfor mener du dette?

___

Resultatene fra oppgave 10b ble som følger:

Tabell 4.10.2: Resultater fra oppgave 10b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hjul A	**12**	**11**	**34**	**17**	**74**	**52,5 %**
Hjul B	4	3	3	5	15	10,6 %
Det samme	8	3	32	9	52	36,9 %
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	69	31	141	100,0 %

Det var en klar nedgang i antall riktige svar i forhold til deloppgave a. Kun litt over halvparten av elevene svarte riktig. Av de gale svarene var det en klar overvekt av elever som svarte at det var samme vinnersjanse på de to lykkehjulene, hele 52 av 67 elever. De resterende 15 elevene mente at det var hjul B som gav størst vinnersjanse. Klassevis er det interessant å se at nesten halvparten av elevene i det rene teoriklassene mente at det var det samme hvilket hjul man spilte på.

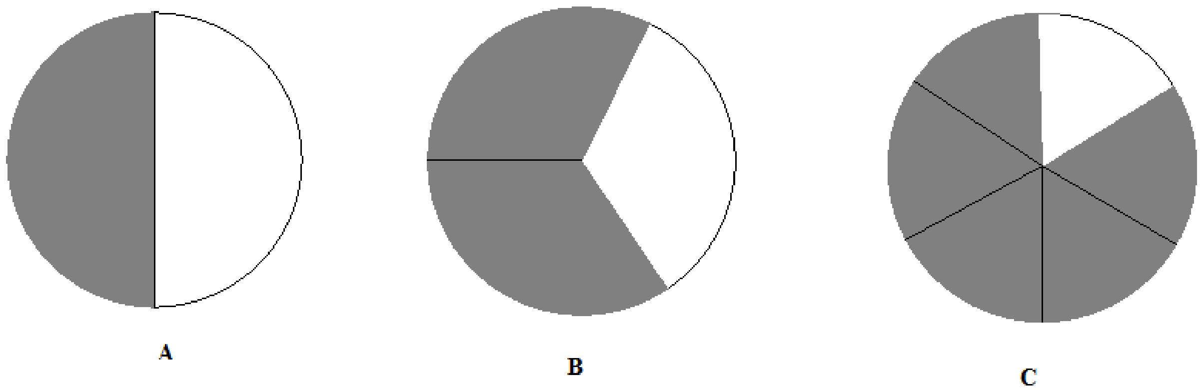
Det var relativt mange elever som mente at det var det samme hvilket hjul man spilte på. Grunnen til at nesten 37 % av elevene mente at det var lik vinnersjanse på de to lykkehjulene kan være at de hvite arealene ble laget litt for like i illustrasjonen til oppgaven. Det er mulig at noen elever trodde at de hvite feltene på hjul B var 1/4. Dette kommer også frem i noen av begrunnelsene til elevene som har krysset av for at sannsynligheten er den samme:

- *Begge lykkehjulene har like mye hvit.*
- *Det er 75 % for sort og 25 % for hvit på begge.*
- *Dei to kvite felta i b blir til saman lika stort som feltet i A.*
- *Hvis vi slår sammen de to hvite på B så blir det like stor som den hvite på A.*

4.11 Oppgave 11

Oppgave 11 var også en lykkehjulsoppgave. Ideen til oppgaven ble hentet fra Breiteigs sannsynlighetstest og har to deloppgaver. Jeg viser begge deloppgavene med en gang her da illustrasjonen i oppgaveteksten hører med til begge:

11. På et tivoli er det tre lykkehjul. En viser blir snurret rundt. Den stopper på et tilfeldig sted. ***Dersom den stopper på et hvitt felt vinner spilleren en premie.*** På lykkehjul A kan man vinne 180 kroner, på B er premien 330 kroner, mens man kan vinne 600 kroner på C.



Hvilket hjul vil du spille på, hvis

a) du bare kan spille en gang?

☐ A

☐ B

☐ C

Hvorfor? ____________________

b) du får lov til å spille 100 ganger?

☐ A

☐ B

☐ C

Hvorfor? ____________________

På denne oppgaven er det vanskelig å si at noe er rett eller galt da det jeg spør om egentlig er hva hver enkelt elev ville ha spilt på om han fikk sjansen. En elev som for eksempel velger hjul A på grunn av at dette gir størst vinnersjanse har jo rett til å mene dette. Jeg ville kanskje også selv ha valgt hjul A dersom jeg fikk sjansen til å spille én gang for å "sikre" at jeg fikk en gevinst. Men jeg har uansett satt hjul B som det riktige svaret i tabellene under da det er dette hjulet som gir mest i forventet tilbakebetaling av de tre hjulene.

4.11.1 Oppgave 11a

Resultatene fra oppgave 11a var som følger:

Tabell 4.11.1: Resultater fra oppgave 11a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hjul A	20	10	38	19	87	61,3 %
Hjul B	**2**	**7**	**21**	**7**	**37**	**26,1 %**
Hjul C	2	0	11	5	18	12,7 %
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	70	31	142	100,0 %

Vi kan se i tabellen at over 61 % av elevene ville valgt å spille med hjul A dersom de fikk sjansen til å spille én gang. Videre ville cirka 26 % valgt hjul B, mens cirka 13 % ville valgt hjul B. Det kan også være verdt å merke seg at 20 av 24 elever, over 83 %, på skole 38 ville valgt hjul A. På skole 39 er det ingen av elevene som ville valgt hjul C, mens faktisk over 40 % ville valgt hjul B.

Grunnen til at så mange elever ville valgt hjul A er bortimot utelukkende at sannsynligheten er størst for å vinne her. De fleste begrunnelsene for valget av hjul A er av liknende sort som denne eleven: *"fordi der er det størst sjangse for å vinne ettersom det kvite feltet er like stort som det svarte"*. Dette tyder på at flesteparten av elevene ville heller gått etter en "sikker gevinst" enn en godt betalt gevinst dersom de fikk muligheten til å spille kun én gang.

De fleste elevene som har valgt hjul B har på en eller annen måte skrevet i begrunnelsen at det er her den forventede tilbakebetalingen er størst. De fleste elevene som har valgt hjul C har valgt dette fordi dette gir den høyeste enkeltpremien av de tre hjulene og de er villige til å satse litt, som følgende elev gir uttrykk for: *"Logikken seier til A, men ein må jo satse litt, eller kva?"*.

4.11.2 Oppgave 11b

Resultatene fra oppgave 11b ble som følger:

Tabell 4.11.2: Resultater fra oppgave 11b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hjul A	6	5	21	9	41	27,5 %
Hjul B	**5**	**4**	**18**	**1**	**28**	**18,8 %**
Hjul C	16	11	32	21	80	53,7 %
Ubesvart	0	0	0	0	0	0,0 %
Sum	27	20	71	31	149	100,0 %

Vi kan se at nesten 19 % av elevene valgte hjul B, noe som er en nedgang på over 7 prosentpoeng i forhold til i oppgave 11a. Over halvparten av elevene krysset av for hjul C, mens 27,5 % hadde krysset av for A. Vi kan også se at summen av antall svar er noe høy på denne oppgaven, dette skyldes at noen av elevene hadde krysset av for flere av alternativene med begrunnelsen at dersom de fikk muligheten til å spille 100 ganger ville de ha prøvd noen ganger på alle de tre hjulene.

Det var litt overraskende at det var færre elever som valgte hjul B på denne oppgaven enn det var i oppgave 11a da det er her vi får store forskjeller i de forventede tilbakebetalingene til de tre forskjellige hjulene. Hjul A ville gitt en forventet tilbakebetaling på 9000 kroner etter 100 spill, hjul B ville gitt 11 000 kroner på 100 spill, mens hjul C ville gitt en forventet avkastning på 10 000 kroner i løpet av 100 spill.

De fleste begrunnelsene for å velge hjul A på denne oppgaven var som i 10a at man hadde størst sannsynlighet for å vinne eller at man kom til å vinne flest ganger. Noen av begrunnelsene for å velge hjul B var:

- *Kjangsen for suksess er ganske stor, samtidig som at premien er bra.*
- *Fordi der ville eg i teorien vunne 9000 kr mens på B 11000 kr og C 10000 kr.*

- *Vinner mest sannsynleg 33 gonger og då vert da mest pengar.*

Her ser vi at noen har regnet på den forventede tilbakebetalingen etter 100 spill og kommet frem til at hjul B gir best avkastning. Andre elever hadde tydeligvis tippet og krysset av for den i midten eller den som hørtes best ut, som den første begrunnelsen er et bevis på. Over halvparten av elevene hadde som nevnt valgt hjul C og noen av begrunnelsene var:

- *Fordi du kan vinne mest der.*
- *Sjølv om du vinn fleire gongar på A, blir den sammenlagte gevinsten størst i C.*
- *Større sum, viss eg spelar 100 gonger aukar sjangsen for å treffe det kvite feltet.*

Her ser vi at elevene mener at premien vil bli størst dersom man spiller på hjul C. Detter tyder på at de kun har sett på premien man får ved én gevinst og ikke regnet ut den forventede premien etter 100 spill.

4.12 Oppgave 12

Oppgave 12 var en litt annerledes myntoppgave. Oppgaven er utviklet av Skaar og Syvertsen (2000).

4.12.1 Oppgave 12a

12. Ola og Kari spiller med fire like mynter, som er merket med tallene 1, 2, 3 og 4. De putter myntene i en krukke og trekker ut to av dem uten å se. Ola vinner hvis summen av de to tallene er et partall, mens Kari vinner hvis summen er et oddetall.

Er dette spillet rettferdig?

☐ Ja

☐ Nei

☐ Vet ikke

Hvorfor mener du dette?

____

I denne oppgaven er det med en forstyrrende faktor som kan være med på å avlede elevene fra det riktige svaret. Denne forstyrrende faktoren er de to odde- og partallene i oppgaveteksten som kan føre til at elevene tenker at dette er et rettferdig spill dersom de ikke leser teksten nøye nok.

Resultatene fra oppgave 12a var:

Tabell 4.12.1: Resultater fra oppgave 12a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Ja	12	8	17	15	52	36,9 %
Nei	**11**	**7**	**47**	**12**	**77**	**54,6 %**
Vet ikke	1	2	5	3	11	7,8 %
Ubesvart	0	0	0	1	1	0,7 %
Sum	24	17	69	31	141	100,0 %

Som vi ser i tabellen svarte cirka 55 % av elevene riktig på denne oppgaven, mens cirka 37 % svarte at de mente myntspillet var rettferdig. 11 elever svarte at de ikke visste om spillet var rettferdig eller ikke, mens en elev ikke svarte i hele tatt på oppgaven. Klassevis var det de rene teoriklassene som gjorde det best med nesten 70 % riktige svar, mens de praktiske klassen fant seg i andre enden med under 40 % riktige svar.

Resultatene fra denne oppgaven stemmer noenlunde bra med hva Skaar og Syvertsen (2000) erfarte. Deres erfaring er at det er omtrent like mange som mener partall vinner som oddetall. Nå er mitt spørsmål noe annerledes enn dette, men vi ser at det er virka halvparten som mener at spillet ikke er rettferdig.

Grunnen til at over en tredjedel av elevene mener at spillet er rettferdig kan være at den nevnte "distractoren" har lurt de. Da har de sannsynligvis ikke lest oppgaven skikkelig og tror at man trekker ut myntene uten å legge sammen tallene. En av elevene som mente at oppgaven var rettferdig skrev denne forklaringen: *"Det er jo like stor sjanse for å få partall som oddetall når det er to av hver"*.

På den andre siden var det også noen elever som tydeligvis hadde lest oppgaven skikkelig, men fremdeles mente at spillet var rettferdig. En elev ved skole 38 som mente at spillet var rettferdig hadde skrevet følgende begrunnelse: *"For halvparten av myntene er partall og den andre halvparten er oddetall, og da burde summene bli noenlunde slik også"*. Her kan det tyde på at eleven ikke har brydd seg med å sjekke hvilke tall vi får ved å legge sammen to og to av tallene 1, 2, 3 og 4, og dermed har vedkommende bare gjettet på at summene også vil bli jevnt fordelt.

Denne oppgaven kan være interessant å bruke i undervisningen. Skaar og Syvertsen (2000) brukte blant annet denne oppgaven som en introduksjon til emnet sannsynlighetsregning fordi de mente det kunne være nyttig med en introduksjon som er noe annet enn der læreren forklarer og elevene regner oppgaver.

4.12.2 Oppgave 12b

De skal spille i alt 24 ganger. Hvor mange ganger tror du Kari vil vinne? ________________

Hvorfor mener du dette?

____

Svarene fra denne oppgaven henger mye sammen med hva elevene har svart på oppgave 12a. Resultatene fra oppgave 12b ble som følger:

Tabell 4.12.2: Resultater fra oppgave 12b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
8 ganger	1	1	6	0	8	5,7 %
12 ganger	13	6	16	10	45	31,9 %
16 ganger	**1**	**1**	**14**	**1**	**17**	**12,1 %**
18 ganger	2	3	8	1	14	9,9 %
Andre svar	6	4	17	13	40	28,4 %
Ubesvart	1	2	8	6	17	12,1 %
Sum	24	17	69	31	141	100,0 %

Som vi ser i tabellen var det svært få elever som svarte riktig på denne oppgaven, kun 17 av 141 elever. Det var cirka 32 % av elevene som svarte at Kari kom til å vinne 12 ganger og det stemmer bra med resultatene fra oppgave 12a, der det var omtrent 37 % som svarte at spillet var rettferdig. Det kan også være verdt å merke seg at over 12 % av elevene lot være å svare på denne oppgaven, noe som er det høyeste i hele testen.

Klassevis er det interessant å se at både skole 38, 39 og de praktiske klassene ved skole 40 hadde alle kun én elev som svarte riktig på denne oppgaven, mens i de rene teoriklassene var det litt over 20 % som svarte riktig. Det kan også være verdt å nevne at ved skole 38 var det 13 elever som mente at Kari kom til å vinne 12 ganger, mens det i oppgave 12a kun var 12 elever som mente at spillet var rettferdig.

Det store antallet elever som valgte å ikke svare på oppgaven i tillegg til det lave antallet riktige svar tyder på at dette var en oppgave som elevene hadde store problemer med. Denne tendensen blir også styrket ved de forskjellige svarene som ble gitt. Det var elever som svarte alt fra at Kari kom til å vinne 0 ganger til at hun kom til å vinne alle 24 gangene.

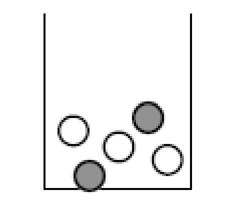
Begrunnelsene fra denne oppgaven er, i likhet med svarene, svært forskjellige. De fleste elevene som tror Kari vil vinne 12 ganger, har naturlig nok begrunnet dette med at Kari og Ola sannsynligvis kommer til å vinne like mange ganger. Det er også svært mange elever som ikke har skrevet noe begrunnelse i hele tatt. Dette styrker teorien om at dette er en vanskelig oppgave som elevene hadde problemer med.

4.13 Oppgave 13

Oppgave 13 var en urneoppgave. Forskjellen fra urneoppgaven i oppgave 1 var at det i denne ikke var med tilbakelegging. Den innholdt to deloppgaver og ideen til oppgaven var hentet fra Breiteig og Green.

4.13.1 Oppgave 13a

13. Vi har ei krukke med hvite og sorte kuler. Den ser slik ut:



Vi trekker tilfeldig ut kuler fra krukken. Etter at vi har trukket, noterer vi ned fargen, ***men legger ikke kulene nedi krukka igjen***.

a) Vi har trukket: H S H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

Resultatet fra oppgaven ble som følger:

Tabell 4.13.1: Resultater fra oppgave 13a

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hvit	0	2	1	4	7	5,0 %
Sort	0	0	1	0	1	0,7 %
Like sannsynlig	**24**	**15**	**67**	**27**	**133**	**94,3 %**
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	69	31	141	100,0 %

Som vi ser var det få elever som hadde problemer med denne oppgaven. Hele 94,3 % av elevene svarte riktig. 7 av de 8 elevene som svarte galt mente at det var størst sannsynlighet for å trekke en hvit kule. Klassevis var det små forskjeller, alle klassene gjorde det veldig bra og spesielt klassen på skole 38 og de tre teoriklassene ved skole 40.

Grunnen til at resultatet ble så bra som det ble her i forhold til oppgave 1 som var en lignende oppgave er nok at her var det ikke med tilbakelegging. Det kan virke som om de aller fleste elevene har fått med seg at det ikke er med tilbakelegging i denne oppgaven. En hovedgrunn til dette kan nok være at jeg har skrevet *"men vi legger ikke kulene nedi krukka igjen"* med uthevet og kursiv skrift i denne oppgaven, i motsetning til hva som var tilfellet i oppgave 1. Av de elevene som har gjort feil kan det se ut som om en del er slurvefeil, mens noen få ikke har lest oppgaven nøyaktig nok og har trodd at også denne oppgaven, i likhet med oppgave 1, var med tilbakelegging. Her har vi et eksempel på en elev som det kan tyde på har trodd at oppgaven var med tilbakelegging: *"Det er størst sjanse for at neste kule er hvit for det er flest hvite kuler"*.

4.13.2 Oppgave 13b

b) Vi har trukket: H H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

Resultatene fra oppgave 13b var som følger:

Tabell 4.13.2: Resultater fra oppgave 13b

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Hvit	0	1	2	2	5	3,5 %
Sort	**24**	**16**	**66**	**28**	**134**	**95,0 %**
Like sannsynlig	0	0	1	1	2	1,4 %
Ubesvart	0	0	0	0	0	0,0 %
Sum	24	17	69	31	141	100,0 %

Som vi ser i tabellen var resultatet, ikke overraskende, omtrent det samme som i oppgave 13a. Det var 134 av 141 elever som svarte riktig, noe som var en mer enn i forrige oppgave. 5 av de 7 elevene som svarte galt svarte at det var størst sannsynlighet for at en hvit kule ble trukket ut neste gang. Med såpass gode resultater er det igjen kun små forskjeller på resultatene til de forskjellige klassene.

Grunnen til at resultatet er så bra på denne oppgaven er nok de samme som i forrige oppgave. Når jeg ser på begrunnelsene til de elevene som har gjort galt tyder det på at de fleste har gjort slurvefeil. Her er to eksempler på begrunnelser som tyder på at eleven har gjort en slurvefeil: 1) *"Det er like stor sjanse for begge da det er like mange sorte og hvite kuler"* og 2) *"Det er størst sjanse for at neste kule er hvit fordi det bare er hvite igjen"*. Det kan tyde på at den andre eleven i dette tilfellet har trodd at det var to sorte kuler som var trukket ut først og at det dermed kun var igjen tre hvite kuler i urnen.

4.14 Oppgave 14

Oppgaven 14 var en oppgave som jeg laget i lag med mine veiledere. Oppgaven skulle være en oppgave som kunne relateres til det daglige liv. Vi endte da opp med en oppgave som handlet om kreft og om sjansen for å få kreft blir større om en person røyker. Hensikten med oppgaven var å teste elevenes forståelse for begrepet tilgjengelighet. Oppgaven var som følger:

14. Det har vært forsket mye på røyking, og skadevirkningene dette kan gi. En av konklusjonene er at røyking gir økt sjanse for lungekreft. Berit sine fire besteforeldre røykte mye, men alle ble over 90 år og døde av andre grunner enn lungekreft. Derfor tror ikke Berit på det som forskerne sier om røyking. Hvem mener du har rett?

☐ Jeg er enig med Berit fordi ______________________________

____

☐ Jeg er enig med forskerne fordi ______________________________

____

Resultatene fra oppgave 14 var som følger:

Tabell 4.14: Resultater fra oppgave 14

	Skole 38	Skole 39	Skole 40		Sum	Prosent
Svar			1T	1P		
Enig med Berit	5	4	7	5	21	14,9 %
Enig med forskerne	**18**	**13**	**60**	**24**	**115**	**81,6 %**
Ubesvart	1	0	2	2	5	3,5 %
Sum	24	17	69	31	141	100,0 %

Vi ser i tabellen at hele 115 av 141 elever har svart at de er enige med forskerne. 21 elever er enige med Berit, mens 5 elever har valgt å ikke svare på oppgaven. Klassevis er det kun små forskjeller, men det kan være verdt å nevne at det er flest elever som er enige med Berit i de to blandede klassene. I begge disse klassene er det over 20 % som mener at Berit har rett.

Ser vi på begrunnelsene til de elevene som er enige med Berit kan det virke som om noen har brukt tilgjengelighet:

- *Jeg tror ikke røyking er så farlig som folk påstår.*
- *Hennes erfaringer med døden av røyk eller ikke snakker imot forskerne.*
- *Fordi alle dei fire dødde når dei blei over 90 år, er det på grunn av alder!*

Her tar muligens elevene resultatet fra Berit sine besteforeldre og blander med eventuelle egne erfaringer, som er lett tilgjengelig, og dette overvinner troen på forskerne siden dette er selvopplevd. Her har kanskje elevene venner eller foreldre som røyker, og som ikke er blitt syke. Dette er et resultat som er lett tilgjenglig for eleven og som gir andre resultater enn forskerne.

4.15 Sammenligning av testresultat blant ulike oppgaver

I denne delen vil jeg prøve å sammenligne testresultatene fra oppgaver som ligner på hverandre og tar for seg noen av de samme begrepene. Jeg har brukt dataprogrammet Minitab som hjelpemiddel for å lage kji-kvadrattester som finner mer ut av sammenhengen mellom oppgaver.

4.15.1 Oppgave 2 mot oppgave 3

Først vil jeg se litt mer på oppgave 2 og 3. Dette er to oppgaver som begge skulle teste elevenes forståelse for tilfeldighet, og derfor kan det være interessant å se om det kan være noen form for sammenheng mellom resultatene fra de to oppgavene. Jeg har satt elevenes resultater inn i en tabell ved hjelp av Minitab og denne ble som følger (se også vedlegg5):

Tabell 4.15.1: Sammenhengen mellom oppgave 2 og 3

		Oppgave 3		
		Galt	Riktig	Sum
Oppgave 2	Galt	22 (19,55)	30 (32,45)	52
	Riktig	31 (33,45)	58 (55,55)	89
	Sum	53	88	141

Vi ser ut fra tabellen at det var 22 elever som svarte galt på begge oppgavene, mens det var 58 elever som svarte riktig på begge oppgavene. Det var 30 elever som svarte galt på oppgave 2 og riktig på oppgave 3, mens 31 elever svarte riktig på oppgave 2 og galt på oppgave 3.

Jeg vil nå teste ut denne eventuelle sammenhengen. Nullhypotesen er at det ikke er noen sammenheng mellom resultatene på de to oppgavene, mens den alternative hypotesen er at det finnes en signifikant sammenheng. Da starter vi med å regne ut oddsforholdet mellom de elevene som har svart riktig og galt. En fullstendig gjennomgang av utregningene for oddsforhold og konfidensintervall kan sees i vedlegg 4.

For å regne ut dette oddsforholdet bruker vi formelen $\psi = \dfrac{p_{11}/p_{12}}{p_{21}/p_{22}}$ der p_{11} står for sannsynligheten for at en elev havner i første rad og første kolonne (altså galt på begge oppgavene), p_{12} står for sannsynligheten for at en elev havner i første rad og andre kolonne (galt på oppgave 2, riktig på oppgave 3) og så videre. Et naturlig estimat for denne ψ er

$$\hat{\psi} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}\cdot n_{22}}{n_{21}\cdot n_{12}}.$$

Ved bruk av denne formelen får vi da følgende resultat på det estimerte oddsforholdet:

$$\hat{\psi} = \frac{n_{11}\cdot n_{22}}{n_{21}\cdot n_{12}} = \frac{22\cdot 58}{30\cdot 31} \approx 1{,}372.$$

Dersom det ikke er noen sammenheng vil $\dfrac{p_{11}}{p_{12}} = \dfrac{p_{21}}{p_{22}} \Rightarrow \psi = 1$.

Dersom det er en positiv sammenheng vil $\dfrac{p_{11}}{p_{12}} > \dfrac{p_{21}}{p_{22}} \Rightarrow \psi > 1$.

Dersom det er en negativ sammenheng vil $\dfrac{p_{11}}{p_{12}} < \dfrac{p_{21}}{p_{22}} \Rightarrow \psi < 1$.

Siden svaret er såpass nærme 1 kan vi ikke si noe om noen sammenheng foreløpig, derfor lager jeg et 95 % konfidensintervall for ψ:

$\hat{\psi} \approx 1{,}372$ og da blir $\log_e \hat{\psi} \approx 0{,}316$. Da får vi følgende konfidensintervall for $\log_e \psi$:

$$\left[0{,}316 \pm 1{,}960\sqrt{\frac{1}{22}+\frac{1}{30}+\frac{1}{31}+\frac{1}{58}}\right] = [0{,}316 \pm 1{,}960 \cdot 0{,}358] = [-0.386, 1.018].$$

95 % konfidensintervallet for ψ er dermed

$$\left[e^{-0.386}, e^{1.018}\right] = [0.680, 2.768].$$

Siden 1 er innenfor dette konfidensintervallet kan vi ikke forkaste nullhypotesen om at det ikke er noen sammenheng. Dette resultatet er noe overraskende da oppgave 2 og 3 har mange likhetstrekk som for eksempel at begge er laget for å teste elevenes forståelse for tilfeldigheter. På forhånd ville jeg trodd at det var flere som hadde enten riktig på begge oppgavene eller galt på begge. Siden begge oppgavene er avkrysningsoppgaver kan en av grunnene til at det ikke er noen signifikant sammenheng mellom resultatene på de to oppgavene være at en del elever har tippet og fått rett på en av de.

4.15.2 Oppgave 4 mot oppgave 5

Både oppgave 4 og oppgave 5 skal teste elevenes forståelse for de store talls lov og derfor tror jeg det kan være interessant å se om det er noen sammenheng mellom resultatene fra disse oppgavene. Til å begynne med har jeg satt resultatene fra de to oppgavene inn i en tabell ved hjelp av Minitab (se også vedlegg 6):

Tabell 4.15.2: Sammenhengen mellom oppgave 4 og 5

		Oppgave 5		
		Galt	Riktig	Sum
Oppgave 4	Galt	61 (52,11)	32 (40,89)	93
	Riktig	18 (26,89)	30 (21,11)	48
	Sum	79	62	141

Som vi ser ut fra tabellen var det 61 elever som svarte galt på begge oppgavene, mens 30 elever svarte riktig på begge. Det var 32 elever som svarte galt på oppgave 4 og riktig på oppgave 5, mens det var 18 elever som svarte riktig på oppgave 4 og galt på oppgave 5.

Nullhypotesen min er som tidligere at det ikke er noen sammenheng mellom resultatene fra de to oppgavene, mens den alternative hypotesen er at det finnes en signifikant sammenheng. Det estimerte oddsforholdet mellom de elevene som har svart riktig og galt blir da (med tallene fra tabell 4.15.2):

$$\hat{\psi} = \frac{n_{11} \cdot n_{22}}{n_{21} \cdot n_{12}} = \frac{61 \cdot 30}{32 \cdot 18} \approx 3{,}177.$$

Siden $\hat{\psi}$ er relativt stor kan det tyde på at vi har en positiv sammenheng, men for å bli noe sikrere så lager jeg et 95 % konfidensintervall for ψ:

$\hat{\psi} \approx 3{,}177$, da får vi $\log_e \hat{\psi} \approx 1{,}156$. får vi følgende konfidensintervall for $\log_e \psi$:

$$\left[1{,}156 \pm 1{,}960\sqrt{\frac{1}{61}+\frac{1}{32}+\frac{1}{18}+\frac{1}{30}}\right] = [1{,}156 \pm 1{,}960 \cdot 0{,}370] = [0.431, 1.881].$$

95 % konfidensintervallet for ψ er dermed

$$\left[e^{0.431}, e^{1.881}\right] = [1.539, 6.560].$$

Dette konfidensintervallet for ψ inneholder ikke 1, noe som betyr at $\hat{\psi}$ er signifikant forskjellig fra 1 ved et 5 % signifikansnivå. Dermed kan nullhypotesen, $\psi = 1$, forkastes og vi kan konkludere med at det er en signifikant positiv sammenheng mellom resultatene fra oppgave 4 og 5. Dette betyr at det var klart flere elever som enten svarte riktig på begge spørsmålene eller svarte galt på begge, enn man kunne forventet dersom det ikke var noen sammenheng mellom oppgavene.

Dette var et resultat som jeg hadde forventet på forhånd da oppgavene er veldig like både i oppbygningen og i det teoretiske innholdet. Begge oppgavene tar for seg de store talls lov der elevene skal avgjøre hvorvidt en av to hendelser som de blir presentert for er mer sannsynlig å inntreffe enn den andre eller om det er like stor sannsynlighet for at de inntreffer.

4.15.3 Oppgave 6a mot oppgave 6b

Dersom vi ser på resultatene fra oppgave 6a og 6b (tabell 4.6.1 og tabell 4.6.2) kan det tyde på at de som har forstått gjennomsnittsbegrepet virkelig har forstått det og har svart riktig på begge deloppgavene, mens de som ikke har en så god og robust forståelse har til dels store problemer og har dermed enten svart galt på begge oppgavene eller ikke svart i hele tatt. Dette har jeg forsøkt å undersøke litt nøyere ved å se på hvor mange elever det faktisk var som svarte riktig eller galt på begge deloppgavene. Nullhypotesen min er at det ikke er noen sammenheng mellom oppgave 6a og 6b, mens den alternative hypotesen er at det er en sammenheng. Jeg fikk følgende tabell fra Minitab (se også vedlegg 7):

Tabell 4.6.3: Sammenhengen mellom Oppgave 6a og 6b

		6b		
		Galt	Riktig	Sum
6a	Galt	38 (12,51)	4 (29,49)	42
	Riktig	4 (29,49)	95 (69,51)	99
	Sum	42	99	141

De øverste tallene i rutene er antall elever som havnet innenfor denne kategorien, for eksempel kan vi se at det var 38 elever som svarte galt på både oppgave 6a og 6b. Tallene under disse, de i parenteser, er de estimerte tallene dersom det ikke hadde vært noen sammenheng mellom disse to oppgavene. Som vi ser i tabellen kan det se ut som det er en sammenheng mellom de to oppgavene da det er relativt stor forskjell på de estimerte tallene og det faktiske antallet elever som havnet innenfor hver gruppe. Det ser ut til at det er en tendens til at elevene enten har svart riktig på begge eller feil på begge. Testen viser med en p-verdi på tilnærmet 0 (se vedlegg 7) at det er en signifikant sammenheng, men den viser ikke hvilken sammenheng det er. Dette kan vi derimot finne ut ved å regne ut oddsforholdet mellom de elevene som har svart riktig og galt, dette er nemlig et mål for hvor sterk sammenhengen er.

Dette oddsforholdet, ψ, er altså definert som $\psi = \frac{p_{11}/p_{12}}{p_{21}/p_{22}}$, men vi bruker følgende estimat for gjøre et anslag for ψ:

$\hat{\psi} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11} \cdot n_{22}}{n_{21} \cdot n_{12}}$. Med tallene fra tabell 4.15.1 får vi da:

$$\hat{\psi} = \frac{n_{11} \cdot n_{22}}{n_{21} \cdot n_{12}} = \frac{38 \cdot 95}{4 \cdot 4} = 225{,}625$$

Som vi ser får vi et veldig høyt tall. Dette tyder på at vi har en positiv sammenheng mellom oppgavene, og en tendens til at mange elever har svart galt på begge oppgavene eller riktig på begge oppgavene.

I dette tilfellet får vi altså et tall som er over 1, og da kan vi konkludere med at vi muligens har en positiv sammenheng mellom oppgavene. Vi kan ikke fastslå det med helt sikkerhet da dette bare er et anslag for ψ, men siden $\hat{\psi} = 225{,}625$ tyder dette på at $\psi \neq 1$.

Bruker vi nå dette til å lage et 95 % konfidensintervall for ψ i oppgave 6 får vi følgende:

$\hat{\psi} = 225{,}625$ og da blir $\log_e \hat{\psi} \approx 5{,}419$. Da blir 95 % konfidensintervallet for $\log_e \psi$:

$$\left[5{,}419 \pm 1{,}960\sqrt{\frac{1}{38} + \frac{1}{4} + \frac{1}{4} + \frac{1}{95}}\right] = [5{,}419 \pm 1{,}960 \cdot 0{,}733] = [3.983, 6.855].$$

Videre får vi følgende 95 % konfidensintervall for ψ:

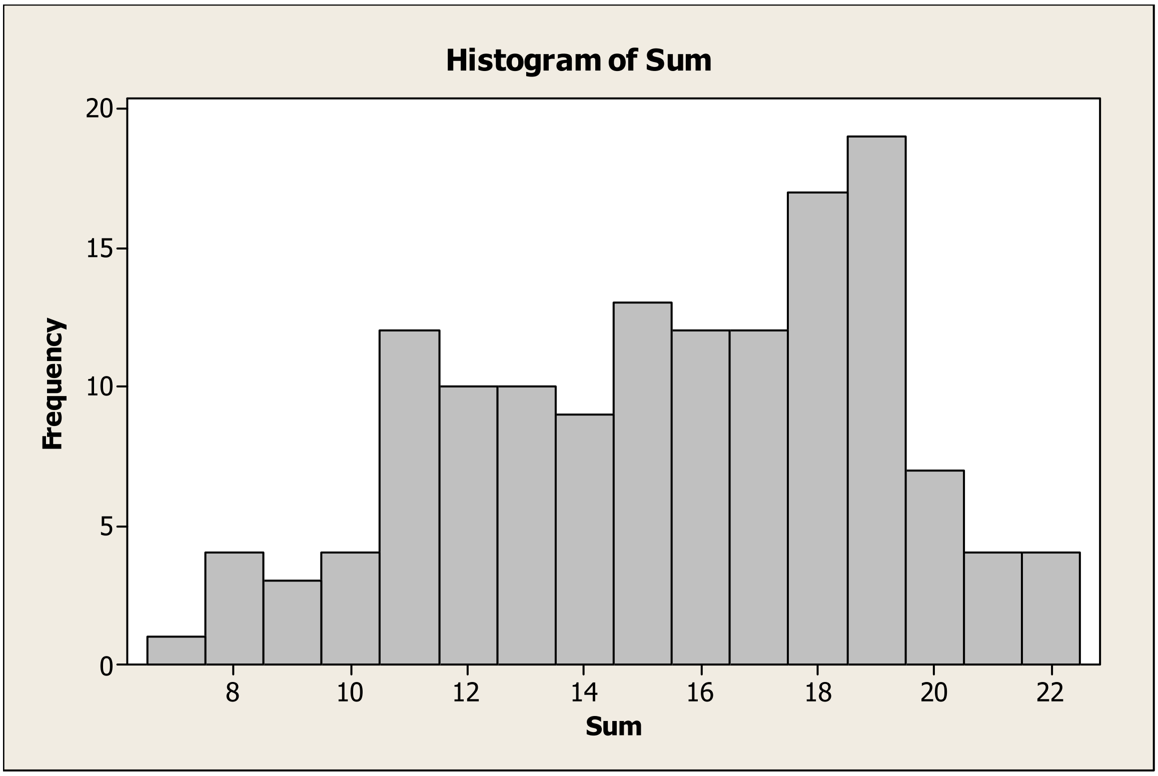
$$\left[e^{3.983}, e^{6.855}\right] = [53.68, 948.61].$$

Siden konfidensintervallet ikke inneholder 1 er det klart at $\hat{\psi}$ er signifikant forskjellig fra 1 på et 5 % signifikansnivå. Dette forteller oss at nullhypotesen min om at det ikke er noen sammenheng mellom oppgave 6a og 6b kan forkastes på ved $\alpha = 0{,}05$ og vi kan konkludere med at $\psi > 1$.

Dette resultatet vil si at vi med 95 % sikkerhet kan si at det var en sammenheng mellom oppgave 6a og 6b, og de som svarte riktig på oppgave 6a, med stor sannsynlighet også ville svare riktig på 6b, og vice versa. Ut fra tabellen kan vi se at det var svært mange elever som enten svarte galt på begge oppgavene eller svarte riktig på begge. Det var faktisk bare 8 av 141 elever som svarte riktig den ene oppgaven og galt på den andre.

4.16 Sammenligning av testresultat

I denne delen av kapittelet vil jeg se på hvordan ulike grupper gjorde det sammenlagt på testen. Jeg vil for eksempel sammenligne hvordan guttene gjorde det i forhold til jentene og hvordan de teoretiske klassene gjorde det i forhold til de praktiske klassene og de blandede klassene. For å gjøre dette har jeg brukt Minitab som et hjelpemiddel og jeg har sett på antall riktige svar som hver elev har hatt i testen. Først vil jeg vise et histogram som viser resultatene til alle de 141 elevene som deltok i testen:



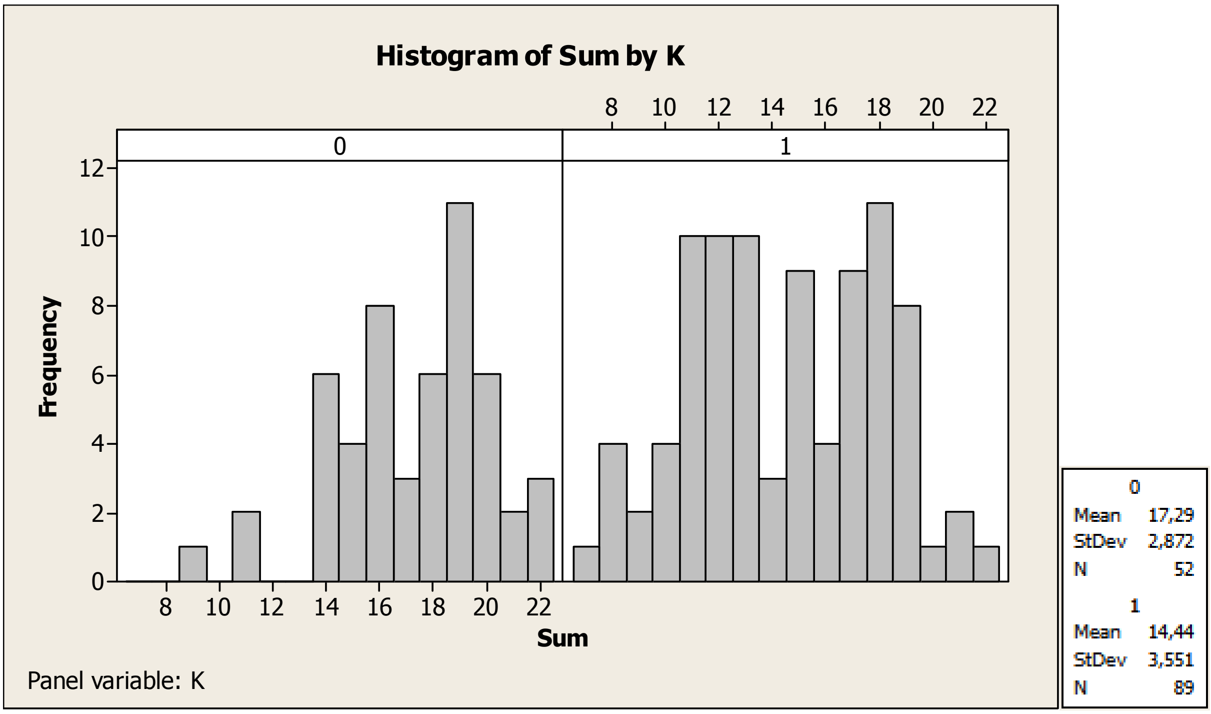
Figur 4.16.1: Antall elever per antall riktig svar.

Det var mulig å få maksimalt 24 riktige svar på denne testen, og som vi ser av histogrammet var det beste resultatet blant elevene 22 riktige, mens det svakeste resultatet var 7 riktige. Gjennomsnittet var cirka 15,49 med et standardavvik på cirka 3,58.

4.16.1 Mellom kjønnene

Først vil jeg se på hvordan guttene gjorde det mot jentene. Dette tror jeg kan være interessant fordi blant annet Green (1982) og Seland (1996) har funnet store forskjeller i prestasjonene i matematikk blant gutter og jenter. I Greens undersøkelse var et av funnene at guttene gjorde det markant bedre enn jentene. Dette synes jeg var et meget interessant funn og vil derfor se om dette også stemmer med resultatene til elevene som har deltatt i min undersøkelse.

Av de 141 elevene som deltok i min undersøkelse var det 52 gutter og 89 jenter. I begge de to blandede klassene og i alle de tre rene praktiske klassene var det et overtall av jenter, mens det var flere gutter enn jenter i alle de tre teoretiske klassene. I histogrammet under ser vi resultatene til guttene og jentene:

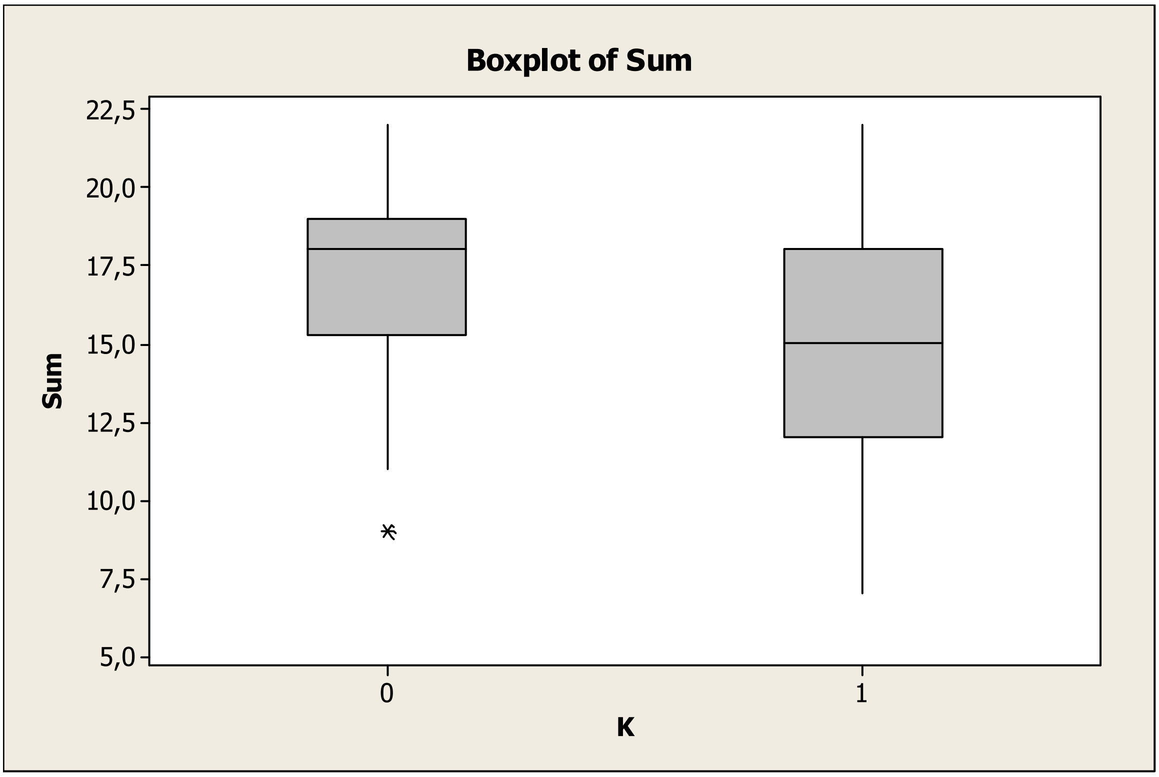


Figur 4.16.1.1: Testresultatene til guttene (0) og jentene(0).

I histogrammet ser vi resultatene til guttene til venstre og resultatene til jentene til høyre. De er kodet som 0 (gutt) og 1 (jente) som vi kan se over søylene og i den lille tabellen til høyre. I denne tabellen ser vi også at gjennomsnittet til guttene var 17,29 riktige svar med standardavvik på 2,872, mens jentene hadde et gjennomsnitt på 14,44 med standardavvik på 3,551. Dette betyr at guttene gjorde det markant bedre enn jentene da de oppnådde et snitt på nesten tre riktige mer.

Vi kan se i histogrammet at det bare var tre gutter som gjorde det dårligere enn 14 riktige. Dette er et meget sterkt resultat. De fleste guttene havnet mellom 14 og 20 riktige svar, mens det var noen få som havnet på 21 og 22 riktige. Blant jentene var det et mye mer spredt resultat. Vi ser at de fleste havnet mellom 11 og 19 riktige svar, mens det var litt flere som havnet under 11 enn det var som hadde 20 eller flere riktige svar. Jentenes resultater ligner mye mer på en normalfordeling enn det guttenes resultater gjør. Ettersom jeg ville forventet noe som lignet en normalfordeling hos begge kjønn kan det tyde på at det er guttene som har prestert meget bra i stedet for at det er jentene som har prestert dårlig.

For å få frem forskjellene mellom guttene og jentene tar jeg også med et boksplott av resultatene. Vi ser igjen guttene til venstre og jentene til høyre:



Figur 4.16.1.2: Boksplott av testresultatene til guttene (0) og jentene (1).

Av boksplottet her ser vi at guttene gjorde det en god del bedre enn jentene. De to boksene viser hvor de midterste 50 % havnet. Det vil si at halvparten av guttene havnet mellom cirka 15 og 19 riktige, mens halvparten av jentene havnet mellom cirka 12 og 18. Videre betyr dette at de 25 % svakeste av guttene fikk dårligere enn 15 riktige, mens den fjerdedelen av guttene som fikk best resultat fikk over 19 riktige. Blant jentene fikk 25 % av de svakeste jentene dårligere enn 12 riktige, mens den fjerdedelen av jentene som gjorde det best fikk over 18 riktige.

Ser vi på boksplottet igjen kan vi se en strek som går igjennom de to boksene. Dette er medianen. For guttene er denne 18, mens jentenes median er 15. Dette viser også hvor stor forskjell det var på resultatene til de to kjønnene. De to strekene som går oppover og nedover på boksene er da som nevnt de 25 % beste og 25 % dårligste resultatene. Blant guttene ser vi også at vi har en stjerne litt under der streken slutter. Dette betyr at det var et resultat som skilte seg noe ut blant datamaterialet. I dette tilfellet var det en av guttene som hadde prestert noe svakere enn de andre.

Dersom vi sammenligner de to boksplottene litt mer kan vi se at det er mer symmetri i boksplottet til jentene. Dette betyr at resultatene er tilnærmet normalfordelt. Boksplottet til guttene er derimot ikke så symmetrisk. Vi har blant annet medianen som ligger veldig høyt inne i boksen, og streken oppover er klart kortere enn streken nedover. Dette betyr at den halvparten av guttene som presterte best var veldig jevn. Disse elevene fikk resultater på mellom 18 og 22 riktige. Videre kan vi se at den halvparten av guttene som presterte dårligst var mer spredt. Disse la seg på et resultat mellom 11 og 18 riktige, i tillegg til dette ene spesialtilfelle med 9 riktige.

Siden det var relativt store forskjeller mellom kjønnene har jeg også laget en tabell som viser prosentandel riktig svar på hver oppgave i testen:

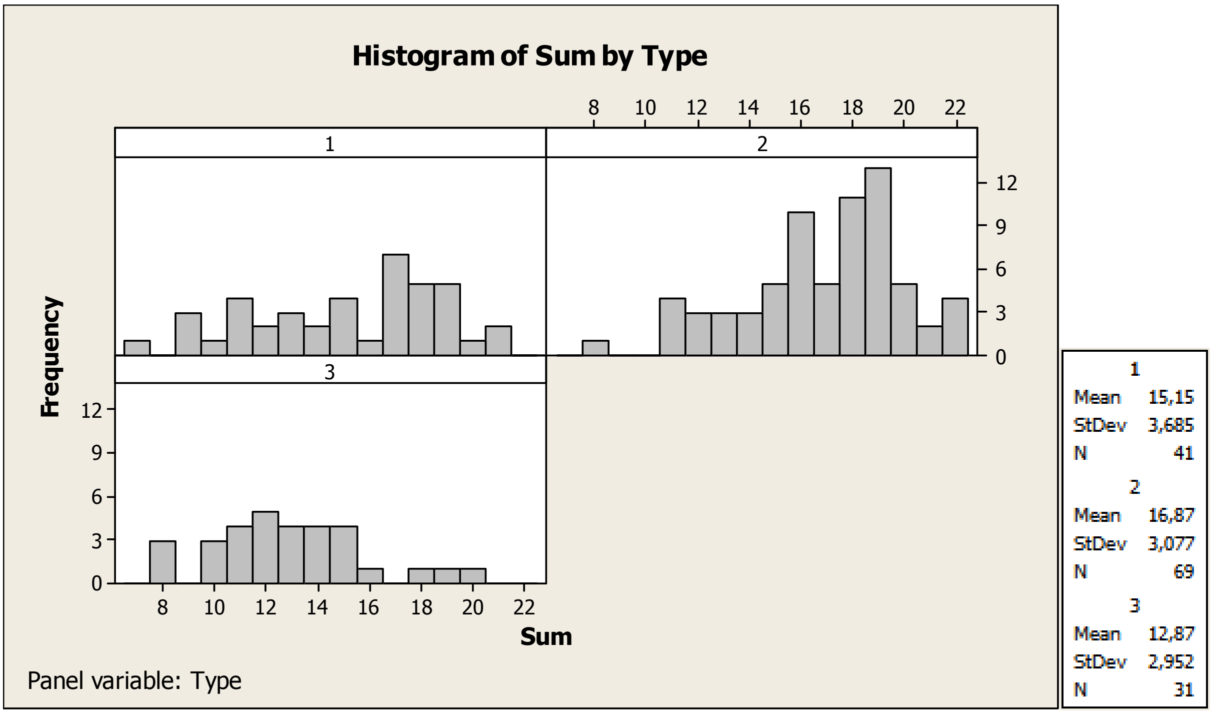
Tabell 4.16.1: Prosentandel riktige svar mellom kjønnene.

	Prosentandel riktige svar	
	Gutt	Jente
Oppg.1a	86,5	62,9
Oppg.1b	84,6	56,2
Oppg.1c	80,8	56,2
Oppg.2	50	70,8
Oppg.3	69,2	58,4
Oppg.4	40,4	30,3
Oppg.5	65,4	31,5
Oppg.6a	90,4	58,4
Oppg.6b	90,4	58,4
Oppg.7a	92,3	97,8
Oppg.7b	26,9	15,7
Oppg.8	42,3	27
Oppg.9a	96,2	94,4
Oppg.9b	94,2	78,7
Oppg.9c	88,5	86,5
Oppg.10a	90,4	71,9
Oppg.10b	51,9	52,8
Oppg.11a	28,8	24,7
Oppg.11b	26,9	11,2
Oppg.12a	61,5	50,6
Oppg.12b	23,1	7,9
Oppg.13a	96,2	93,3
Oppg.13b	96,2	94,4
Oppg.14	84,6	79,8

Som vi ser var gutten jevnt over bedre enn jentene, noen plasser betydelig bedre som i oppgave 1, 5, 6 og 10a. Jentene gjorde det betydelig bedre enn guttene på oppgave 2, mens de var noe bedre på oppgave 7. Noen av disse resultatene vil jeg komme tilbake til i diskusjonskapittelet.

4.16.2 Mellom blandede, teoretiske og praktiske klasser

Jeg vil nå sammenligne testresultatene mellom de ulike klassene. Jeg har valgt å dele de inn i tre grupper. De tre teoretiske klassene er en gruppe, de tre praktiske klassene er en gruppe og de to blandede klassene er den siste gruppen. Det første jeg har gjort er å lage et histogram, ved hjelp av Minitab, som viser hvordan resultatene fordelte seg i disse tre gruppene. I histogrammet er gruppene kodet som 1 (de blandede klassene), 2 (de teoretiske klassene) og 3 (de praktiske klassene).:

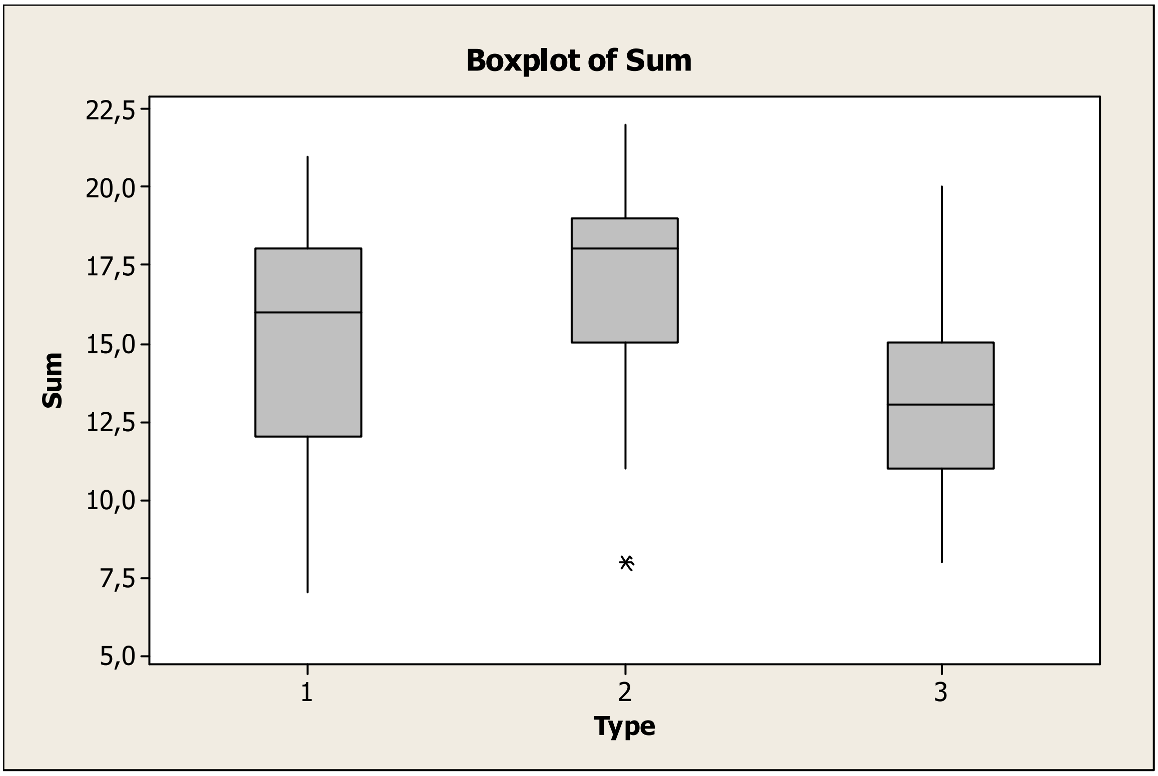


Figur 4.16.2.1: Testresultatene fordelt klassevis.

Vi kan se i den lille tabellen til høyre at det var henholdsvis 41, 69 og 31 elever i de tre gruppene. Det var ikke overraskende elevene fra de rene teoretiske klassene som gjorde det best. De fikk en gjennomsnittsscore på 16,87. De to blandede klassene var like bak med 15,15 riktige svar i gjennomsnitt, mens de rene praktiske klassene havnet litt bak med en middelverdi på 12,87 riktige svar. Det er heller ikke overraskende at standardavviket er klart størst i de blandede klassene da disse inneholder elever fra både praktiske og teoretiske klasser. Dette blir også bekreftet av histogrammet der vi kan se at resultatene til elevene var veldig spredt i de blandede klassene, vi hadde få elever innenfor hver søyle. Elevene spredte seg fra 7 riktige og opp til 21 riktige svar, men det var flest elever med 17, 18 og 19 riktige.

Dersom vi ser på histogrammet til de teoretiske klassene er dette mer samlet. De fleste elevene i disse klassene fikk mellom 15 og 20 riktige svar på testen, med et maksimum på 22 riktige. Det var kun 1 av disse 69 elevene som gjorde det svakere enn 11 riktige. Denne eleven havnet på 8 riktige svar. Ser vi på histogrammet til de praktiske klassene kan vi blant annet se at det kun var 4 av 31 elever som oppnådde mer enn 15 riktige svar. De fleste elevene innen denne gruppen havnet på mellom 10 og 15 riktige svar. Maksimumet var 20 riktige som én elev oppnådde, mens minimumet for denne gruppen var 8 riktige. Tre elever oppnådde dette resultatet.

I tillegg til histogrammet har jeg også laget et boksplott for resultatene til disse tre gruppene. Dette har jeg også laget ved hjelp av Minitab og gruppene er kodet som i histogrammet (1, 2 og 3):



Figur 4.16.2.2: Boksplott av grupperesultatene.

Ser vi på alle tre boksplottene og sammenligner de er det ikke den store overraskelsen. De teoretiske klassene gjorde det best, mens de praktiske klassene fikk det svakeste resultatet. Ser vi på den svakeste fjerdedelen i hver gruppe kan vi se at de blandede klassene strekker seg fra 7 til 12 riktige, de teoretiske klassene strekker seg fra 11 til 15 riktige i tillegg til et "utspring" med 8 riktige, mens de praktiske klassene strekker seg fra 8 til 11 riktige. Går vi videre til de 50 % i midten ser vi at de blandede klassene går fra 12 til 18 riktige med en median på 16 riktige, de teoretiske klassene går fra 15 til 19 riktige med en median på 18 riktige og tilslutt har vi de praktiske klassene som går fra 11 til 15 riktige med en median på 13 riktige. Den beste fjerdedelen blant de blandede klassene går fra 18 riktige og opp til 21 riktige. Blant de teoretiske klassene strekker denne fjerdedelen seg fra 19 riktige og opp til 22 riktige, mens den går fra 15 til 20 riktige blant de praktiske klassene.

Ser vi på symmetrien i disse tre boksplottene er det boksplottet til de blandede og praktiske klassene som passer best til en normalfordeling. Boksplottet til de teoretiske klassene kan minne litt om boksplottet til guttene i figur 4.16.1.2. Vi har en litt skjev fordeling der den halvparten av elevene som oppnådde det beste resultatet er mye mer komprimert enn den andre halvparten. Vi kan se i boksplottet at den øvre halvdelen har fordelt seg mellom 18 og 22 riktige, mens den nedre halvdelen ligger fra 11 til 18 riktige, i tillegg til denne ene eleven med 8 riktige.

4.16.3 Mellom hvordan de selv mener de presterer i matematikkfaget

I denne delen vil jeg sette resultatene til elevene opp mot hvordan de svarte på oppgave 15 i testen. Oppgave 15 var som følger:

15. Hvordan synes du selv at du gjør det i matematikkfaget?

☐ Veldig bra

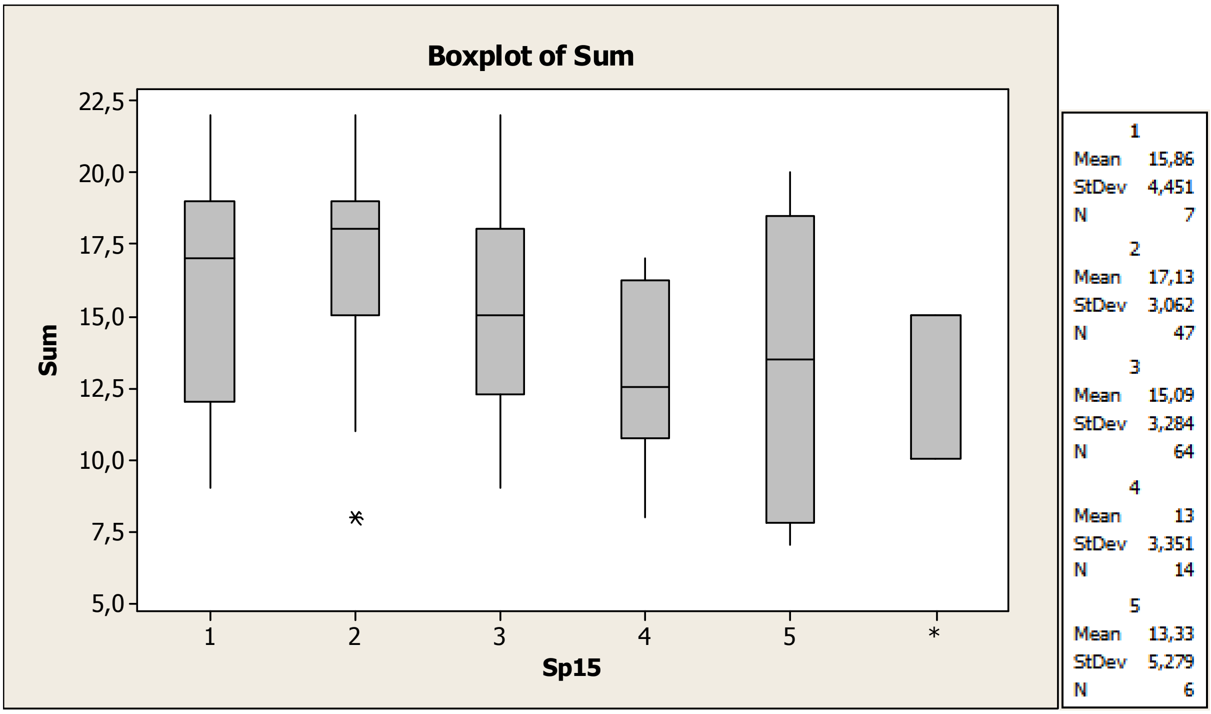
☐ Bra

☐ Middels

☐ Dårlig

☐ Veldig dårlig

Her skulle altså elevene vurdere seg selv i matematikkfaget. Jeg har satt elevenes resultater opp mot hva de svarte på denne oppgaven. Her er svarene kodet som 1 (veldig bra), 2 (bra), 3 (middels), 4 (dårlig) og 5 (veldig dårlig):



Figur 4.16.3.1: Resultater for presteringsgruppene.

Som vi kan se i tabellen til høyre var det flest elever som mente at de gjorde det bra eller middels i matematikkfaget, henholdsvis 47 og 64. De tre andre gruppene hadde alle få elever. Det var 7 elever som mente at de gjorde det veldig bra, 14 elever som mente at de gjorde det dårlig, mens det var 6 elever som mente de gjorde det veldig dårlig. Ser vi på gjennomsnittene til de ulike gruppene kan vi se at det var de elevene som mente at de gjorde det bra i matematikkfaget som hadde det høyeste gjennomsnittet med 17,13 riktige. Deretter fulgte de som hadde krysset av for veldig bra med 15,86 riktige i gjennomsnitt foran de som mente de gjorde det middels, disse hadde et gjennomsnitt på 15,09. Det dårligste gjennomsnittet fikk faktisk de som mente de gjorde det dårlig i matematikk med 13 riktige i gjennomsnitt, mens de som mente de gjorde det veldig dårlig havnet like foran med 13,33.

Dersom vi ser på boksplottene til 1 og 3, har disse klare likhetstrekk noe som er litt overraskende. Begge har store spredninger i antall riktige. Man har elever som hadde fra 9 riktige og opp til 22 riktige i begge grupper og den midterst halvparten av elever ligger innenfor noenlunde like grenser. Også gruppe 5 hadde relativt stor spredning. Dette var elevene som mente at de gjorde det veldig dårlig i matematikk. Her hadde vi elever som fikk fra 7 til 20 riktige på testen.

4.16.4 Mellom hvordan de liker matematikkfaget

I denne delen vil jeg sette resultatene til elevene opp mot hvordan de svarte på oppgave 16 i testen. Oppgave 16 var som følger:

16. Hva synes du om matematikkfaget?

☐ Liker det veldig godt

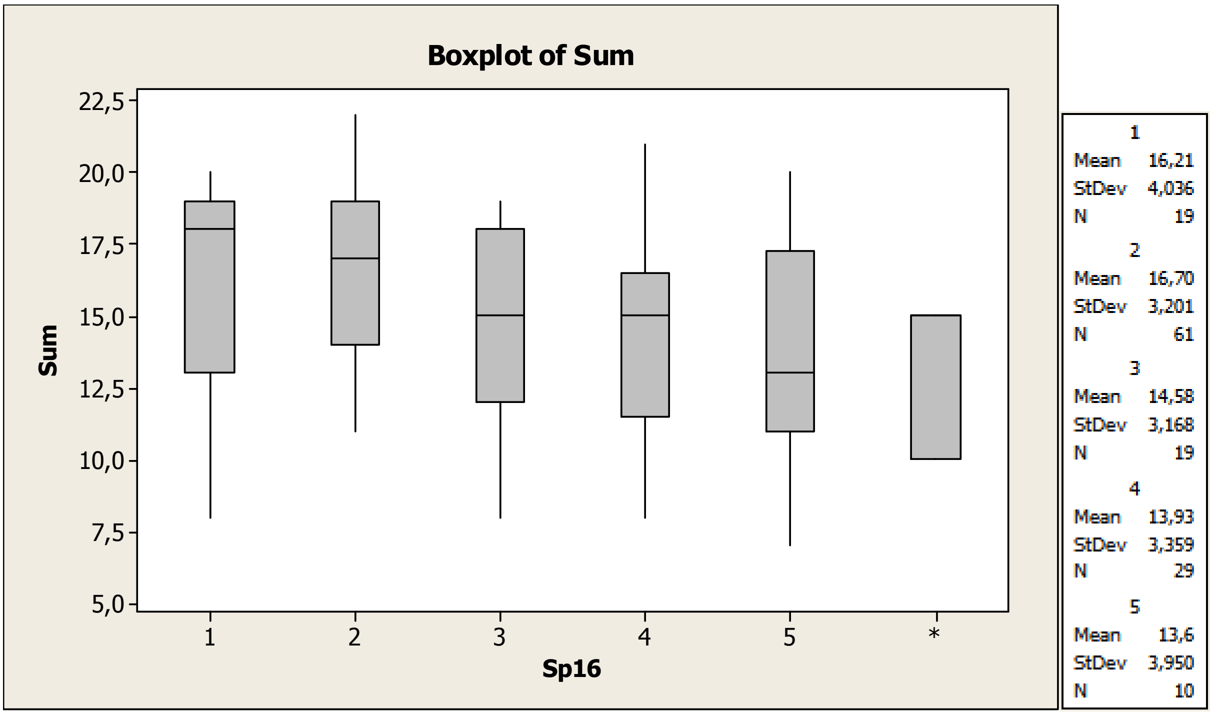
☐ Liker det godt

☐ Liker det dårlig

☐ Liker det veldig dårlig

☐ Ingen formening

Her skulle elevene svare på hva de synes om matematikkfaget, om de liker det eller ikke. Jeg har satt elevenes resultater i testen opp mot hvordan de liker matematikkfaget. Her er svarene kodet som 1 (veldig godt), 2 (godt), 3 (ingen formening), 4 (dårlig) og 5 (veldig dårlig):



Figur 4.16.4.1: Resultater satt opp mot hvordan elevene liker matematikkfaget.

Som vi ser har vi en del likheter med figur 4.16.3.1. Elevene er mer fordelt på denne oppgaven enn på den forrige. Vi har flest elever som liker matematikk godt, 61, mens det er

29 elever som liker matematikk dårlig. Det er 19 elever som liker matematikk veldig godt og samme antall elever har ingen formening. Ti elever liker matematikk veldig dårlig.
Vi kan se av boksplottet at den halvparten av de som liker matematikk veldig godt har gjort det veldig bra på testen. Alle disse ligger på 18 riktige eller bedre. På tross av dette er de elevene med flest riktige blant de elevene som liker matematikk godt. Til og med blant elevene som liker matematikk dårlig er det elever som har gjort det bedre enn de beste i gruppe 1.

Vi ser at det er spesielt stor spredning blant resultatene i både gruppe 1, 4 og 5. Alle disse gruppene inneholder elever med både 20 eller flere riktige og ned til 8 eller færre riktige. Dette ser vi også igjen i standardavviket i tabellen til høyre i figuren. Spesielt gruppe 1 og 5 har høyt standardavvik. En annen interessant ting vi kan lese ut fra boksplottet er at elevene som liker matematikk dårlig og elevene som ikke har noen formening har relativt like resultater. Medianen er 15 hos begge grupper, og den dårligste halvparten hos begge grupper er omtrent lik. Det er i den øvre halvparten vi har noen forskjeller. Det tredje kvartilet er noe bedre i gruppe 3, mens de aller beste i gruppe 4 er noe bedre enn de beste i gruppe 3.

Som en oppsummering av oppgave 15 og 16 har jeg laget en tabell over hvordan elevene svarte på disse:

Tabell 4.16.4.1: Resultater fra oppgave 15 og 16.

		Oppg.16					
		Veldig godt	Godt	Ingen formening	Dårlig	Veldig dårlig	Sum
Oppg.15	Veldig bra	3	3	0	1	0	7
	Bra	11	29	4	2	1	47
	Middels	5	28	12	16	3	64
	Dårlig	0	1	3	7	3	14
	Veldig dårlig	0	0	0	3	3	6
	Sum	19	61	19	29	10	138

Som vi ser fra tabellen var det flest elever som vurderte prestasjonene sine som bra eller middels i kombinasjon med at de liker matematikkfaget godt. I tillegg var det en del elever som mente de presterte middels i matematikkfaget, men samtidig likte faget dårlig eller ikke hadde en formening om hvordan de likte det. Det var også en del elever som mente de gjorde det bra i matematikkfaget og likte faget veldig godt. På de andre kombinasjonene var det få elever. I summen kan vi at det er 138 elever som ligger inne i tabellen. Grunnen til at dette tallet ikke er 141 er at det var tre elever som ikke svarte på disse to oppgavene.

5 Diskusjon

I dette kapittelet vil jeg diskutere noen av resultatene som jeg gjennomgikk i analysekapittelet. Jeg vil ikke kommentere alle resultatene fra alle oppgavene, men vil prøve å velge ut de resultatene som jeg synes er de viktigste og mest interessante. Disse resultatene vil jeg så prøve å begrunne på bakgrunn av teorien som er gjennomgått i kapittel 2 og tidligere undersøkelser.

5.1 Forholdsbegrepet

Det første begrepet elevene møtte i testen var *forholdsbegrepet*. Ut fra resultatene på oppgavene i testen som handlet om begrepet forhold (oppgave 1, 9, 10 og 13) kan det virke som om dette er et begrep som elevene har relativt god kontroll på. På disse oppgavene var prosentandelen riktige svar følgende:

Tabell 5.1: Prosentandel riktige svar på forholdsoppgaver

Oppgave	Prosentandel riktige svar
1a	71,6
1b	66,7
1c	65,2
9a	95,0
9b	84,4
9c	87,2
10a	78,2
10b	52,5
13a	94,3
13b	95,0

Her ser vi at prosentandelen var klart lavere på oppgave 1 og 10b enn de andre oppgavene. Som jeg nevnte i analysekapittelet kan mye av grunnen til at mange elever svarte feil på oppgave 1 være at de misforsto oppgaven og trodde uttrekningen var med tilbakelegging. Det er mulig at prosentandelen ville vært noenlunde lik som på de andre oppgavene dersom dette ikke hadde vært tilfelle. I oppgave 10b var muligens illustrasjonen noe uklar. De hvite feltene i de to lykkehjulene (se kapittel 4.10.2) var muligens laget for like slik at mange elever trodde begge hjulene hadde 25 % sannsynlighet for suksess.

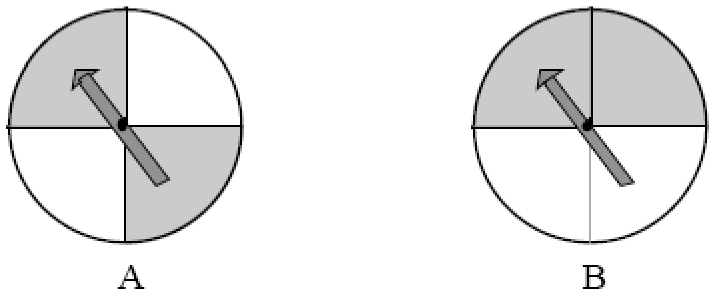
Green (1982) fant store mangler i elevenes forståelse for begrepet forhold. Han testet elever i alderen 11-16 år og resultatene viste en bedring etter hvert som elevene ble eldre, men mine resultater viser at elevene klarte seg bedre enn de eldste elevene i Greens undersøkelse. Dermed kan jeg ikke si, ut fra mine resultater, det samme om elevene fra min undersøkelse. Tabell 5.1 viser at det var gode resultater på oppgavene som handlet om forhold. I tillegg kan som nevnt de fire dårligste resultatene forklares med misforståelser av oppgavetekst eller illustrasjoner til oppgavene.

Begrunnelsene til elevene på de forskjellige forholdsoppgavene tyder også på at forhold er et begrep som elevene har relativt god kontroll på. De viser også at flere elever bruker sammenligning som strategi når de løser oppgavene. Noen utvalgte begrunnelser fra forholdsoppgavene er:

- *Fordi det er i boks A er det $\frac{1}{4}$ og i boks B $\frac{2}{9}$* (Oppgave 9b).
- *Fordi det er en mer hvit kule i forhold til svarte kule oppe i krokka* (Oppgave 1a).
- *Det er $\frac{2}{3}$ svarte kuler* (oppgave 13b).
- *Hvis vi slår sammen de to hvite på B så blir de like stor som den hvite på A* (Oppgave 10b)

Den første begrunnelsen viser bruk av forhold for å løse oppgaven. Begrunnelse to og tre viser elever som har sammenlignet antallet hvite og sorte kuler. Den siste begrunnelsen som jeg har tatt med viser en elev som har svart galt, men ut fra hva han skriver tolker jeg det som at forholdsbegrepet er forstått fordi han ser på arealene av de hvite feltene i forhold til hverandre og sammenligner disse. Her tyder begrunnelsen på at det er illustrasjonen til oppgaven som er misforstått og dermed er grunnen til at eleven har svart galt.

Selv om resultatene mine tyder på at forholdsbegrepet er godt forstått, tyder de også på at elevene har noe større problemer når de skal sammenligne områder i stedet for kuler eller lignende konkrete gjenstander. Oppgave 10 var den eneste forholdsoppgaven i testen der elevene skulle sammenligne to områder i forhold til hverandre. Resultatene fra deloppgave a er noe dårligere enn på de andre. Flere enn 1 av 5 elever mener sannsynligheten for å treffe hvit felt på følgende to lykkehjul ikke er den samme:



Figur 5.1: Illustrasjon fra oppgave 10a.

Blant begrunnelsene til disse elevene hadde noen skrevet at sannsynligheten var større når de hvite feltene var spredt, mens andre hadde skrevet det motsatte, at sannsynligheten var større når de hvite feltene var samlet. Dette tyder på at elevene bruker en årsakstenkning. De leter etter en årsak for å løse oppgaven. Dette viser at selv om forholdsbegrepet generelt sett er forstått, er det en del 16-17 åringer som har problemer med å bruke begrepet.

5.2 Representativitet og tilfeldighet

Det neste elevene ble testet i var *representativitet* og *tilfeldighet*. Både oppgave 2 og 3 tok for seg disse begrepene, og ser vi på resultatene fra oppgavene virker det som om dette var noe som elevene hadde problemer med. Nesten halvparten av elevene på første steget i den videregående skolen ville ifølge mine resultater sagt at en spesiell rekke med myntkast var mer sannsynlig enn en annen. Videre mener også nesten halvparten av elevene at en spesiell lottorekke er mer sannsynlig å bli trukket ut enn en annen.

På bakgrunn av nevnte resultater kan det virke som om mange av elevene har en ufullstendig forståelse for begrepet tilfeldighet og bruker representativitetsstrategien slik at resultatet blir galt. På oppgave 2 svarer mange elever at sannsynligheten er størst for å få enten en av de to rekkene KMKMKM og MKKMKM eller begge to. Disse to rekkene representerer tre kron og tre mynt, altså en lik fordeling mellom antall kron og mynt ved seks kast. Ut fra flere av begrunnelsene til elevene virker det som mange tror at disse to rekkene har størst sannsynlighet for å inntreffe fordi det er størst sannsynlighet for å få tre kron og tre mynt på seks kast. En elev som krysset av for begge disse rekkene hadde skrevet følgende begrunnelse:

> *"Det er to sider av en mynt. Altså 50 % sjangse for hver av sidene. Egentlig er det vel like stor sannsynlighet på alle, fordi en mynt er veldig tilfeldig, men i prinsippet skal det bli like mange K som M".*

Slik som jeg tolker denne begrunnelsen ville jeg sagt at eleven bruker først en sannsynlighetstenkning som sier at egentlig er det like stor sannsynlighet for alle rekkene. Deretter tar årsakstenkningen over, han leter etter en årsak for hvorfor en av rekkene skal være mer sannsynlig enn de andre. Eleven kommer da frem til at det i prinsippet skal bli like mange kron og mynt. Denne årsakstenkningen overvinner sannsynlighetstenkningen og han krysser av for de to nevnte rekkene. En slik årsakstenkning kunne også ses igjen hos noen andre elever i deres begrunnelser.

Flere elever krysset av for rekken KMKMKM og nesten samtlige av disse begrunnet valget med at det er like stor sannsynlighet for å få kron som det er for å få mynt. Da tolker jeg det slik at de ser på rekken KMKMKM som den rekken som representerer denne sannsynligheten på best mulig måte. De bruker altså det Kahnemann og Tversky kaller representativitet. De ser på tre kron og tre mynt som mest representativ for den 50-50 fordelingen som myntkast er. Samtidig ser de på annenhver kron og mynt som mest representativ for den tilfeldige prosessen som myntkast er. De viser også med dette at de bruker det som kalles "the negative recency effect" eller "the gambler's fallacy" – spillerens mistak. De tror at det som har skjedd før vil få konsekvenser for det som skjer senere, at mynten har et slags "minne". Dersom det første kastet er kron vil det være større sannsynlighet for å få mynt neste gang.

Også de fleste elevene som svarte riktig, at alle fire rekkene er like sannsynlige, brukte begrunnelsen at det er like stor sannsynlighet for å få kron som det er for å få mynt. Disse har da forstått at en bestemt rekke ikke er mer sannsynlig enn en annen selv om det er tre eller seks kron i rekken. Flere nevner også at hvert kast er tilfeldig og uavhengig.

På oppgave 3 var de fleste gale svarene at en eller begge de to rekkene uten spesielt mønster hadde størst sannsynlighet for å bli trukket ut. Igjen kan det virke som om disse bruker representativitet når de kommer frem til det de tror er det riktige svaret. Det virker som om mange av eleven tror at rekken med tallene 1-7 har veldig liten sannsynlighet for å bli trukket ut, mens de to andre rekkene har større sannsynlighet. Elevene ser da på de to andre rekkene som representanter for tilfeldige blandede rekker uten noe mønster. Flere av begrunnelsene tyder også på dette:

- *Fordi dette er ei blanding av både einsiffra og tosiffra tal.*
- *Fordi det står ikkje i rekkfølge, dei er sprett.*
- *Sannsynet for at tala kjem etter kvarandre er tilnærma lik null.*

Som vi ser av begrunnelsene tror elevene at rekker som inneholder både ensifrede og tosifrede tall og rekker som ikke har tall i rekkefølge har større sannsynlighet for å bli trukket ut. En annen grunn til at de mener dette kan være at de bruker en tilgjengelighetsstrategi (Shaughnessy & Bergman, 1993). De har kanskje sett på lottotrekningen på fjernsynet og lagt merke til at det som regel er en rekke uten noe form for mønster som blir trukket ut. Dermed tror de at en bestemt oppgitt representant for en slik rekke har større sannsynlighet for å bli trukket ut.

Selv om oppgave 2 og 3 er to relativt like oppgaver i teorien fant jeg i kapittel 4.15.1 et interessant funn. Hypotesetesten der viste at det ikke var noen signifikant sammenheng mellom resultatene fra de to oppgavene. Mange elever hadde svart riktig på den ene oppgaven, men galt på den andre. Kun 58 av 141 elever hadde svart riktig på begge oppgavene (se tabell 4.15.1). Dette er med på å styrke teorien om at en del elever ikke har en fullgod forståelse for begrepet tilfeldighet.

5.3 De store talls lov

Neste begrep elevene ble testet i var *de store talls lov*. Både oppgave 4 og 5 tok for seg dette begrepet og ut fra resultatene mine virker det som om dette er et begrep eleven har store problemer med. Prosentandelen riktige svar på disse to oppgavene var henholdsvis 34,0 og 43,7, ikke ulikt resultatene til Breiteig (2002).

På oppgave 4 svarte over halvparten av elevene at de to hendelsene var like sannsynlige. Dette resultatet stemmer godt overens med resultatene til både Shaughnessy og Bergman (1993), Breiteig (2002) og Kahnemann og Tversky (referert i Shaughnessy og Bergman, 1993). Over halvparten av elevene mener altså at det er like sannsynlig at de blir født 7 jenter blant 10 nyfødte, som at det fødes 70 jenter blant 100 nyfødte. De tar ifølge Breiteig (2002) ikke hensyn til den 50-50 fordelingen som kommer frem fra befolkningsstatistikk og biologien. De leter etter en årsak og finner at forholdet 7 av 10 er det samme som 70 av 100. Argumentet deres blir dermed at de to hendelsene er like sannsynlige. Årsakstenkningen overvinner sannsynlighetstenkningen.

Resultatene fra oppgave 5 var noe bedre enn i oppgave 4, men de viser fortsatt at elevene hadde store problemer med de store talls lov. Cirka en fjerdedel av elevene svarte at de to hendelsene var like sannsynlig. Dette er under halvparten så mange som det var som svarte det samme i oppgave 4. Grunnen til dette tror jeg er at elevene som brukte en årsakstenkning i oppgave 4 har vanskeligere for å finne en årsak i oppgave 5. De har ikke lenger to like forhold som overbeviser de om at sannsynlighetene er like. Ut fra resultatene og begrunnelsene virker det som om mange av disse elevene som bruker en årsakstenkning har svart at sannsynligheten er størst for at gjennomsnittshøyden til en tropp på 30 mann er over 185 cm. Begrunnelsene de bruker for dette valget er at dess flere det er i troppen, dess større er sannsynligheten for at gjennomsnittet blir høyt. Noen av begrunnelsene som støtter denne teorien er:

- *Fordi med 30 mann kan gjennomsnittet bli høgare enn med 3 mann.*
- *Da er det flere folk.*
- *Fordi eg meiner at viss det er fleir folk vert det meir sannsynlig.*
- *Fordi det er mange rekrutter og det må nokk ein del til for å dra gjennomsnittet opp for det er sikkert mange lave mennesker som deltar i militæret.*
- *Jo fleire folk, jo større sjans er det for at fleire er høgare enn 182, og kan auka gjennomsnittet.*

Her ser vi begrunnelser som tyder på at eleven har brukt en årsakstenkning. De mener at dersom vi får flere folk inn i troppen vil sannsynligheten for å få mange høye mennesker øke, noe som igjen vil føre til at sannsynligheten for å øke gjennomsnittet stiger. De leter etter en årsak, finner denne og overser da at sannsynligheten er større for å få inn mennesker som er under 185 cm enn det er for å få folk som er over. Noen av begrunnelsene, som i den første og fjerde begrunnelsen over, tyder også på at begrepet gjennomsnitt ikke er godt forstått.

Oppgave 4 og 5 var altså to oppgaver som i teorien var ganske like, men der kontekst og formulering var forskjellig. I kapittel 4.15.2 viste analysen at det var en signifikant positiv sammenheng mellom resultatene fra oppgavene. Det vil si at det var en overvekt av elever som hadde svart enten riktig eller galt på begge oppgavene. I tabell 4.15.2 kan vi se at det var kun 30 av 141 elever som hadde svart riktig på begge oppgavene. Dette er litt over en femtedel av elevene og styrker klart teorien om at de store talls lov er et begrep som elevene har store problemer med.

De store talls lov var et begrep der det var relativt stor forskjell mellom kjønnene. Både på oppgave 4 og 5 gjorde guttene det klart bedre enn jentene, og spesielt på oppgave 5 (se tabell 4.16.1). Disse resultatene kan tyde på at guttene har en bedre forståelse av de store talls lov enn jentene.

5.4 Middelverdi

Middelverdi var også et begrep som elevene ble testet i. Oppgave 6, med sine to deloppgaver, var i hovedsak oppgaven som tok for seg dette begrepet. Ut fra resultatene i kapittel 4.6 virker det som om gjennomsnitt er et begrep som mange elever har god kontroll på. På begge deloppgavene svarte cirka 7 av 10 elever riktig.

I kapittel 4.15.3 analyserte jeg sammenhengen mellom resultatene fra oppgave 6a og 6b. Oppgavene er helt like, eneste forskjell er at et tall er endret på. Resultatene fra analysen viste da også at det var en signifikant positiv sammenheng mellom resultatene fra de to oppgavene. Kun 8 av 141 elever hadde svart riktig på én av oppgavene, mens 38 hadde svart galt på begge og 95 hadde svart riktig på begge. Disse resultatene tyder på altså på at 3 av 10 elever på første steget i den videregående skolen ikke klarer å regne ut størrelsen på en gjenstand når du vet størrelsen på de andre gjenstandene og gjennomsnittet alle skal ha.

Ut fra tabell 4.16.1 kan vi se at det var store forskjeller blant kjønnene på oppgave 6. Det viste seg at guttene gjorde det klart bedre enn jentene på denne oppgaven. Cirka 9 av 10 gutter svarte riktig på gjennomsnittsoppgaven, mens tallet for jentene var cirka 6 av 10. Dette er en relativt klar indikasjon på at guttene har en bedre forståelse for middelverdi enn jentene.

5.5 Konjunksjonsloven

Konjunksjonsloven var et annet begrep som ble testet. Det var i hovedsak oppgave 8 som tok for seg dette begrepet. Resultatet fra denne oppgaven visste at cirka 3 av 10 elever svarte riktig. Over halvparten av elevene svarte at damen sannsynligvis ville bli over 90 år, mens cirka 17 % mente at det var størst sannsynlighet for at hun ble over 100 år. Dette resultatet stiller seg bak resultatet til blant annet Kahnemann og Tversky (referert i Shaughnessy og Bergman, 1993) som har gjort funn som tyder på at mange ikke har en god forståelse for denne loven.

Ut fra begrunnelsene til elevene er det derimot vanskelig å si noe spesielt om forståelsen for akkurat konjunksjonsloven på denne oppgaven. De fleste begrunnelsene tyder på at elevene har trodd at svaralternativene på oppgaven er intervaller, noe som gjør at jeg ikke kan si så mye konkret om hvordan elevene har tenkt i forhold til konjunksjonsloven. Derimot kan jeg si at begrunnelsene til de fleste elevene tyder på at de har brukt en årsakstenkning når de har løst oppgaven. Noen utvalgte begrunnelser, med svaralternativet i parentes, fra oppgave 8 er:

- *Fordi hun trener hver dag og holder seg i form* (Hun blir over 90 år).
- *Ho holder seg i god form og har alle forutsetninger for å leve* (Hun blir over 100 år).
- *Fordi ho er så sprek. Det er litt uvanleg å kome over 100 år* (Hun blir over 90 år).
- *Fordi det kan skje masse uventa ting sjølv om ho virker relativt sunn* (Hun blir over 85 år).

Disse begrunnelsene tolker jeg dit hen at elevene bruker det faktum at hun er sprek og frisk samt at hun trener som en årsak til å kunne si at sannsynligheten er størst for at hun skal bli enten over 85, 90 eller 100 år. De aller fleste som hadde svart riktig på denne oppgaven hadde brukt en begrunnelse av lignende type som den fjerde som er nevnt over her. Noen få elever viste bruk av konjunksjonsloven i begrunnelsen sin som følgende elev: *"Fordi når du er over 85, kan du og vera over 100"*. Denne begrunnelsen tyder på at konjunksjonsloven er forstått, men som nevnt var det får elever som viste dette i begrunnelsen sin.

5.6 Betinget sannsynlighet

Oppgave 7 var en oppgave mange elever hadde problemer med. Deloppgave a var en relativt enkel oppgave der hele 135 av 141 elever svarte riktig, men på deloppgave b var det derimot større problemer. Under en femtedel av elevene svarte riktig. Grunnen til dette lave tallet kan være at elevene blandet inn betinget sannsynlighet når de løste oppgaven. Som jeg nevnte i analysekapittelet kunne flere begrunnelser tyde på dette.

5.7 Forventning

Resultatene fra oppgave 11 og 12 viste at elevene hadde problemer med begrepet forventning. Dette er et vanskelig tema som de ikke har vært mye borti tidligere på skolen, noe som nok er en av hovedgrunnene til at elevene viste svake resultater.

I oppgave 11 skulle elevene velge mellom tre lykkehjul som alle gav ulike forventede tilbakebetalinger. På de to deloppgavene valgte henholdsvis 26,1 % og 18,8 % av elevene hjulet som gav høyest forventet tilbakebetaling. Det var i tillegg kun et fåtall av disse elevene som begrunnet valget sitt med at den forventede tilbakebetalingen ville være størst ved spill på dette lykkehjulet. På 11a fikk elevene kun lov til å "spille" én gang og resultatet viste at over 6 av 10 elever valgte det hjulet som gav størst sannsynlighet for gevinst. Flesteparten av disse begrunnet valget sitt med at akkurat dette. På 11b fikk elevene muligheten til å "spille" 100 ganger, og her valgte over halvparten av elevene det hjulet som gav størst enkeltgevinster. Mange av disse elevene begrunnet valget sitt med at her fikk man størst premie. Dette tyder på at de ikke har forstått forventning da de velger hjulet som gir størst enkeltpremie, og ikke hjulet som gir størst forventet tilbakebetaling.

I oppgave 12 skulle først elevene avgjøre om et spill var rettferdig eller ikke. Cirka 37 % av elevene svarte at de mente at spillet var rettferdig. Som jeg nevnte i analysekapittelet kan noen av disse ha blitt "lurt" av den forstyrrende faktoren – de fire tallene i oppgaveteksten. I det tilfellet har ikke elevene analysert oppgaveteksten nøye nok. Dette resultatet er omtrent som det Skaar og Syvertsen (2007) har erfart. Resultatene fra oppgave 12b viste, i likhet med oppgave 11, at begrepet forventning var et begrep elevene hadde problemer med. Elevene skulle avgjøre hvor mange spilleomganger en av spillerne kunne forvente å vinne av 24 omganger. Cirka 12 % av elevene svarte riktig, dette er et svakt resultat selv om vi tar med i betraktningen at det er cirka 37 % av elevene som svarte galt på grunn av en følgefeil fra 12a. Dette støtter teorien og resultatene fra oppgave 11.

6 Konklusjon

I dette kapittelet vil jeg oppsummere noen av de viktigste resultatene fra diskusjonskapittelet. Jeg vil også svare på forskningsspørsmålene som jeg stilte i begynnelsen av oppgaven (se kapittel 1.2).

Det første forskningsspørsmålet som jeg stilte var følgende:

- Hvilke misoppfatninger innenfor noen utvalgte begreper innen statistikk og sannsynlighet finnes det blant elevgrupper i den videregående skolen og hvor utbredt er disse?

Ut fra mine resultater vil jeg si at en stor andel av elevene har en misoppfatning av de store talls lov. Litt over 1 av 5 elever svarte riktig på begge de to oppgavene som tok for seg de store talls lov. Både Shaughnessy og Bergman (1993), Breiteig (2002) og Kahnemann og Tversky (referert i Shaughnessy og Bergman, 1993) har funnet lignende resultater. Dette tyder på at de store talls lov er et begrep som det bør jobbes mye mer med i den norske skolen.

Representativitet er et annet begrep som ut fra mine resultater virker å være noe utbredt strategi blant elever. Både oppgave 2 og 3 viste at elever bruker representativitet når de løser sannsynlighetsoppgaver, noe som viser seg å lede til problemer en del ganger. Disse oppgavene viste også at en del elever har problemer med begrepet tilfeldighet.

Neste forskningsspørsmål var som følger:

- Hvilke begreper har elevene god kontroll på?

Resultatene miner viser at forhold er et begrep som elevene har en god forståelse for. Både på urneoppgavene (oppgave 1 og 13) og oppgave 9 viste at de fleste elevene kunne sammenligne forhold og bestemme hva som hadde størst sannsynlighet for å bli trukket ut. Dersom vi tar med elevene som misforstod oppgaven og ikke begrepet, klarte minst 8 av 10 elever å løse disse oppgavene.

Middelverdi var et annet begrep som det ut fra mine resultater viste seg at mange elever hadde god forståelse for. Cirka to tredjedeler av elevene svarte riktig på både oppgave 6a og 6b, som var de to oppgavene som i hovedsak tok for seg middelverdi. Spesielt guttene hadde god kontroll på dette begrepet. Hele 9 av 10 gutter hadde riktig svar på begge deloppgavene.

Det neste forskningsspørsmålet jeg stilte var følgende:

- Hva kan gjøres for å forbygge disse misoppfatningene?

Her bør blant annet media være mer kritiske til hva de skriver og hvilke illustrasjoner de bruker. Samtidig bør samfunnet være kritiske til statistikk- og sannsynlighetsstoff som de blir presentert for både i media og dagliglivet ellers. Begreper fra emnet må bli brukt på riktig måte, da vil elevene plukke opp dette og tilegne seg gode begrepsforståelser.

Det ligger også et viktig ansvar på oss lærere og skolen. Det må legges til rette for en god og motiverende undervisning. Et viktig element her tror jeg er hvilke oppgaver vi gir elevene. Jeg tror det er viktig at elevene får oppgaver som er motiverende og spennende, men det viktigste av alt er at de får arbeide med oppgaver der de får oppdage selv hva de ulike begrepene innebærer.

Det siste forskningsspørsmålet mitt var:

- Hva kan gjøres for å bygge robuste begreper innen emnet?

Jeg mener det allerede er tatt et steg i riktig retning ved at statistikk og sannsynlighet har blitt en mye større del av matematikkfaget i den siste læreplanen. Dette fører til at vi får flere undervisningstimer som igjen gjør det mulig å gå mer grundig inn på emnet.

For å bygge opp robuste begreper må vi også forebygge misoppfatninger innen de samme begrepene. Derfor går dette punktet hånd i hånd med forrige punkt, og dermed gjelder også det som ble nevnt under forrige forskningsspørsmål under dette.

Å ha varierende innfallsvinkler i undervisningen er et annet viktig grep man bør gjøre. Bruk av forskjellige undervisningsmetoder og medier fører til at man treffer flere elever med undervisningen da elever lærer på forskjellige måter.

7 Pedagogiske implikasjoner og videre forskning

I dette kapittelet vil jeg komme med undervisningsforslag innen undervisning av statistikk og sannsynlighet. Jeg vil også komme inn på hvilke videre forskning som trengs som bakgrunn for skolens arbeid for å utvikle en god begrepsforståelse blant elever.

Statistikk og sannsynlighet har allerede tatt et stort steg ved sin nye posisjon i matematikkfaget i den nye læreplanen. Dette gjør at elevene får flere undervisningstimer i emnet, noe som gjør at lærere får bedre tid til å bygge opp robuste begreper. For å gjøre dette har jeg noen forslag.

Som jeg nevnte i konklusjonen tror jeg et viktig element er at man har gode og interessante oppgaver. Som Breiteig (2002) nevner kan man bruke oppgaver med tema fra dagliglivet. Et eksempel fra Freudenthal-instituttet i Utrecht som han nevner er:

> *"En klasse skal til Kitzbühl i vinterferien. Bør de tegne reiseforsikring? Opplysninger innhentes. De finner at 7 % av elever på slike turer får bruk for en forsikring, de kan bli frastjålet kamera, lommebok, de kan falle i slalåmbakken og måtte gipse beinet. Utgiftene ved slike skader beregnes etter innhentede opplysninger, i gjennomsnitt 2300 kr for hvert tilfelle. Er da en forsikringspremie på 350 kroner rimelig? Burde elevene heller samle inn dette beløpet av hver til en felles sikkerhetskasse?"* (Breiteig, 2002, s. 9).

En annen ide er å gjøre en oppgave om til et eksperiment. Et eksempel på en aktuell oppgave er oppgave 12 i testen min, som Skaar og Syyvertsen (2007) har brukt til akkurat dette som en innledning på emnet statistikk og sannsynlighet. Breiteig (2002) nevner en struktur med fire punkter (etter at man har presentert oppgaven):

1) Elevene kan gjette sannsynligheten. Drøft ulike forslag.
2) Prøv ut! Bruk egnede hjelpemidler som lommeregner, datamaskin eller annet.
3) Reflekter over resultatet. Fremstillinger som grafer og diagrammer kan være til hjelp her.
4) Arbeid med lignende oppgaver for å drille begreper.

For å avdekke ulike tenkemåter og eventuelle misoppfatninger av begreper kan man bruke ulike diagnostiske oppgaver. I tillegg til dette kan man prøve å få til en dialog med elevene og eventuelt få elevene til å fortelle om hvordan de tenker.

Dersom man skal gjennomføre et forsøk eller en simulering vil det være lurt å tenke ut metoden på forhånd. For eksempel vil det være både tidkrevende og slitsomt dersom hver enkelt elev eller gruppe skal simulere 200 myntkast. Det man kan gjøre da er å la enkeltelever eller grupper gjøre for eksempel 20 myntkast og etterpå kan hver elev eller gruppe presentere resultatet sitt. Underveis kan man for eksempel se på hvordan frekvensene forandrer seg. Dette vil være en god måte å illustrere variasjon og de store talls lov på.

Det synes å være behov for mer forskning på begrepsforståelsen til elever i emnet statistikk og sannsynlighet. Resultatene fra min undersøkelse viste at flere elever har problemer med ulike begreper, men jeg tok bare for meg noen få utvalgte. En undersøkelse der man tar for seg noen andre begreper og ulike tenkemåter bør være aktuell.

Det kunne også vært interessant med en lignende undersøkelse som denne for å ha sett om man fant de samme resultatene som jeg har kommet frem til. Her kunne man eventuelt testet ut enda flere elever eller elever i andre aldre.

8 Refleksjoner over eget arbeid

Arbeidet med denne masteroppgaven har vært veldig lærerikt og interessant på mange måter. På den ene siden har jeg lært mye om emnet statistikk og sannsynlighet. Her har undervisning og læring i emnet vært de to hovedtingene jeg har lært om. Jeg har tatt med meg flere gode ideer til undervisningsmetoder og gode didaktiske oppgaver som kan brukes i undervisningen. På den andre siden har jeg lært veldig mye om prosessen rundt det å jobbe med en slik oppgave. Jeg har fått innsikt i ulike forskningsmetoder og lært spesielt mye om forskningsmetodene som jeg har brukt i denne oppgaven.

Skulle jeg gjort dette arbeidet om igjen er det en del ting jeg ville endret på. Etter å ha jobbet med diskusjonen og konklusjonen er det blant annet en del litteratur jeg kunne tenkt å lese og eventuelt tatt med i teorikapittelet. Jeg ser også nå i etterkant at det er noen oppgaver i testen som med fordel kunne vært byttet ut eller endret på. Oppgave 8 tror jeg ble for uklar, her kunne oppgaveteksten med fordel vært noe endret, da begrunnelsene tydet på at de fleste elevene trodde svaralternativene var intervaller. Også oppgave 10b var en oppgave som kunne vært noe endret. Resultatene viste at de hvite feltene ble for like slik at mange elever misforsto oppgaven.

Oppgave 14 er en oppgave som muligens kunne vært byttet ut. Den ble muligens noe for opplagt. Denne kunne med fordel vært byttet ut med en annen oppgave med samme formål, men der konteksten var endret.

Intervjuene gikk, som nevnt i metodekapittelet, ikke helt som planlagt. De ble gjort like etter testen var gjennomført. Hadde jeg fått gjort de om igjen ville jeg ha gjennomført de på et noe senere tidspunkt slik at jeg hadde hatt en mulighet til å studere svarene på forhånd. På denne måten ville jeg visst hvilke oppgaver som hadde gitt interessante funn og hvilke spørsmål som kunne vært aktuelle å stille.

Referanseliste

Breiteig, T. (2002). Regn med usikkerhet: Sjanse, risiko og matematikklæring. NSMO, Trondheim.

Breiteig, T. & Brekke, G. (1994) Matematikk – kunnskaper og holdninger. Kompendium i prosjektet: Ansvar for egen læring i matematikk. Program for utdanningsforskning, Norges forskningsråd.

DeBeres, R. (1988). Statistics for college-bound students: Are the secondary schools responding? *School Science & Mathematics, 88,* s. 200-209.

Falk, R. & Konold, C. (1992). The psychology of learning probability. I F. Gordon & S. Gordon (red.), *Statistics for the twenty-first century,* (s. 151-164). Washington, DC: Mathematical Association of America.

Green, D. R. (1982). *Probability concepts in 11-16 year old pupils*. Centre for Advancement of Mathematical Education in Technology, University of Technology Loughborough, Loughborough.

Henry, M. (2003). Probabilistic thinking (introduction). CERME-3, Thematic working group 5, Stochastic Thinking, Theme: Probabilistic Thinking.

Holmes, P. (1986). *The best of teaching statistics.* Sheffield: Teaching Statistics Trust.

Hundeland, P. S. (1996). *Fra sannsynlighetsregningens historie, utviklingstrekk, anvendelser og uavhengighet.* Hovedoppgave i matematikkdidaktikk, Kristiansand, Høgskolen i Agder.

Kahnemann D. & Tversky A. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90* (4), s. 293-315.

Kapadia, R. & Borovcnik, M. (red.) (1991). *Chance encounters: Probability in education.* Dordrecht: Kluwer Academic Publishers.

KIM-prosjektet (2005). Kvalitet i matematikkundervisningen. Hentet 18.mai 2008 fra http://kim.tfn.no/pmwiki.php?n=Main.HomePage

Konold, C. (1989). Informal conceptions of probability. *Cognition and instruction, 6,* s. 59-98.

KUF (1996). *Læreplanverket for den 10-årige grunnskolen.* Oslo: KUF.

Mertens, D. M. (2005). *Research and evaluation in education and psychology*. Thousand Oaks, California: Sage Publications.

Rolandsen, W. (2001). *Statistikk og sannsynlighetsregning i videregående skule.* Hovedoppgave i matematikkdidaktikk, Kristiansand, høgskolen i Agder.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. I D. Grows (red.), *Handbook of research in mathematics teaching and learning* (s.465-494). New York: MacMillan.

Shaughnessy, J. M. & Bergman, B. (1993). Thinking about uncertainty: probability and statistics. I P. Wilson (red.), *Research ideas for the classroom: High school mathematics* (s. 177-197). New York: MacMillan.

Skaar, B. & Syvertsen, H. T. (2007). Undersøkende aktiviteter med oppgavekort. I B. Jaworski, A. B. Fuglestad, R. Bjuland, T. Breiteig, S. Goodchild, B. Grevholm (red.) *Læringsfellesskap i matematikk* (s. 101-109). Bergen: Caspar Forlag AS.

Syvertsen, H. T. (1999). *Fra sannsynlighetsregning på 1700-tallet til norsk skole år 2000.* Hovedoppgave i matematikkdidaktikk, Kristiansand, Høgskolen i Agder.

Tamhane, A. C. & Dunlop, D. D. (2000). *Statistics and Data Analysis.* Upper Saddle River, New Jersey: Prentice Hall, Inc.

Telemarksforsking Notodden. Prosjekt 284: KIM – Kvalitet i matematikkundervisningen – Digitalisering. Hentet 18.mai 2008 fra http://www.tfn.no/article/articleview/548/1/8/

Utdanningsdirektoratet. Kunnskapsløftet – fag og læreplaner. Hentet 18.mai 2008 fra
http://www.udir.no/templates/udir/TM_UtdProgrFag.aspx?id=2103

Oversikt over vedlegg

1	Informasjonsskriv til skolene
2	Testen
3	Kodeskjema
4	Utregning av oddsforhold og konfidensintervall
5	Kji-kvadrattesten til oppgave 2 og 3
6	Kji-kvadrattesten til oppgave 4 og 5
7	Kji-kvadrattesten til oppgave 6a og 6b

Vedlegg 1: Informasjonsskriv til skolene

Informasjon vedrørende test i statistikk og sannsynlighet

Jeg er en mastergradsstudent i matematikkdidaktikk ved Universitetet i Agder. I mastergradsoppgaven min skal jeg skrive en oppgave om elevers begrepsforståelse i statistikk og sannsynlighet. Jeg er spesielt interessert i elevers misoppfatninger innen emnet, da tidligere undersøkelser har vist at mange elever sliter med dette. En kartlegging av disse misoppfatningene kan i fremtiden være til hjelp i undervisningen.

I den forbindelse med denne kartleggingen skal jeg gjennomføre en test på et utvalg av elever som går i første klasse i den videregående skolen. Testen vil basere seg på generell forståelse og ungdomsskolepensumet.

Testen vil være anonym, men jeg vil be om at elevene vurderer seg selv i matematikkfaget og hva de synes om matematikk. Dette gjør jeg for å kunne sammenligne svarene og se om det er noe systematisk mønster blant de som liker matematikk, de som synes matematikk er vanskelig osv. Dersom det er mulig kunne jeg også tenke meg å få noen få ord om hvordan nivået til klassen er av læreren, dette for at jeg skal kunne ha en mulighet til å se om det er noen forskjeller på svarene til klasser som ligger på ulike nivåer.

Den endelige oppgaven min vil bli sendt til skolen som takk for hjelpen.

Hilsen
Lars Aga

Vedlegg 2: Testen

<u>Sannsynlighet</u>

☐Gutt Skole: ____________________________

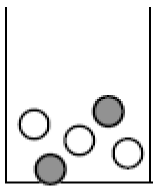
☐Jente Klasse: ___________________________

Les oppgavene nøye før du svarer. Selv om du er i tvil, så skal du ikke hoppe over noen oppgaver. ***Det er viktig at du svarer selv om du er usikker på svaret.***

Der det er ruter setter du kryss i ruta du mener er riktig. Mener du det er flere riktige svar, setter du bare flere kryss.

2. Vi trekker tilfeldig ut kuler fra en krukke. I krukka har vi både hvite og sorte kuler. Etter at vi har trukket en gang, noterer vi ned fargen og legger kula oppi krukka igjen. Vi har trukket ut noen kuler, og skal til å trekke igjen.

a

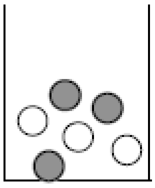


Vi har trukket: H S H ……

☐ Det er størst sjanse for at neste kule er hvit
☐ Det er størst sjanse for at neste kule er sort
☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

b 

Vi har trukket: S S H S ……

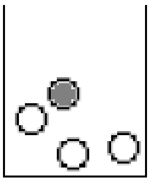
☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

c 

Vi har trukket: H H H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

2. Du gjør seks tilfeldige myntkast. Hvilken av de følgende rekkefølgene av mynt (M) og kron (K) er mest sannsynlig eller er de like sannsynlige?

☐ MKMMKM er mest sannsynlig

☐ MMMMKK er mest sannsynlig

☐ KMKMKM er mest sannsynlig

☐ MKKMKM er mest sannsynlig

☐ Alle fire rekkefølgene er like sannsynlige

Hvorfor mener du dette?

____

3. Hvilken av følgende Lottorekker tror du har størst sjanse for å bli trukket ut som vinnerrekke?

☐ 1 2 3 4 5 6 7

☐ 7 13 14 19 25 31 34

☐ 3 6 9 15 27 30 33

☐ Alle har like stor sjanse

Hvorfor mener du dette?

____

4. På et stort sykehus registrerer man alle barn som blir født et år. Følgende kan da hende:

A: Av de første 10 barna som blir født, er det 7 jenter.

B: Av alle de 100 nyfødte, er det 70 jenter.

☐ Hendelse A er mer sannsynlig enn hendelse B.

☐ Hendelse B er mer sannsynlig enn A.

☐ A og B er like sannsynlige.

Hvorfor mener du dette?

____

5. Gjennomsnittshøyden til alle rekrutter i Norge er et år 182 cm. Hva er da mest sannsynlig av følgende hendelser:

☐ At gjennomsnittshøyden til en liten tropp på 3 mann er over 185 cm

☐ At gjennomsnittshøyden til en tropp på 30 mann er over 185 cm

☐ Begge hendelsene er like sannsynlige

Hvorfor mener du dette?

____

6. Ole skal pakke fire egg som skal veie i gjennomsnitt 70 gram. Hva må det siste egget veie, dersom de tre første veier:

A: 70, 70 og 68 gram **Svar:** ____________

B: 69, 68 og 70 gram **Svar:** ____________

Forklar hvordan du tenkte:

A: ____

B: ____

7. I en klasse er det 13 gutter og 16 jenter. Klassen har fått to gratisbilletter til en konsert. Klassen vil trekke lodd om billettene. Hver elevs navn skrives på en lapp. Lappene legges så oppi en boks, og læreren trekker ut to lapper uten å se.

a)

☐ Det er størst sjanse for at den første som trekkes ut er en gutt

☐ Det er størst sjanse for at den første som trekkes ut er ei jente

Hvorfor mener du dette?

____

b)

☐ Det er størst sjanse for at de to som trekkes ut er en gutt og ei jente

☐ Det er størst sjanse for at de to som trekkes ut er to jenter

☐ Det er størst sjanse for at de to som trekkes ut er to gutter

Hvorfor mener du dette?

____

8. En kvinne fyller 80 år. Hun er sprek og frisk, og daglig går hun turer og svømmer. Hvilken av følgende hendelser er mest sannsynlig:

☐ Hun blir over 85 år.

☐ Hun blir over 90 år.

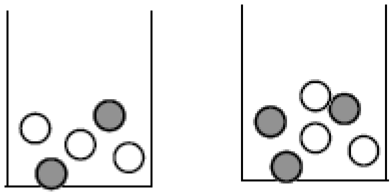
☐ Hun blir over 100 år.

Hvorfor mener du dette?

____

9. På et tivoli skal Petter trekke ei kule fra en boks uten å se. Boksen inneholder et visst antall hvite og sorte kuler, ***og han får en premie dersom han trekker en hvit kule***. På forhånd får Petter velge mellom to bokser, der han får vite antallet hvite og sorte kuler i hver boks.

a) Hvilken boks er det lurest å trekke fra, A eller B, eller er det det samme?



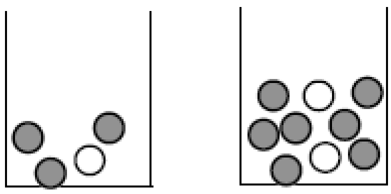
A B

Svar: ___________________________

Hvorfor mener du dette?

____

b) Hvilken boks er det lurest å trekke fra, A eller B, eller er det det samme?



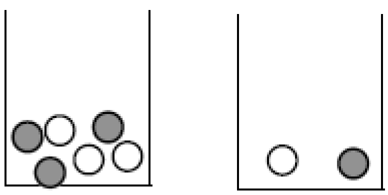
A B

Svar: ___________________________

Hvorfor mener du dette?

____

c) Hvilken boks er det lurest å trekke fra, A eller B, eller er det det samme?



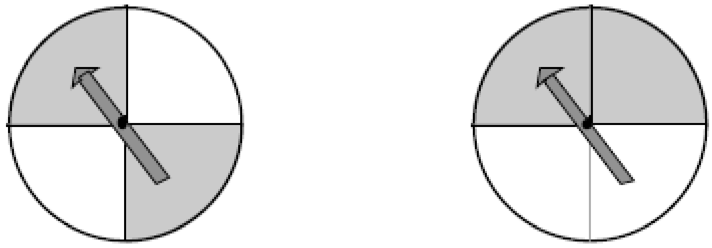
A B

Svar: ___________________________

Hvorfor mener du dette?

____

10. På et tivoli er det flere lykkehjul. En viser blir snurret rundt. Den stopper på et tilfeldig sted. ***Dersom den stopper på et hvitt felt vinner spilleren en premie.***



A B

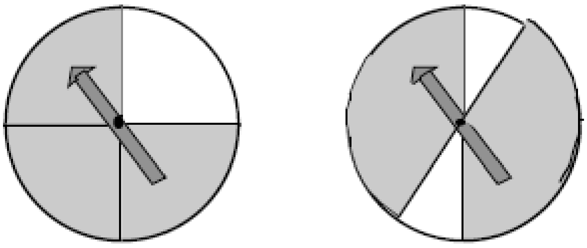
Hvilket hjul gir størst vinnersjanse?

☐ A

☐ B

☐ Samme vinnersjanse på begge hjul

Hvorfor mener du dette?



A B

Hvilket hjul gir størst vinnersjanse?

☐ A

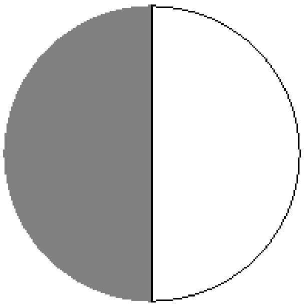
☐ B

☐ Samme vinnersjanse på begge hjul

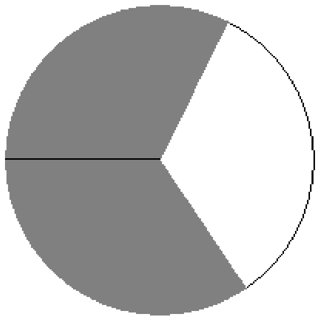
Hvorfor mener du dette?

---

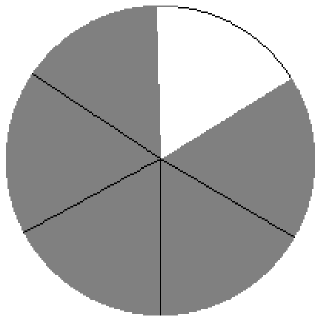
11. På et tivoli er det tre lykkehjul. En viser blir snurret rundt. Den stopper på et tilfeldig sted. ***Dersom den stopper på et hvitt felt vinner spilleren en premie.*** På lykkehjul A kan man vinne 180 kroner, på B er premien 330 kroner, mens man kan vinne 600 kroner på C.



A



B



C

Hvilket hjul vil du spille på, hvis

a) du bare kan spille en gang?

☐ A

☐ B

☐ C

Hvorfor? ____

b) du får lov til å spille 100 ganger?

☐ A

☐ B

☐ C

Hvorfor? ____

12. Ola og Kari spiller med fire like mynter, som er merket med tallene 1, 2, 3 og 4. De putter myntene i en krukke og trekker ut to av dem uten å se. Ola vinner hvis summen av de to tallene er et partall, mens Kari vinner hvis summen er et oddetall.

Er dette spillet rettferdig?

☐ Ja

☐ Nei

☐ Vet ikke

Hvorfor mener du dette?

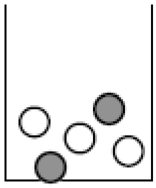
____

De skal spille i alt 24 ganger. Hvor mange ganger tror du Kari vil vinne? ________________

Hvorfor mener du dette?

____

13. Vi har ei krukke med hvite og sorte kuler. Den ser slik ut:



Vi trekker tilfeldig ut kuler fra krukken. Etter at vi har trukket, noterer vi ned fargen, ***men legger ikke kulene nedi krukka igjen***.

a) Vi har trukket: H S H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

b) Vi har trukket: H H ……

☐ Det er størst sjanse for at neste kule er hvit

☐ Det er størst sjanse for at neste kule er sort

☐ Det er like stor sjanse for begge

Hvorfor mener du dette?

____

14. Det har vært forsket mye på røyking, og skadevirkningene dette kan gi. En av konklusjonene er at røyking gir økt sjanse for lungekreft. Berit sine fire besteforeldre røykte mye, men alle ble over 90 år og døde av andre grunner enn lungekreft. Derfor tror ikke Berit på det som forskerne sier om røyking. Hvem mener du har rett?

☐ Jeg er enig med Berit fordi ____

____

☐ Jeg er enig med forskerne fordi ____

____

15. Hvordan synes du selv at du gjør det i matematikkfaget?

☐ Veldig bra

☐ Bra

☐ Middels

☐ Dårlig

☐ Veldig dårlig

16. Hva synes du om matematikkfaget?

☐ Liker det veldig godt

☐ Liker det godt

☐ Liker det dårlig

☐ Liker det veldig dårlig

☐ Ingen formening

Vedlegg 3: Kodeskjema

Kodeskjema

Skolene
38
39
40

Kjønn
X – jente
Y – gutt

Oppgavene
1 – riktig
11 – galt
0 – Ubesvart

Vedlegg 4: Utregning av oddsforhold og konfidensintervall

p_{11}	p_{12}
p_{21}	p_{22}

For å regne ut oddsforholdet bruker vi formelen $\psi = \frac{p_{11}/p_{12}}{p_{21}/p_{22}}$ der p_{11} står for sannsynligheten for at en elev havner i første rad og første kolonne, p_{12} står for sannsynligheten for at en elev havner i første rad og andre kolonne og så videre. Et naturlig estimat for denne ψ er

$$\hat{\psi} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11} \cdot n_{22}}{n_{21} \cdot n_{12}}.$$

Dette estimatet kommer av at hver p_{rk}, der r = rad og k = kolonne, er estimert ved n_{rk}/n (Tamhane & Dunlop, 2000, s. 327-328).

Dersom det ikke er noen sammenheng vil $\frac{p_{11}}{p_{12}} = \frac{p_{21}}{p_{22}} \Rightarrow \psi = 1$.

Dersom det er en positiv sammenheng vil $\frac{p_{11}}{p_{12}} > \frac{p_{21}}{p_{22}} \Rightarrow \psi > 1$.

Dersom det er en negativ sammenheng vil $\frac{p_{11}}{p_{12}} < \frac{p_{21}}{p_{22}} \Rightarrow \psi < 1$.

Dersom svaret hadde blitt nær 1 ville vi ikke fått en signifikant sammenheng mellom oppgavene. Dersom vi får et tall under 1 tyder det på en negativ sammenheng mellom oppgavene, en tendens til at mange elever har svart riktig på én av de to deloppgavene.

Vi kan nå regne ut et konfidensintervall for ψ, og da nytter vi følgende fremgangsmåte:

$E(\log_e \hat{\psi}) \cong \log_e \psi$ og $Var(\log_e \hat{\psi}) \cong \frac{1}{np_{11}} + \frac{1}{np_{12}} + \frac{1}{np_{21}} + \frac{1}{np_{22}}$

$Var(\log_e \hat{\psi})$ kan estimeres ved:

$\hat{Var}(\log_e \hat{\psi}) \cong \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}$.

Dette gir oss følgende tilnærmede 100(1-α) % konfidensintervall for $\log_e \psi$:

$$\left[\log_e \hat{\psi} - z_{\alpha/2}\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}, \log_e \hat{\psi} + z_{\alpha/2}\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}\right].$$

Dersom vi kaller dette konfidensintervallet for $[n, \phi]$, der n er nedre grense og ø er øvre grense, gir det oss konfidensintervallet $[e^n, e^\phi]$ for ψ.

Vedlegg 5: Kji-kvadrattesten til oppgave 2 og 3

Tabulated statistics: Sp2; Sp3

```
Rows: Sp2   Columns: Sp3

            0      1     All

0          22     30      52
        19,55  32,45   52,00

1          31     58      89
        33,45  55,55   89,00

All        53     88     141
        53,00  88,00  141,00

Cell Contents:      Count
                    Expected count


Pearson Chi-Square = 0,782; DF = 1; P-Value = 0,377
Likelihood Ratio Chi-Square = 0,778; DF = 1; P-Value = 0,378
```

Vedlegg 6: Kji-kvadrattesten til oppgave 4 og 5

Tabulated statistics: Sp4; Sp5

Rows: Sp4   Columns: Sp5

	0	1	All
0	61	32	93
	52,11	40,89	93,00
1	18	30	48
	26,89	21,11	48,00
All	79	62	141
	79,00	62,00	141,00

Cell Contents: Count
Expected count

Pearson Chi-Square = 10,141; DF = 1; P-Value = 0,001
Likelihood Ratio Chi-Square = 10,173; DF = 1; P-Value = 0,001

Vedlegg 7: Kji-kvadrattesten til oppgave 6a og 6b

Tabulated statistics: Sp6a; Sp6b

```
Rows: Sp6a   Columns: Sp6b

            0      1     All

0          38      4      42
        12,51  29,49   42,00

1           4     95      99
        29,49  69,51   99,00

All        42     99     141
        42,00  99,00  141,00

Cell Contents:      Count
                    Expected count


Pearson Chi-Square = 105,343; DF = 1; P-Value = 0,000
Likelihood Ratio Chi-Square = 111,828; DF = 1; P-Value = 0,000
```