Bruk av terninger i statistikkundervisning for å øke forståelsen for enkelte terskelbegrep

Med praktiske eksempler fra basisundervisning for tannlegestudenter

Obligatorisk oppgave i basismodulen i pedagogikk (UPED620)
Stein Atle Lie, Institutt for Klinisk Odontologi, UiB

Innledning

Tannlegestudenter på ulike nivå skal ha en viss mengde statistikk undervisning. Grunnen til dette er at de i forbindelse med oppgaver i studiet (masteroppgaver, spesialistkandidat-oppgaver eller en PhD-grad) ofte skal gjennomføre en klinisk studie der bearbeiding av et tallmateriale er en del av oppgaven. Videre vil de ofte i sin kliniske praksis møte nye teknikker, instrument, produkt og materialer der en studie argumenterer for dette nye, ofte ved bruk av statistiske argument. En utfordring for meg som underviser i statistikk er at de fleste studentene ikke har hatt matematikk eller tilsvarende teoretiske fag på mange år. De fleste har ikke hatt dette siden første år på tannlegestudiet, den gang da matematikken var ferskvare fra videregående skole. Å definere et læringsutbytte i et teoretisk fag når studentene ikke (lenger) er vant til det kan være en utfordring.
Statistikk oppstod som et fag rundt terningspill, med en historie tilbake til blant annet Galileo Galilei. Regnestykkene den gang var blant annet å se på de ulike mulighetene ved kast av 3 terninger. Terninger har siden den gang hatt en naturlig plass i introduksjonen til statistikkfaget.
Terninger benyttes ofte i undervisning til å demonstrere begrep som sannsynlighet, utfall og sannsynlighetsfordeling. Dette gjøres ofte uten at terningen fysisk benyttes i selve undervisningen, men at det er en teoretisk betraktning rundt dette. Intensjonen ved bruk av terninger er som regel at de gir håndfaste eksempler for de nevnte terskelbegrepene, på det gjenkjennelige objektet som en terning er. Utfordring er likevel å omsette bruken av terninger til praktiske situasjoner der man observerer ulike størrelser for pasienter, tenner i munnen hos en person, en størrelse målt på et laboratorium eller annet.
Bakgrunnen for at denne oppgaven fokuserer på terskelbegrep er fordi det er mange av dem i løpet av undervisningen i statistikk og det har vist seg at det kan være utfordrende for studentene å få med seg de mest sentrale av disse. Viktigheten av å definere terskelbegrep og hvorledes disse formidles i en undervisningssituasjon er også noe jeg selv mener jeg har hatt utbytte av i basismodulen (se også Meyer og Land 2003).
Utprøvingen i undervisningssituasjon, som er nevnt i denne oppgaven, ble gjennomført høsten 2014 i kurset ODO-STAT1/06.

Noen terskelbegrep

Terninger har som nevnt vært benyttet i lang tid til illustrasjon for noen enkle og sentrale begrep i statistikk. For eksempel introduseres ofte studenter til begrepet **sannsynlighet** ved at en (rettferdig) terning, med 6 sider, der sannsynligheten for et gitt tall (for eksempel «5») vil være 1/6 diskuteres.
Dette gir også en enkelt introduksjon til begrepet **sannsynlighetsfordeling**, der man for eksempel ser på alle mulige utfall for en enkelt terning (1, 2, 3, 4, 5, 6) og sannsynligheten for disse utfallene.
Når man så kaster terningen en gang observerer man **et utfall** for denne sannsynlighetsfordelingen. Således vil man da også definere den enkelte observasjonen (**utfallet**) man gjør i et forsøk.
En sentral størrelse i statistikk er det **empiriske gjennomsnittet**. Det vil si gjennomsnittet av flere observerte **utfall** fra en **sannsynlighetsfordeling.** Gjennomsnittet er en størrelse de fleste ha et begrep om, da det ofte benyttes i dagligtale. Som underviser i statistikk gjennom en årrekke har jeg observert at det for de aller fleste studenter innen kliniske fag, som for eksempel tannleger, vil være en betydelig utfordring å forstå hva de statistiske egenskapene til gjennomsnittet er. Jeg har også gjennomgående inntrykk av at bruken av enkle 6-kantede terninger gir en forståelig innføring til noen begrep, siden terningen er et kjent objekt. Samtidig kan dette gi en barriere når man prøver å dra lærdommen fra bruken av terninger over til eksempler fra en praktisk (for eksempel klinisk) problemstilling. Grunnen til dette er, etter min mening, at studentene ofte blir låst til kjennskapen de har til den klassiske og trivielle terningen og ikke forstår overgangen når man diskuterer data fra praktiske situasjoner.
Et begrep som er av betydelig viktighet innen statistikk er **sentralgrenseteoremet**. I den undervisningen jeg gjennomfører i statistikk for tannlegestudentene vil dette først og fremst dreie seg rundt *hva som skjer med et gjennomsnitt* når antall utfall fra en sannsynlighets-fordeling øker. Overgangen fra de kjente terningene, med 6 sider, til denne abstrakte størrelsen, som så anvendes på praktiske eksempler fra en klinisk setting blir da stor! Ved å utvide bruken av terninger fra bare de kjente terningene med 6 sider, til også uvante terninger, vil studentene i større grad utfordres slik at det ikke blir terningene i seg selv, men prinsippene som demonstreres og diskuteres.

Materiell

Med bakgrunn i undervisningen i basismodulen i pedagogikk, høsten 2014, har jeg gått til innkjøp av klassiske terninger med 6 sider og terninger med 20 sider (fra -9 til +10). For disse terningene er det mange nok til at studentene kan arbeide i par (eller flere i gruppe) med samme problemstilling (bilde 1). I tillegg har jeg anskaffet et sett med terninger med flere ulike fasonger. Disse terningene kan brukes på ulike måter i undervisningen (bilde 2).

Bilde 1: Klassiske terninger med 6 sider (til høyre) og terninger med 20 sider (fra -9 til +10)

Bilde 2: Terninger med ulikt antall sider, fra 4 til 20.

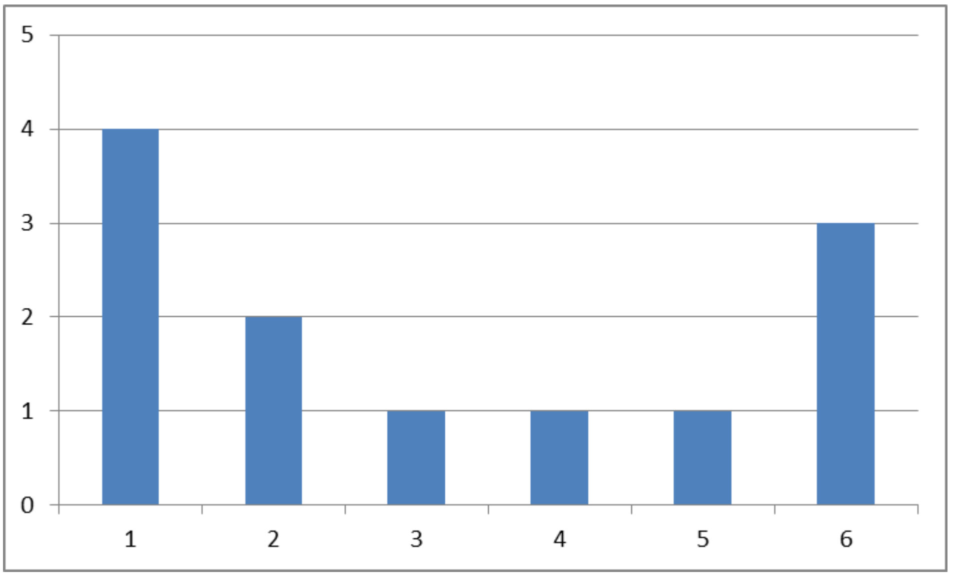
Forslag til praktiske øvelser i undervisningssituasjonen.

I dette kapittelet vil jeg komme med noen konkrete eksempler. Overskriften for eksemplene vil være hvilket terskelbegrep det er meningen at studentene skal forstå. Noen av eksemplene er gjennomført og erfaringer fra disse vil diskuteres senere, mens andre eksempler ikke er gjennomført.

Sentralgrenseteoremet

Vi lar 2 (evt 3) studenter jobbe sammen med 1 tradisjonell terning med 6 sider.
Først skal studentene trille terningen 12 ganger, telle opp hvor mange ganger de ulike sidene på terningen forekommer, og så tegne et stolpediagram for dette. Teoretisk skal det nå være 2 forekomster for hver side av terningen, men det vil selvsagt (pga tilfeldigheter) ikke være slik, men stolpediagrammet vil i større eller mindre grad ha en flat (uniform) fasong (se eksempel i figur 1)

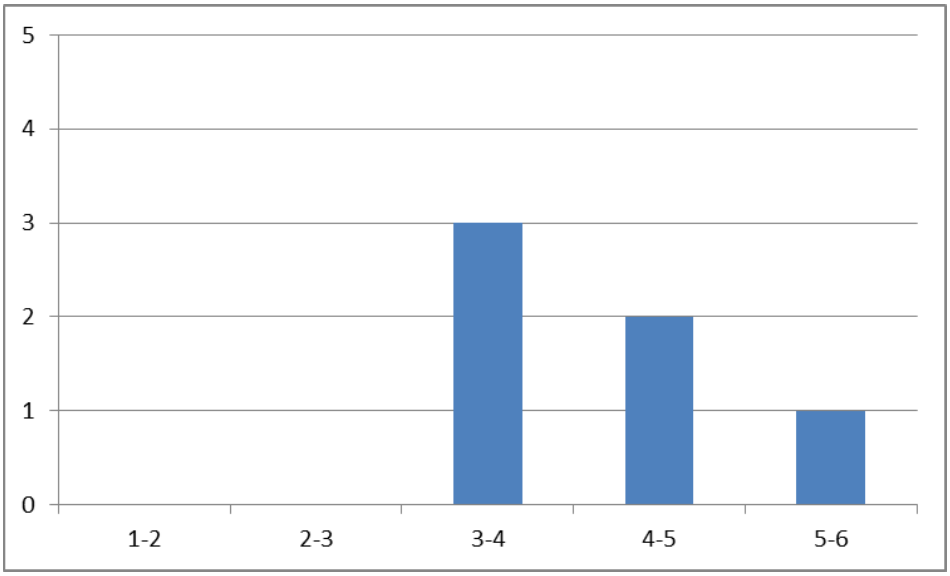
Figur 1: Eksempel på forekomst av antall sider på en standard terning med 6 sider, når terningen kastes 12 ganger.



Når studentene er ferdig med denne delen blir de utfordret til å diskutere forskjellen mellom denne figuren og den teoretiske figuren (der antallet av alle sidene skulle forekommet eksakt 2 ganger). De fleste studentene bør ved denne delen av øvelsen skjønne forskjellen mellom den teoretiske fordelingen og den fordelingen de observerer ved faktisk å trille terningen.

Nå skal studentene igjen trille terning 12 ganger, men denne gangen skal de trille terningen først 2 ganger, så skal de regne gjennomsnittet av de to kastene (f.eks (2+5)/2=3.5). Til sammen vi de nå altså få 6 gjennomsnitt. Så skal de igjen lage et stolpediagram for de 6 verdiene de nå har fått (se eksempel i figur 2).

Figur 2: Eksempel på gjennomsnitt for kast med standard terning (med 6 sider) 2 ganger når dette gjentas 6 ganger. Terningen kastes da totalt 12 ganger.

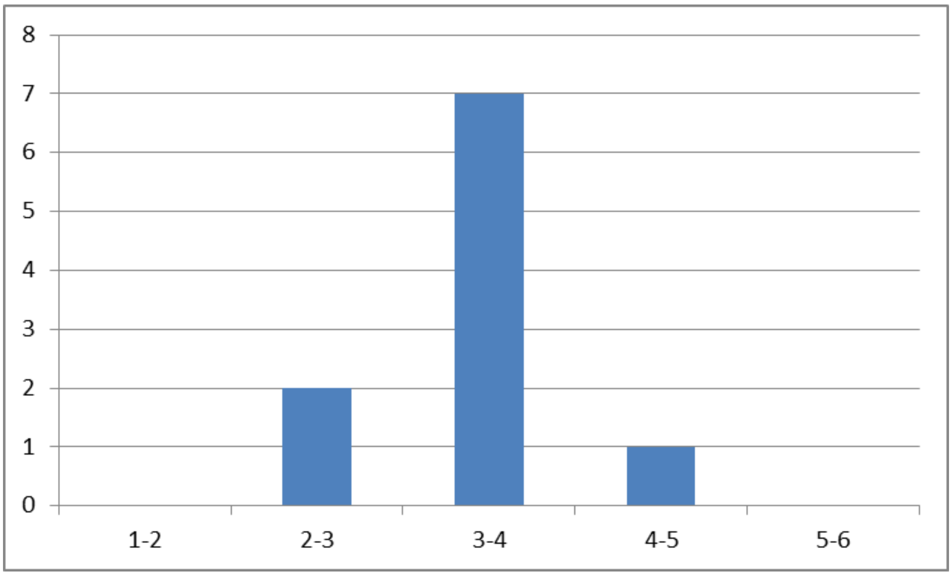


Etter denne delen av øvelsen skal studentene igjen diskutere. Hensikten her er for studentene å oppdage at figuren de får i større grad vil ha en topp, eller høyest stolpe, mot midten. Sjansen for å få tallet 1 eller 6 er nå liten (begge terningene må da enten være 1 eller 6). Tallet 3.5 vil det derimot være størst sjanse for å få.

Hittil bør studentene ha fått med seg forskjeller mellom den teoretiske antagelsen for 1 terning (dvs at alle sider har like stor sannsynlighet) og det man observerer når man faktisk kaster en terning. Videre bør de ha observert at hvis man tar gjennomsnitt av 2 kast med 1 terning, så vil det skje noe med symmetrien og det vil være en tendens til en topp mot midten av figuren de har laget.

Som siste del av denne øvelsen fikk jeg nå studentene i par til å kaste terning 10 ganger. Den ene studenten kastet og den andre skrev ned resultatet. De skulle nå regne gjennomsnittet av de 10 kastene. Tallet 10 ble valgt, fordi det da er enkelt å regne gjennomsnitt (det er da summen av alle tallene de har fått dividert med 10). Denne gangen skal de rapportere tallene til foreleser som lager en felles figur på tavlen. Høsten 2014 hadde jeg kun 6 studenter, mens til vanlig er det rundt 20 studenter. Figuren i eksempelet (figur 3) tar derfor utgangspunkt i at det er 10 par av studenter. For høsten 2014 fikk jeg studentene til å gjøre forsøket 2 ganger, slik at figuren da ble basert på 6 rapporterte tall til foreleser.

Figur 3: Eksempel på gjennomsnitt for 10 kast med en standard (med 6 sider) når dette gjentas 6 ganger.



Etter denne delen av øvelsen ble det i plenum diskutert hvorfor denne figuren har en markert topp mot midten. Sannsynligheten for å observere gjennomsnitt som er 1 eller 6 er nå svært liten (alle de 10 kastene med terningene måtte da i så fall blitt 1 eller 6). Tallet 3.5 (eller noe i nærheten) vil det derimot være stor sannsynlighet for å få.

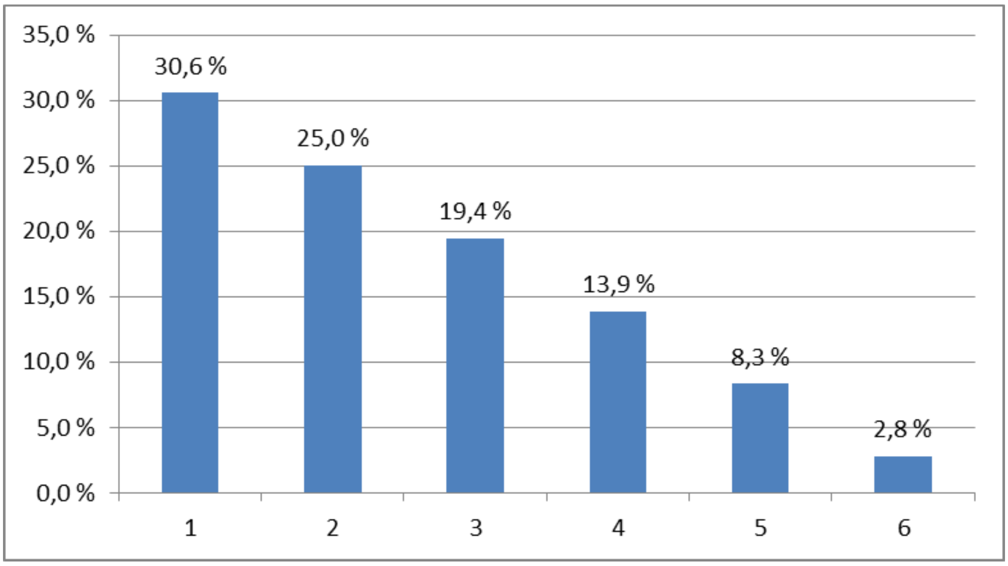
Forslag til andre gruppebaserte øvelser under forelesning

For kliniske data er det ofte slik at de er skjevfordelt. For eksempel vil antall hull i tenner for en gruppe barn ofte være «0», noen barn vil ha 1 hull, enda færre vil ha 2 osv. Således er standard betraktning for en vanlig terning dårlig egnet i undervisning, siden sannsynlighetsfordelingen for denne er uniform (det vil si at alle sider ha samme sannsynlighet). For at studentene skal oppdage at denne typen skjeve fordelinger ikke bare er noe man observerer i reelle data, men også noe man relativt enkelt kan teoretisk konstruere kan man sette studenter sammen i grupper (2 eller fler), gi dem 2 terninger og se om de klarer å konstruere en slik skjev fordeling med å kaste disse.
Den enkleste løsningen på dette er at studentene kaster de 2 terningene og så velger den terningen som har færrest antall øyne. Valgene de da vil gjøre vil være slik som skissert i tabell 1 og den teoretiske fordelingen for dette er vist i figur 4.

Tabell 1: Minimum antall øyne, når man kaster 2 terninger med 6 sider på hver terning.

	Terning-2					
Terning-1	**1**	**2**	**3**	**4**	**5**	**6**
1	1	1	1	1	1	1
2	1	2	2	2	2	2
3	1	2	3	3	3	3
4	1	2	3	4	4	4
5	1	2	3	4	5	5
6	1	2	3	4	5	6

Figur 4: Den teoretiske fordelingen fra kast med 2 terninger, når man for hvert kast velger den terningen med færrest antall øyne.



Hypotesetesting

For terningene med ulikt sett av antall øyne gjennomførte jeg også i løpet av undervisningen en øvelse der 2 studenter fikk velge ut en terning hver. De skulle da ikke vise denne terningen til meg. Så skulle de trille terningen 10 ganger og oppgi de 10 resultatene (**utfallene**) de fikk fra dette. Idéen her var todelt. Først var det å demonstrere med bruk av terninger at man som forsker ikke vet hvilken bakenforliggende fordeling dataene man observerer har. Det vil si at jeg som foreleser ikke vet hvilke terninger de 2 studentene har valgt.
Dernest var idéen å regne gjennomsnitt for resultatene fra de to studentene. Så ble det diskutert på bakgrunn av de to gjennomsnittene og den erfaring man nå hadde opparbeidet gjennom de foregående øvelsene egenskapene til disse to gjennomsnittene. Gjennomsnittene vil i seg selv nå være normalfordelt (dvs symmetrisk fordelt) og gi et uttrykk for sentraltendensen for den terningen som lå til grunn for kastene som ble gjort. Så ble den statistiske størrelsen (som studentene hadde fått presentert tidligere i kurset) beregnet. Dette ble gjort for at det til slutt kunne gjennomføres den statistiske testen t-test. Ved denne testen

vil jeg som forsker anta at de to terningene som studentene har valgt er like. Basert på de 10 kastene, som studentene har gjort, gjøres da disse beregningene (jeg valgte å gjøre dette med Excel på lerret, for å ikke bruke for mye tid til selve beregningene på tavlen). Testen gav meg da en sannsynlighet for om de to terningene studentene hadde valgt var like. Siden den sannsynligheten var stor antok jeg at terningene måtte være like.
De to terningene studentene hadde valgt var en terning med 18 sider og en terning med 20 sider. Det var altså ikke to helt like terninger, men terninger som var relativt like…

Personlige erfaringer

Som underviser i innføringskurs i statistikk gjennom flere år har det alltid vært en utfordring å engasjere studentene og gi dem eksempler og oppgaver der de kan koble det de lærer med den praktiske situasjonen de har, enten dette er i en vitenskapelig eller en klinisk situasjon. De fleste lærebøker i statistikk gir en innføring i de teoretiske (bla de matematiske) sidene ved statistikkfaget før de gir eksempler fra praktiske situasjoner. For svært mange studenter er det en utfordring å koble det de har lært fra det teoretiske med de praktiske eksemplene. Dette betyr nødvendigvis ikke at studentene ikke har lært det de skulle lære, men at det er et gap mellom forutsetningene som legges når man underviser statistikk og hvorledes studentene tar dette med seg når de for praktiske oppgaver, enten på kurset eller seinere. Læringsmålet kan sånn sett være veldefinert og studentene kan ha oppnådd dette, men det kan likevel være en avstand til hvorledes man ønsker at studentene skal bruke det de har lært videre. For studenter innen anvendte og kliniske felt trenger derfor de vanlige læringsmålene ved et grunnkurs i statistikk ikke å være de riktige læringsmålene.
Øvelsene der studentene sitter i grupper og triller terning og diskuterer figurene de lager og hvorledes disse endres med økende antall kast for et gjennomsnitt, syns jeg fungerte godt etter intensjonen og studentene gav gode tilbakemeldinger på disse øvelsene. Læringsmålet synes sånn sett klart å definere og terskelbegrepene er relativt enkle å forklare.
Når det gjelder gjennomføringen av hypotesetestingen der 2 studenter skulle trille skjulte terninger og jeg som foreleser regner på tavlen erfarte jeg at det ble en for lang tankerekke og vanskelig for studentene å henge med på hele resonnementet.

Diskusjon

I innførende emner i statistikk, særlig for studenter innen anvendte og praktiske fag slik tannlegestudiet og medisin er, kan praktiske øvelser som gir en overføringsverdi til reelle data som ligner det studentene senere vil få befatning med være betydningsfullt. Terninger er i så måte egnet, siden de kan benyttes direkte til «datainnsamling» i løpet av undervisningen. På den annen side kan det være utfordrende for studenter å se sammenhenger med de trivielle terningene og data fra en situasjon der data fra pasienter eller et labbforsøk skal samles inn. I tidsskriftet *Journal of Statistics Education* er det flere eksempler på bruk av ulike kjente gjenstander i en undervisningssituasjon. For eksempel viser Schwartz et mulig bruk av M&M sjokolade for å gi en innføring i den statistiske metoden variansanalyse (ANOVA) (Schwartz 2013). Artikkel *Team-Based Learning in a Statistical Literacy Class* i samme tidsskrift viser

også bruken av gruppe basert undervisning for begynnerkurs i statistikk. Gruppebaserte undervisningsmetoder må sies å være noe uvanlig i generell undervisning i statistikk (St. Clair og Chihara 2012) bortsett fra når det kommer til regneøvelser som ofte går parallelt med undervisningen. Dette har nok historiske årsaker, men er også kanskje noe overraskende for noen likevel, særlig siden flere har vist at aktivisering av studentene, for eksempel gjennom gruppebaserte øvsler, kan gi et økt læringsutbytte (se for eksempel Biggs 1999).
Bruken av terninger i statistikkundervisning er utbredt, men ofte benyttes det mer som en teoretisk betraktning og ikke direkte med fysiske terninger. I denne oppgaven har jeg prøvd ut enkelte muligheter for å bruke terninger mer aktivt og mer variert for å få studentene til å få en mer aktiv forståelse og oppdage noen grunnleggende terskelbegrep i statistikk.

Referanser

Biggs J. (1999) What the Student Does: teaching for enhanced learning. Higher Education Research & Development, 18: 1, 57 — 75

Meyer, J.H.F. og Land, R. (2003) Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practicing within the disciplines. Occasional Report 4, School of Education, University of Edinburgh

Schwartz T.A. (2013) Teaching Principles of One-Way Analysis of Variance Using M&M's Candy Journal of Statistics Education, Volume 21, Number 1

St. Clair K og Chihara L. (2012) Team-Based Learning in a Statistical Literacy Class. Journal of Statistics Education, Volume 20, Number 1