Betinget sannsynlighet, Bayes' setning og hypergeometrisk sannsynlighetsmodell

Av Marie Schau og Elisabeth Syrstad

MAT3010

13.05.2015

Innhold

1 Innledning

Vi vil se nærmere på deler av kompetansemålet «kombinatorikk og sannsynlighet» fra læreplanen i matematikk R1 for videregående skole. Dette er stoff mange elever synes er vanskelig, og det kan være utfordrende å finne forklaringsmodeller innenfor sannsynlighetsregning som appellerer til elevenes intuitive forståelse. Svarene man får oppleves ofte som overraskende og kontraintuitive. Dette må læreren være klar over når undervisningen skal legges opp.

Oppgaven er delt i to deler. I den første delen vil hovedvekten være på betinget sannsynlighet og Bayes` setning. Den andre delen omhandler hypergeometrisk sannsynlighetsmodell.

2 Betinget sannsynlighet og Bayes' setning

Ifølge første del av kompetansemålet skal eleven kunne: «gjøre rede for begrepene uavhengighet og betinget sannsynlighet, og utlede og anvende Bayes' setning på to hendelser» (1).

2.1 Sannsynlighet

Det kan være greit å starte med å definere hva vi mener med sannsynlighet:

«*Probability P* is a numerical valued function that is defined on the events in a sample space *S* and that satisfies the following conditions:

(i) $0 \leq P(A) \leq 1$ for all events A
(ii) $P(S) = 1$
(iii) For disjoint events $A_1, A_2, \ldots\ldots, A_n$

$P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = P(A_1) + P(A_2) + \ldots\ldots + P(A_n)$ » (3) (der ∪ leses «union)

For å oppfylle komplementsregelen, merker vi oss at A og A^C er to disjunkte hendelser, og at $A \cup A^C = S$. Vi har at

$P(A \cup A^C) = P(A) + P(A^C)$
$P(S) = 1$
$P(A^C) = 1 - P(A)$

De følgende tre reglene/lovene for sannsynlighet gjelder for S:

$P(A) = 1 - P(A^C)$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ kalles også for den generelle addisjonssetningen
$P(A \cap B) = P(A|B) \cdot P(B)$ kalles også for den generelle produktsetningen,
der ∩ leses «snitt» og P(A|B) leses «A gitt B».

Disse reglene er viktig verktøy vi trenger når vi skal regne, og som vi vil bruke når vi senere skal se på Bayes' setning.

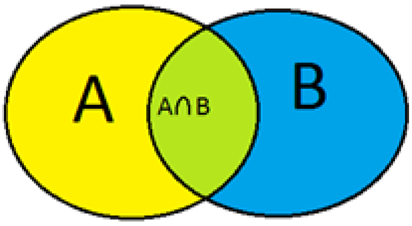
2.2 Betinget sannsynlighet

Ofte kan vi oppleve at sannsynligheten til en hendelse B må modifiseres etter at vi får informasjon om at en relatert hendelse A har inntruffet eller ikke. Sannsynligheten P(B|A) er sannsynligheten for at hendelsen B skal inntreffe når vi vet at hendelsen A har inntruffet. Vi kan omforme den generelle produktsetningen for å få følgende uttrykk:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad \text{der } P(A) > 0$$

Bevis (2) for betinget sannsynlighet:

Vi forutsetter at P(A) > 0, og vi regner med en uniform sannsynlighetsmodell. Der er alle utfallene like sannsynlige. Her forutsetter vi også at vi har et endelig utfallsrom S.



Figur 1, Venndiagram

La A og B være to hendelser i et utfallsrom S med N mulige utfall. A består av N(A) utfall, og B består av N(B) utfall. Det er N(A ∩ B) utfall i A ∩ B. Når vi vet at hendelsen A har inntruffet, står vi igjen med N(A) mulige utfall. Det er nå bare utfallene i A ∩ B som gir hendelsen B. Dermed er det N(A ∩ B) gunstige utfall for hendelsen B. Da er

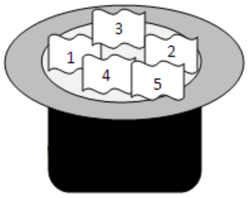
$$P(A) = \frac{N(A)}{N(S)} \qquad \text{og} \qquad P(A \cap B) = \frac{N(A \cap B)}{N(S)}$$

$$P(B|A) = \frac{\text{tallet på gunstige utfall}}{\text{tallet på mulige utfall}} = \frac{\frac{N(A \cap B)}{N(S)}}{\frac{N(A)}{N(S)}} = \frac{N(A \cap B)}{N(A)}$$

Vi dividerer med N i teller og nevner. Det gir

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad \text{forutsatt at } P(A) > 0$$

Eksempel (4) på betinget sannsynlighet:



I dette eksemplet skal vi se på hvordan opplysninger om at en hendelse har inntruffet eller ikke vil innvirke på sannsynligheten til en relatert hendelse. Vi tenker oss at Per og Kari skal trekke hver sin lapp fra en krukke som inneholder fem lapper med tallene 1 til 5. De kan enten trekke et partall eller et oddetall. De trekker en gang hver, og Per skal trekke få trekke først. Vi definerer hendelsene:

A: Per trekker en lapp med partall
B: Kari trekker en lapp med partall

Per trekker først: P(A)= $\frac{2}{5}$

Når Kari så skal trekke, vil resultatet Per fikk påvirke resultatet hun får:

Dersom Per trakk et partall: P(B) = $\frac{1}{4}$
Dersom Per trakk et oddetall: P(B) = $\frac{2}{4} = \frac{1}{2}$

Vi har at $P(B|A) = \frac{1}{4}$ og at $P(B|A^C) = \frac{1}{2}$

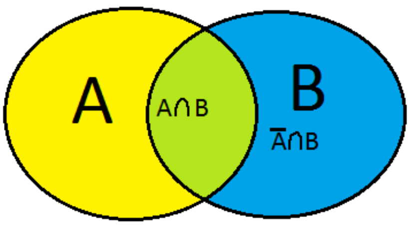
Hva er så sannsynligheten for at både Per og Kari trekker et partall?

$P(A \cap B) = P(A) \cdot P(B|A) = \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}$

Vi får det samme resultatet dersom Kari er den første som trekker, fordi $P(A \cap B) = P(B \cap A)$

2.3 Total sannsynlighet

Bevis (2) for total sannsynlighet:



Figur 2, Venndiagram

Vi lar A og B være to hendelser. Hendelsen B kan vi dele i to deler. $A \cap B$ er utfallene i B som også er med i A. $A^C \cap B$ er utfallene i B som ikke er med i A. Til sammen blir dette hele B.

$B = (A \cap B) \cup (A^C \cap B)$

De to delene $A \cap B$ og $A^C \cap B$ har ingen felles utfall. Derfor kan vi finne sannsynligheten for B ved å summere sannsynligheten for de to delene. Vi får

$P(B) = P(A \cap B) \cup (A^C \cap B) = P(A \cap B) + P(A^C \cap B)$

Ifølge den generelle produktsetningen er $P(A \cap B) = P(A) \cdot P(B|A)$ og $P(A^C \cap B) = P(A^C) \cdot P(B|A^C)$. Dermed er den totale sannsynligheten for hendelsen B gitt ved følgende formel:

$P(B) = P(A) \cdot P(B|A) + P(A^C) \cdot P(B|A^C)$

Eksempel (2):

I et lotteri er $\frac{1}{3}$ av loddene røde. Sannsynligheten for å vinne på et rødt lodd er $\frac{1}{10}$. Sannsynligheten for å vinne på et av de andre loddene er $\frac{1}{12}$. Hva er da sannsynligheten for å vinne på et tilfeldig valgt lodd?

Vi får disse hendelsene:

R: Loddet er rødt
G: Loddet gir gevinst

$P(R) = \frac{1}{3}$ og $P(G|R) = \frac{1}{10}$ og $P(G|R^C) = \frac{1}{12}$ og $P(R^C) = 1 - P(R) = 1 - \frac{1}{3} = \frac{2}{3}$

Da er den totale sannsynligheten for gevinst:

$P(G) = P(R) \cdot P(G|R) + P(R^C) \cdot P(G|R^C) = \frac{1}{3} \cdot \frac{1}{10} + \frac{2}{3} \cdot \frac{1}{12} = \frac{1}{30} + \frac{1}{18} = \frac{3}{90} + \frac{5}{90} = \frac{8}{90} = \frac{4}{45}$

2.4 Bayes' setning

Thomas Baye (1702-1761) var en engelsk prest som regnes som opphavsmannen til Bayes` setning (eller Bayes teorem). Bayes` statistikk er mye brukt innenfor områder som farmasi, medisin, og markedsføring. Innenfor disse fagområdene er det vanlig at det foretas en subjektiv utvelgelse av hvilke hypoteser man ønsker å gå videre med. En farmasøyt skal for eksempel utvikle en ny medisin. Det er flere kjemiske substanser å velge mellom, og det må tas et valg basert på et faglig skjønn når det bestemmes hvilke ingredienser som skal testes videre i utviklingen av medisinen. Det kan være så små numeriske forskjeller mellom de ulike substansene at man ikke kan ta et objektivt valg kun basert på målbare data. Innenfor markedsføring kan valget av hvilke farger eller bilder som skal brukes i en kampanje baseres på synsing eller følelser, først etter en evaluering av kampanjen kan man bedømme om valget var fornuftig.

Sannsynligheten som settes eller velges på en hypotese før man gjør et forsøk eller innhenter informasjon kalles en «prior» -sannsynlighet. Etter at forsøk et utført eller mer utfyllende informasjon er innhentet, justeres sannsynligheten for hypotesen til en «posterior» -sannsynlighet.

En retning innenfor statistikk- og sannsynlighets-regning kalles derfor for subjektivistisk Bayesisk sannsynlighet. Det er ikke kun objektive, målbare data man baserer seg på, men subjektive valg foretatt underveis i en prosess.

Bevis (2, 3) for Bayes' setning :

Her vises to måter å bevise Bayes' setning på. Først en litt enklere versjon der vi tar utgangspunkt i at $P(A \cap B) = P(B \cap A)$:

Vi har at

$P(A \cap B) = P(B \cap A)$

$P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

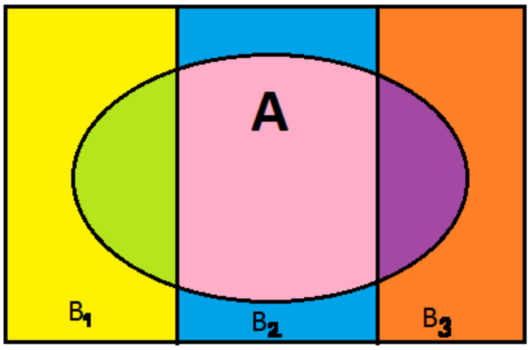
$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$ der $P(A) > 0$

Så viser vi en noe mer avansert metode:

Vi vil finne en formel for $P(B_1|A)$, og skriver

$P(B_1|A) = \frac{P(A \cap B1)}{P(A)}$

A er unionen av de parvis disjunkte hendelsene $(A \cap B_1)$, $(A \cap B_2)$ ……. $(A \cap B_k)$.



Figur 3, Illustrasjon av strukturen av Bayes' setning, der k=3 (k er tilfeldig valgt som et eksempel).

Da får vi at

$P(A) = P(A \cap B_1) \cup P(A \cap B_2) \cup ... \cup P(A \cap B_k) = P(A \cap B_1) + P(A \cap B_2) + ... + P(A \cap B_k)$

$= P(B_1 \cap A) + P(B_2 \cap A) + ... + P(B_k \cap A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + ... + P(B_k) \cdot P(A|B_k)$

$= \sum_{j=1}^{k} P(Bj) \cdot P(A|Bj)$

Hvert ledd i denne summen kan skrives som

$P(A \cap B_j) = P(B_j)\ P(A|B_j)$

Nå kan vi skrive Bayes' setning på generell form:

$P(B_1|A) = \frac{P(B1) \cdot P(A|B1)}{\sum_{j=1}^{k} P(Bj) \cdot P(A|Bj)}$

Eksempel 1 (4) på bruk av Bayes' setning:

En skole har 840 elever, 360 jenter og 480 gutter. 5 jenter er fargeblinde, 34 gutter er fargeblinde. Vi trekker tilfeldig en elev ved skolen, og denne eleven er fargeblind. Hva er sannsynligheten for at denne eleven i tillegg er jente?

Vi har følgende hendelser:

J: Eleven er jente
F: Eleven er fargeblind

Her ønsker vi å finne $P(J|F)$. Da må vi først sortere opplysningene opplysningen vi har fått.

$P(J) = \frac{360}{840}$ $\quad$ $P(F) = \frac{39}{840}$ $\quad$ $P(F|J) = \frac{5}{360}$

$P(J|F) = \frac{P(F|J) \cdot P(J)}{P(F)} = \frac{\frac{360}{840} \cdot \frac{5}{360}}{\frac{39}{840}} = \frac{5}{39}$

Vi kan sette opp en krysstabell, som en kontroll for at vi har regnet riktig:

	Jenter	**Gutter**	**Sum**
Fargeblinde	5	34	39
Normalt fargesyn	355	446	801
Sum	360	480	840

Vi finner sannsynligheten ved å lese av i tabellen hvor mange fargeblinde jenter vi har, og hvor mange elever som er fargeblinde. Disse tallene settes så inn i formelen

P(fargeblind jente) = $\frac{\text{gunstige (antall fargeblinde jenter)}}{\text{mulige (antall elever som er fargeblinde)}} = \frac{5}{39}$. Vi ser at dette stemmer med det vi regnet ut ovenfor.

Eksempel 2 (4) på bruk av Bayes' setning:

Vi antar at 5 % av befolkningen lider av en sykdom. Vi bruker en test som gir utslag for 96 % av dem som er syke, og 10 % av dem som er friske. Hva er sannsynligheten for at en person som tester positivt på testen, virkelig er syk?

Ser vi på 1000 personer, kan vi anta at 50 er syke, og 950 er friske. Testen vil gi utslag for $50 \cdot 0{,}96 = 48$ personer som er syke. Den vil gi utslag for $950 \cdot 0{,}10 = 95$ personer som er friske.

Det betyr at av $48 + 95 = 143$ personer som får positivt utslag på testen, har kun 48 sykdommen.

Her passer det godt å bruke Bayes` setning. Vi definerer de to hendelsene vi trenger:

S: En person har sykdommen
U: Testen gir positivt utslag

$P(S) = 0{,}05$ $P(U|S) = 0{,}96$ $P(U|S^C) = 0{,}10$ $P(S^C) = 1 - P(S) = 0{,}95$

Vi må også finne sannsynligheten for at testen slår ut på en tilfeldig person:

$P(U) = P(S \cap U) + P(S^C \cap U)$
$P(U) = P(S) \cdot P(U|S) + P(S^C) \cdot P(U|S^C)$
$P(U) = 0{,}05 \cdot 0{,}96 + 0{,}95 \cdot 0{,}10 = 0{,}143$

$$P(S|U) = \frac{P(U|S) \cdot P(S)}{P(U)} = \frac{0{,}05 \cdot 0{,}96}{0{,}143} = 0{,}336$$

Det er altså bare 33,6% sjanse for at en person som tester positivt på testen faktisk har sykdommen.

Eksempel 3 (5) på bruk av Bayes' setning:

En taxi er involvert i en ulykke om natten. I byen der ulykken skjer, har alle taxier enten fargen blå eller fargen grønn. Vi får oppgitt følgende:

85% av taxiene har fargen grønn, og 15% har fargen blå. Et øyenvitne identifiserer ulykkestaxien til å være blå. Påliteligheten til øyenvitnet ble testet under tilsvarende forhold som ulykkesnatten. Da så vitnet korrekt farge på bilen i 80 % av tilfellene, og feil farge i 20 % av tilfellene.

Hva er sannsynligheten for at taxien som var involvert i ulykken virkelig var blå?

Psykologer har funnet at mange spurte personer mener at det er omtrent 80% sjanse for at øyenvitnet har sett riktig. Vi skal bruke Bayes` setning for å finne det korrekte svaret. Vi har en «prior» -sannsynlighet for at taxien er blå på 15 %.

Vi definerer følgende hendelser:

A: Ulykkestaxien er blå
B: Øyenvitne sier at ulykkestaxien er blå

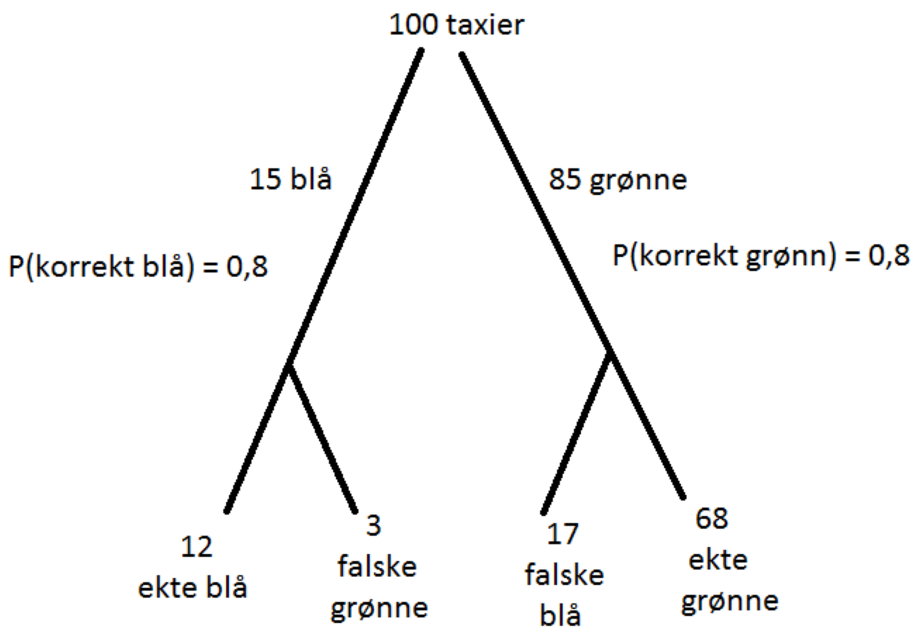
Vi har at $P(A) = 0{,}15$ og $P(B|A) = 0{,}80$

$P(B) = P(A) \cdot P(B|A) + P(A^C) \cdot P(B|A^C) = 0{,}15 \cdot 0{,}80 + 0{,}85 \cdot 0{,}20 = 0{,}29$

$P(A|B) = \frac{0{,}80 \cdot 0{,}15}{0{,}29} = 0{,}41$

Vi får nå en «posterior» -sannsynlighet på 41 %. Et betydelig høyere tall enn de 15 % vi hadde før utregningen, og et betydelig lavere tall enn det mange ville antatt på forhånd.

Et valgtre kan være til hjelp for å skaffe oss en oversikt over problemstillingen. Når vi tar $\frac{\text{gunstige}}{\text{mulige}} = \frac{12}{29}$ = 0,41, ser vi at vi får det samme svaret som ovenfor.



Figur 4, Valgtre som viser Bayes' setning på eksempel med blå og grønne taxier.

Eksempel 4 (6) på bruk av Bayes' setning:

«Bayes goes to Hollywood»
Dette eksemplet er tatt med for å vise hvor viktig det er å lese statistikk og tall med et kritisk blikk, og ikke la seg villede av tilsynelatende overbevisende tall-argumenter.

OJ. Simpson er en amerikansk fotballspiller og skuespiller bosatt i Hollywood. I 1994 ble han arrestert, mistenkt for å ha drept sin kone Nicole Brown Simpson. Nicoles blod ble funnet i huset der de bodde, i bilen hans og på sokkene hans. Påtalemakten fokuserte på at OJ. Simpson var voldelig, og at han tidligere hadde mishandlet sin kone. Forsvarsadvokaten argumenterte for at tidligere voldsbruk var irrelevant for saken. De presenterte følgende statistikk for å underbygge dette:

4 millioner kvinner i USA blir utsatt for vold av ektemannen i året. Og videre: 1432 kvinner (i 1992-tall) ble drept av sine ektemenn. Av ektemenn som slår kona, er det bare 1 av 2500 som dreper henne ($\frac{1432}{4\,000\,000} \approx \frac{1}{2500}$).

Regningen er riktig, og resonnementet virker klart. Kunnskap om betinget sannsynlighet og Bayes` setning hjelper oss til å innse hvorfor dette er tullete regning.

Forsvarsadvokaten regner ut P(drept|utsatt for vold av ektemannen), men vi lurer ikke på om Nicole ble drept. Det vet vi fra før. Vi lurer heller ikke på om hun ble utsatt for vold av ektemannen. Det vi er interessert i å vite er dette: Hva er sannsynligheten for at du blir drept av mishandleren når du blir drept?

I USA er stistikken klar: 90 % av drepte kvinner som har blitt mishandlet, blir drept av mishandleren.

2.5 Uavhengighet

Vi må være oppmerksomme på situasjoner der den betingede sannsynligheten $P(A|B)$ viser seg å bli den samme som den uavhengige sannsynligheten $P(A)$. Da vil informasjon om forekomsten av hendelse B ikke innvirke på fastsettingen av sannsynligheten av hendelse A. Når vi har at $P(A|B) = P(A)$, sier vi at hendelsene A og B er uavhengige.

To hendelser A og B er uavhengige dersom:

$P(A|B) = P(A)$ der $P(A) > 0$ og $P(B) > 0$.

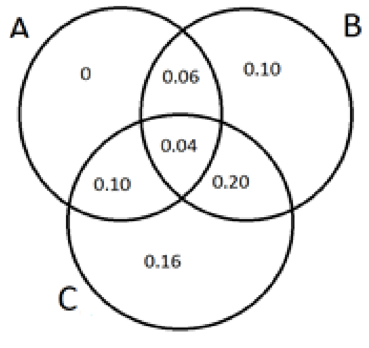
$P(A|B) = \frac{P(A \cap B)}{P(B)}$, da er $P(A|B) = P(A)$ ekvivalent med $P(A \cap B) = P(A) \cdot P(B)$, som viser at definisjonen av uavhengighet er symmetrisk i A og B.

Vi har videre at

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A) \cdot P(B)}{P(A)} = P(B)$$

Uavhengighet for tre hendelser:

Vi må være spesielt oppmerksomme når vi ser på flere enn 2 hendelser, og skal vurdere uavhengighet. Det er ikke sikkert at hendelsene er parvis uavhengige. Vi ser på et eksempel:



Figur 5, Venndiagram, vurdering av uavhengighet ved 3 ulike hendelser

Selv om vi har at $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) = 0{,}04$, er ikke de andre hendelsene parvis uavhengige:

$P(A \cap B) = 0.10$, men $P(A) \cdot P(B) = 0{,}20 \cdot 0{,}40 = 0{,}08$

$P(B \cap C) = 0.24$, men $P(B) \cdot P(C) = 0{,}40 \cdot 0{,}50 = 0{,}20$

$P(C \cap A) = 0.14$, men $P(C) \cdot P(A) = 0{,}50 \cdot 0{,}20 = 0{,}10$

Vi har at:

De n hendelsene $\{A_1, A_2, \ldots, A_n\}$ er gjensidig uavhengig hvis og bare hvis det for hver k ($k = 2,3,\ldots,n$) og hver delmengde $\{ i_1, i_2, \ldots, i_n \}$ av k distinkte verdier trukket fra mengden av de n første naturlige tallene,

$$P(A_{i1} \cap A_{i2} \cap \ldots \cap A_{ik}) = P(A_{i1}) \cdot P(A_{i2}) \cdot \ldots \cdot P(A_{ik}) \qquad (7)$$

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ er derfor en nødvendig, men ikke tilstrekkelig betingelse for at 3 hendelser skal være gjensidig uavhengig.

3 Hypergeometrisk sannsynlighetsmodell

I kompetansemålene for «Kombinatorikk og sannsynlighet» matematikk R1 i «Læreplan i matematikk for realfag – programfag i studiespesialiserende utdanningsprogram» står det:

Mål for opplæringen er at elevene skal kunne:

- drøfte kombinatoriske problemer knyttet til ordnede utvalg med og uten tilbakelegging og uordnede utvalg uten tilbakelegging, og bruke dette til å utlede regler for beregning av sannsynlighet

Forklaring av den hypergeometriske fordelingen

Den hypergeometriske fordelingen har følgende karakteristiske trekk:

- En populasjon bestående av N elementer inneholder a elementer med en spesiell egenskap
- Man foretar n trekninger uten tilbakelegging. Det vil si at sannsynligheten endrer seg for hvert trekk man tar
- x er antall elementer man trekker med den spesielle egenskapen

$$P(X = x) = \frac{\binom{a}{x} \cdot \binom{N-a}{n-x}}{\binom{N}{n}}$$

Eksempel

Du har en kortstokk med 52 kort. Av disse er 13 hjerterkort. Du får helt tilfeldig utdelt 13 kort fra bunken. Hva er sannsynligheten for at det er nøyaktig 5 hjerterkort blant de 13 kortene?

$$P(X = 5) = \frac{\binom{13}{5} \cdot \binom{52-13}{13-5}}{\binom{52}{13}} = \frac{\binom{13}{5} \cdot \binom{39}{8}}{\binom{52}{13}} = \frac{1287 \cdot 61523748}{635013559600} = 0{,}125$$

3.1 Hypergeometrisk fordeling med flere variable

En kan også ha hypergeometrisk fordeling i en populasjon hvor det er flere variable med en bestemt egenskap. Formelen for den hypergeometriske fordelingen med tre variable, hvor hele populasjonen har spesielle egenskaper, blir seende slik ut (8):

$N = a + (N - a)$

$a = x + (n - x)$

$N = n_1 + n_2 + n_3$

$n = x_1 + x_2 + x_3$

$$P(x_1, x_2, x_3) = \frac{\binom{n_1}{x_1} \cdot \binom{n_2}{x_2} \cdot \binom{n_3}{x_3}}{\binom{N}{x}}$$

Eksempel

Et elevråd på en videregående skole består av 15 elever. Elevene har 6 representanter fra VG1, 4 representanter fra VG2 og 5 representanter fra VG3. De skal trekke 6 elever som skal være med å arrangere OD-dagen. Hva er sannsynligheten for at de 6 elevene vil bestå av to elever fra hvert klassetrinn?

$$P(2\ elever\ fra\ hvert\ klassetrinn) = \frac{\binom{6}{2}\cdot\binom{4}{2}\cdot\binom{5}{2}}{\binom{15}{6}} = \frac{15\cdot 6\cdot 10}{5005} = 0{,}18$$

3.2 Bevis for at den hypergeometriske går mot den binomiske fordelingen

Når populasjonen blir stor, går den hypergeometriske fordelingen mot den binomiske fordelingen. For at man skal kunne bruke den binomiske fordelingen må forsøkene være uavhengige av hverandre og sannsynligheten for at en hendelse skal inntreffe må være lik i alle forsøk. Når populasjonen, N, er stor i forhold til antall trekninger, n, blir denne sannsynligheten tilnærmet lik sannsynligheten for verdien x i en binomisk fordeling, $p_x(k) = \binom{n}{k}p^k(1-p)^{n-k}$, med $p = \frac{a}{N}$ (9).

En kan også bevise dette med grenseverdien når N går mot uendelig (10).

La

$$f(x) = \frac{\binom{a}{x}\cdot\binom{N-a}{n-x}}{\binom{N}{n}}$$

med $p = \frac{a}{N}$ bevis at

$$\lim_{N\to\infty} f(x) = \binom{n}{x}p^x(1-p)^{n-x}$$

$$f(x) = \frac{\binom{a}{x}\cdot\binom{N-a}{n-x}}{\binom{N}{n}} = \frac{a!}{x!\,(a-x)!}\ \frac{(N-a)!}{(n-x)!\,[(N-a)-(n-x)]!}\ \frac{n!\,(N-n)!}{N!}$$
$$= \frac{n!}{x!\,(n-x)}\ \frac{a!\,(N-a)!\,(N-n)!}{(a-x)!\,(N-a-n+x)!\,N!}$$
$$= \binom{n}{x}\frac{[a(a-1)\ldots(a-x+1)][(N-a)(N-a-1)\ldots(N-a-1)\ldots(N-a-n+x+1)]}{N(N-1)\ldots(N-n+1)}$$

for alle x = 0, 1, 2, …, n. a→∞ fordi a=pN og N→∞. Med N-a ≤N-n og n kan bli neglisjert siden a og N går mot uendelig, sette $q = \frac{1}{p} = \frac{N}{a}$ og får

$$f(x) = \binom{n}{x}\frac{[a(a-1)\ldots(a-x+1)]\big[(q-1)a\big((q-1)a-1\big)\ldots((q-1)a-n+x+1)\big]}{qa(qa-1)\ldots(qa-n+1)}$$
$$= \binom{n}{x}\left(\frac{1}{q}\right)^x\left(\frac{q-1}{q}\right)^{n-x} = \binom{n}{x}p^x(1-p)^{n-x}$$

Som er sannsynlighetsfunksjonen til den binomiske fordelingen.

Referanser

1 http://www.udir.no/kl06/MAT3-01/Hele/Kompetansemaal/Matematikk-R1/

2 Sinus matematikk R1, Oldervoll T., Orskaug, O., Vaaje, A., Svorstøl, O., Hals, S., Cappelen Damm, 2013

3 Statistical Concepts and Methods, Bhattacharyya, G., K., Johnson, R., A., John Wilet & Sons, 1977

4 Nasjonal Digital Læringsarena, www.ndla.no, https://ndla.no/nb/node/108221?fag=57933

5 Personlig overlevering fra Helmer Aslaksen, fra boka «Chance Rules» av Everitt, kapittel «Conditional Probability and the Reverend Thomas Bayes», avsnitt, «An Unsung Statistican».

6 Sommervoll, D., E., Mattespettboka statistikk, Gyldendal akademisk, 2013.

7 *George, Glyn, "Testing for the independence of three events," Mathematical Gazette 88, November 2004, 568.*

8 http://ndla.no/nb/node/103756

9 Bjørnstad, Jan. Hypergeometrisk Fordeling. I Store norske leksikon. https://snl.no/hypergeometrisk_fordeling.8

10 http://www.math.wm.edu/~leemis/chart/UDR/PDFs/HypergeometricBinomial.pdf