

Personlig stemmeafgivning

Efter valget i 2005¹ har man udspurgt en mindre del af de deltagende, om de har stemt personligt.

Man har svar fra 1131 mænd (hvoraf 54 % har stemt personligt - ifølge eget udsagn) og 928 kvinder (hvoraf 52 % har stemt personligt - ifølge eget udsagn).

Hvis alle husker rigtigt og alle er sanddru og alle personer er udvalgt helt tilfældigt, kan vi betragte svarene som en *stikprøve* – som kan give et billede af, hvorledes hele vælgerbefolkningen har stemt.

Dette billede kan være mere eller mindre rigtigt. I teorien kan man sagtens tænke sig, at der i stikprøven har været 54 %, selvom næsten ingen mænd i hele *populationen* har stemt personligt: Tilfældigvis (og uheldigvis) er der bare udvalgt en delgruppe, som ikke er *repræsentativ* for hele populationen.

Ofte gives der nogenlunde gode beskrivelser af populationen ved hjælp af stikprøver; dog er nøjagtigheden af målingerne tit overvurderet. Men en gang i mellem vil stikprøver være misvisende – på grund af tilfældigheder.

Valg 2005: Faktiske tal (fra stikprøven)

Køn	Personlig stemme	Ikke personlig stemme	I alt	Frekvens
Mand	607	524	1131	0,54
Kvinde	482	446	928	0,52
I alt	1089	970	2059	0,53

Der er en mindre men dog klar forskel på de to køn i stikprøven: vi vil *teste hypotesen* (påstanden):

” H_0 : Der er ingen forskel på mænds og kvinders tilbøjelighed til at stemme personligt”

I stikprøven er der en forskel på de to *frekvenser*, men det vi vil udtale os om, er ikke stikprøvens, men populationens frekvenser. Og som tidligere nævnt: Principielt kan vi aldrig vide, hvordan frekvenserne er i populationen. Men vi kan lave en model, hvor vi går ud fra H_0 . Vi kan i denne beregne (eller eksperimentere) os frem til, hvad der vil ske, hvis man lavede sådan en stikprøve uendeligt mange gange (eller i hvert fald et meget stort antal gange.)

Hvis H_0 gælder, kan vi beregne en fælles frekvens for mænd og kvinder. Det bliver i dette tilfælde 0,53 (eller 53 %.) For vor stikprøve kan vi nu beregne en *teststørrelse*, der fortæller

¹ Se <http://bank1.surveybank.aau.dk/webview/>

hvor langt udfaldet i stikprøven er fra det forventede udfald (idet vi forudsætter at nulhypotesen er sand.) Vi deler nu alle udfald i to dele: en med små afvigelser og en med store afvigelser. Vi sætter grænsen mellem små og store således, at kun 5 % af udfaldene vil være store (hvis man tænkte sig, at stikprøven blev udtaget uendeligt mange gange.)

Hvis vi i vor stikprøve har en lille afvigelse ræsonneres: Nulhypotesen H_0 *accepteres*. At stikprøven er lidt anderledes end det forventede skyldes (nok) bare tilfældigheder. At acceptere betyder ikke, at man tror, at hypotesen er 100 % rigtig, men at man indtil videre ikke ved bedre.

Derimod: Hvis vi i stikprøven har en stor afvigelse ræsonneres: Nulhypotesen H_0 *forkastes*. At afvigelsen er så stor er usandsynligt: det sker kun 5 ud af 100 gange. Vi tror ikke på det usandsynlige. Vi tror ikke på at meget sjældne begivenheder rammer netop os. Men da det trods alt sker 5 ud af 100 gange, vil vi sommetider bedømme hypoteser forkert. Det kunne vi hindre ved at forlange en større afvigelse, så vi fx kun tog fejl 1 ud af 100 gange. Ulempen er, at så bliver vi nødt til at acceptere mange forkerte hypoteser, som ellers kunne afvises.

Valg 2005: Forventede tal (hvis nulhypotesen er sand)

Køn	Personlig stemme	Ikke personlig stemme	I alt	Frekvens
Mand	598,18	532,82	1131	0,53
Kvinde	490,82	437,18	928	0,53
I alt	1089	970	2059	0,53

Med tallene fra forrige tabel beregnes en frekvens for dem, der stemmer personligt uden hensyn til køn:

$$f = (607+482)/(1131+928) = 1089/2059 = 0,53$$

Med denne frekvens beregnes fx det forventede antal mænd, der stemmer personligt som $m_p = 1131 \cdot 0,53 = 598,18$

Tilsvarende kunne det forventede antal kvinder beregnes; de sidste 3 tal findes dog nemmest, når man husker, at både række- og søjlesummer stadig skal være de samme som før.

For hvert af de 4 realiserede tal, vil vi beregne afstanden mellem faktisk (*fak*) og forventet værdi (*fv*) ifølge denne formel:

$$\text{Afvigelse} = \frac{(fak - fv)^2}{fv}$$

Adderes afvigelserne fås teststørrelsen X

$$X = \sum \frac{(fak - fv)^2}{fv}$$

Denne størrelse er et tal knyttet til vores stikprøve. Men da udtrækningen af stikprøven beror på tilfældigheder fås der ikke altid samme teststørrelse for hver stikprøve – selvom populationen er den samme. Teststørrelsen er en *stokastisk variabel*. Den realiserede værdi

af teststørrelsen skal sammenlignes med alle de mulige resultater fra uendeligt mange (teoretiske) stikprøver baseret på samme hypotese. Disse følger en kendt χ^2 -fordeling. Med denne afgøres om den samlede afvigelse er stor eller lille.

Valg 2005: Beregning af afvigelser

Køn	Personlig stemme	Ikke personlig stemme	I alt
Mand	0,13	0,15	0,28
Kvinde	0,16	0,18	0,34
I alt	0,29	0,32	0,61

Den først afvigelse beregnes som: $Afvigelse = \frac{(607 - 598,18)^2}{598,18} = 0,13$

Tilsvarende for de tre andre.

Teststørrelsen $X = 0,61$

En forudsætning for testen er, at alle forventede størrelser er mindst 5.

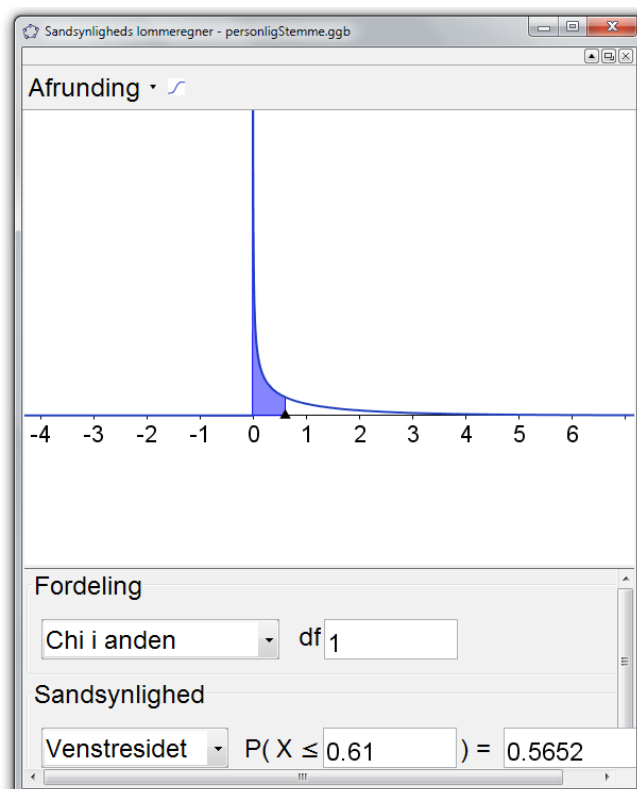
En anden ting er, at der er forskellige χ^2 -fordelinger: Vi beregner eller tæller, hvor mange tal i tabellen, der kan bestemmes, før resten er givet. Det er antallet af frihedsgrader (eng.: degrees of freedom eller df). I denne opgave er der 1 frihedsgrad. I denne opgave-type med r rækker og s søjler kan antallet af frihedsgrader beregnes som $(r-1)(s-1)$.

I GeoGebra kan vi med sandsynligheds-lommeregneren finde sandsynligheden for, at en afvigelse er mindre end den fundne afvigelse:

Her er sandsynligheden = 56,5 %. Det vil sige, at teststørrelsen befinder sig mellem de små afvigelser; sandsynligheden for at få en større teststørrelse er 100 % - 56,5 %; dvs. 43,5 %. Det sidste tal kaldes *p-værdien* og først når p-værdien er under 5 % forkastes hypotesen (når vi arbejder med *signifikansniveauet* 5 %.)

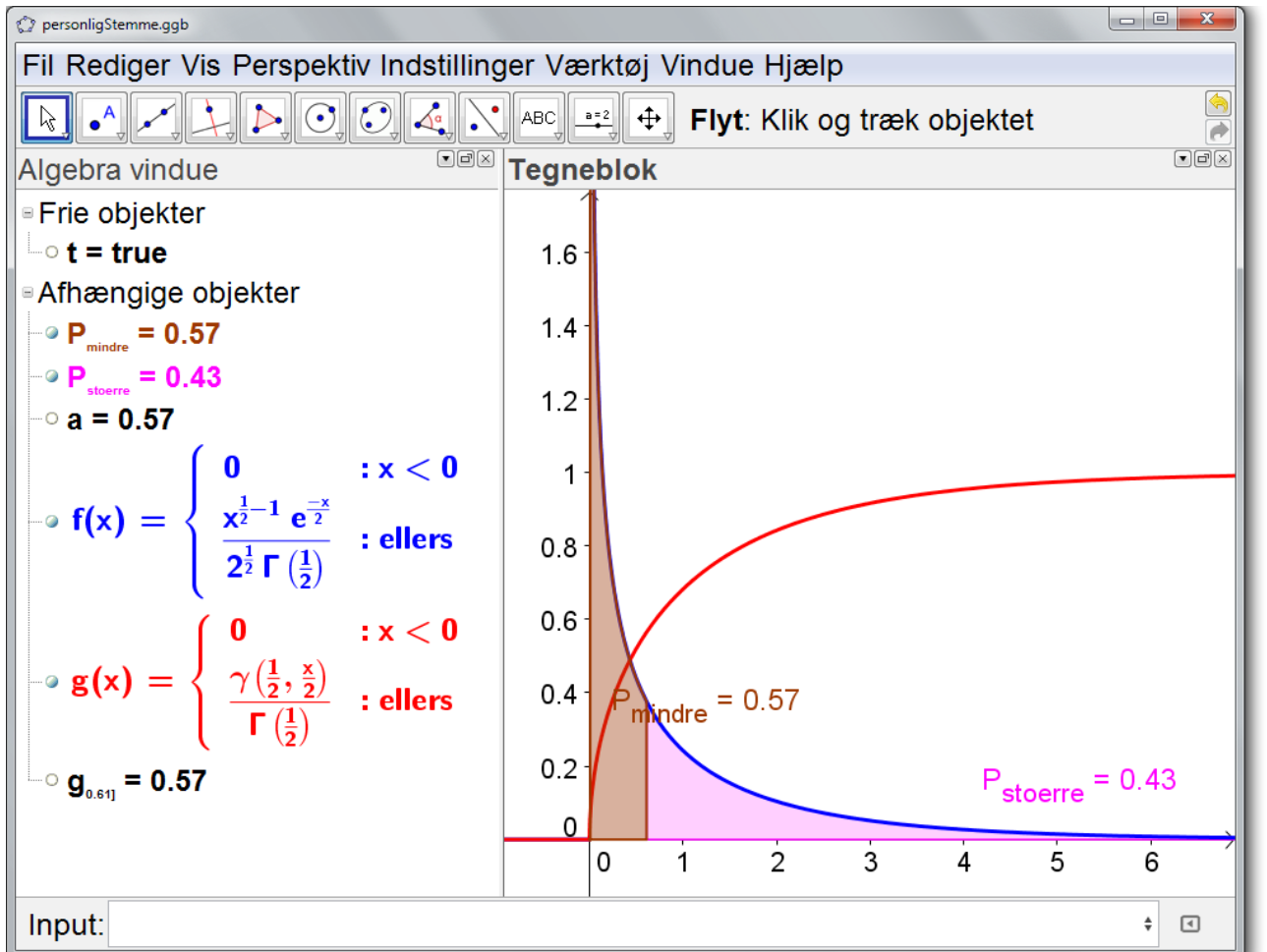
Den teststørrelse, der får os til at forkaste hypotesen, kaldes *den kritiske værdi*. Den afhænger derfor af signifikansniveauet. Er det 5 %, vil den kritiske værdi være 3,84, da

$P(X \leq 3,84) = 0,95$ og $P(X > 3,84) = 0,05$
(Jævnfør også tabellen i Bilag 2.)



Konklusionen er, at nulhypotesen accepteres.

Bilag 1: Grafisk illustration af forskellige sandsynligheder i GeoGebra



Bilag 2: p-værdier

Degrees of freedom (df)	χ^2 value ^[12]										
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88
10	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59
P value (Probability)	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001
	Insignificant								Significant		

Bilag 3: GeoGebra

Personlig stemmeafgivning (Valg 2005)

	A	B	C	D	E
1	Valg 2005	Faktiske tal			
2					
3	Køn	Personlig st...	Ikke personli...	I alt	Frekvens
4	Mand	607	524	1131	0.54
5	Kvinde	482	446	928	0.52
6		1089	970	2059	
7	Fælles ps frekvens	0.53			
8					
9	Tabel forudsat nulhypotesen: ingen forskel på M ...				
10					
11	Valg 2005	Beregneede tal			
12					
13	Køn	Personlig st...	Ikke personli...	I alt	
14	Mand	598.18	532.82	1131	
15	Kvinde	490.82	437.18	928	
16		1089	970	2059	
17					
18	Valg 2005	Forskelle ...			
19					
20	Køn	Personlig st...	Ikke personli...	I alt	
21	Mand	0.13	0.15	0.28	
22	Kvinde	0.16	0.18	0.34	
23		0.29	0.32	0.61	
24					
25					

Bilag 4: Ordliste

stikprøve en del af en population

population den mængde, vi vil udtale os om (endelig eller uendelig, nutidig eller fremtidig)

repræsentativ ligner nogenlunde

teste undersøge om en påstand er rigtig

hypotesen en formodning eller påstand, der skal undersøges

nulhypotesen en hypotese, hvor en parameter antages at have en kendt værdi eller som her: at frekvensen for kvinder og mænd er ens. Nulhypotesen betegnes: H_0 .

H_0 *accepteres* hvis afvigelsen fra det forventede er lille

H_0 *forkastes* hvis afvigelsen fra det forventede er usandsynlig stor, fx så stor at det kun vil ske i 5 % af alle tilfælde (eller evt. 1 % eller 0,1 %)

observationer både stikprøver og populationer kan beskrives ved hjælp af observationerne, det være sig alder, køn eller om man stemmer personligt

observationssæt er mængden af observationer

hyppighed er antallet af en bestemt type observation

observationssættets størrelse = n er antallet af observationer i observationssættet eller summen af hyppigheder

frekvens er $\text{hyppighed}/n$; frekvensen kan skrives som , decimalbrøk eller procent

kumulerede frekvenser er den brøkdel af observationerne, der er mindre end eller lig med en given størrelse

stokastiske eksperimenter er eksperimenter med et tilfældigt *udfald* som fx terningekast (med *udfaldsrummet* $U=\{\text{plat, krone}\}$.)

sandsynlighed er et mål for den forventede frekvens af et udfald eller en delmængde af udfald ved et bestemt eksperiment

stokastisk variabel er en funktion af et stokastisk eksperiments udfald; det kunne være antal "plat" ved 10 terningekast eller som her: den beregnede afvigelse mellem de forventede og de faktiske værdier i tabellerne; funktionsværdien betegnes X

$P(X \leq t) = p$ læses: sandsynligheden ² for at den stokastiske variabel er mindre end eller lig med t er p (som er et tal i intervallet $[0, 1]$.)

χ^2 - *fordeling* en samling funktioner, der for frihedsgraderne 1,2,3, ... kan angive $P(X \leq t)$, hvor X er den stokastiske variabel, der måler afvigelse i vort eksempel

p-værdien $P(X > t)$ eller $1,00 - P(X \leq t)$

2 (eng.) probability

signifikansniveau er den procentdel af udfaldene, der anses som "usandsynlige"; her fx valgt som de 5 % mest ekstreme

den kritiske værdi er så grænsen mellem de sandsynlige og de usandsynlige udfald; den afhænger naturligvis af hvilket signifikansniveau der er valgt