

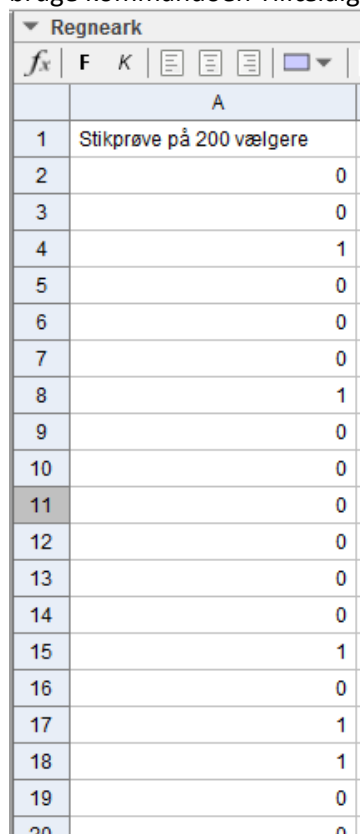
Ekspirement med konfidensintervaller

Konfidensintervaller – 95% af mange konfidensintervaller indeholder den sande andel

Ekspirementet i dette papir ligger som video på <https://youtu.be/UeoGL5oAMm4>.

Del 1: Vi har en stikprøve på fx 200 vælgere, og vi kender den sande andel af vælgere, der stemmer på et bestemt parti fx 23%.

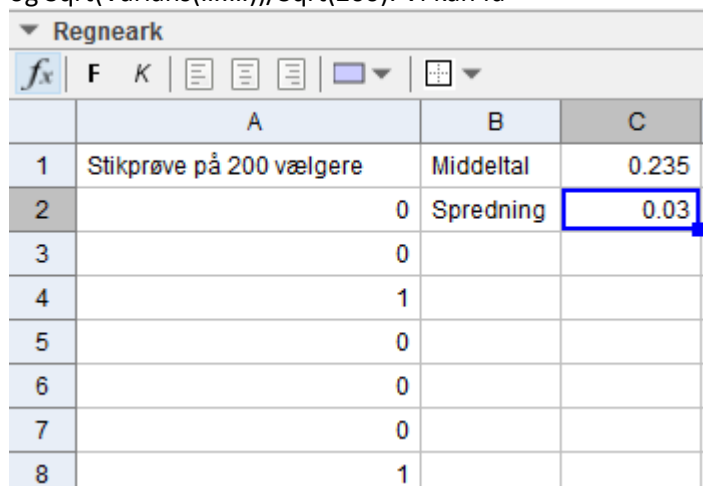
Vi vil starte med at simulere en sådan stikprøve, hvor vi i Geogebra opretter en liste. Hvert enkelt element er en simulering af en vælger, hvor sandsynligheden for at stemme på det pågældende parti er 0,23. Vi kan bruge kommandoen TilfældigBinomial(1,0.23) i hver celle. Vi kan få en stikprøve, hvor en del ser ud som



Regneark	
A	
1	Stikprøve på 200 vælgere
2	0
3	0
4	1
5	0
6	0
7	0
8	1
9	0
10	0
11	0
12	0
13	0
14	0
15	1
16	0
17	1
18	1
19	0
20	0

Hvis vi taster Crtl R, så får vi hver gang en ny simulering af en stikprøve.

Del 2: For hver stikprøve kan vi udregne middeltallet og spredningen med kommandoerne Middeltal(.....) og Sqrt(Varians(.....))/Sqrt(200). Vi kan få



Regneark			
	A	B	C
1	Stikprøve på 200 vælgere	Middeltal	0.235
2	0	Spredning	0.03
3	0		
4	1		
5	0		
6	0		
7	0		
8	1		

For hver genberegning med Crtl R for vi en ny stikprøve og dermed et nyt middeltal og en ny spredning.

Del 3: For hver stikprøve kan vi dermed bestemme venstre og højre grænse for konfidensintervallet ud fra de to værdier.

Vi kan få

▼ Regneark					
fx F K [] [] [] [] [] []					
	A	B	C	D	E
1	Stikprøve på 200 vælgere	Middeltal	0.3	Konfidensinterval	
2	0	Spredning	0.032	Venstre grænse	Højre grænse
3	1			0.235	0.365
4	0				
5	0				
6	1				
7	1				

Hvis vi anvender Ctrl R mange gange, så ser vi, at vi ofte for et interval, som den "sande" andel tilhører.

Del 4: Vi vil nu lave fx 300 simuleringer og for hver simulering bestemme grænserne for konfidensintervallet. Vi kan højreklikke på hver af de to celler D3 og E3 og vælge "Optag i regneark".

Vi kan få

▼ Regneark								
fx F K [] [] [] [] [] []								
		B	C	D	E	● II F	● II G	
1	re	Middeltal	0.3	Konfidensinterval		D3	E3	
2	0	Spredning	0.032	Venstre grænse	Højre grænse			
3	1			0.235	0.365			
4	0							
5	0							
6	1							
7	1							

Tryk på Ctrl R 300 gange. Vi kan få

292						0.148	0.262	
293						0.139	0.251	
294						0.139	0.251	
295						0.175	0.295	
296						0.175	0.295	
297						0.217	0.343	
298						0.166	0.284	
299						0.157	0.273	
300						0.148	0.262	

Del 5: Vi kan nu undersøge for hvert konfidensinterval om den sande værdi for andelen ligger i intervallet med kommandoen Hvis(F2<=0.23<=G2,1,0).

Vi kan få

291						0.17	0.29	1
292						0.148	0.262	1
293						0.139	0.251	1
294						0.139	0.251	1
295						0.175	0.295	1
296						0.175	0.295	1
297						0.217	0.343	1
298						0.166	0.284	1
299						0.157	0.273	1
300						0.148	0.262	1
301								286

Konklusion: Vi får i disse 300 simuleringer, at i $\frac{286}{300} = 0,953$. Dvs. vi eksperimentelt har vist, at vi arbejder med et 95%-konfidensinterval.