# Normalfordelingsapproksimationen

En stokastisk variabel er binomialfordelt med sandsynlighedsparameter og antalsparameter hvis

Vi har desuden en kort notation: .

Man kan vise at middelværdien er og spredningen er .

Hvis og , så er normalfordelingen med samme middelværdi og spredning en god approksimation til binomialfordelingen.

### Opgave 1

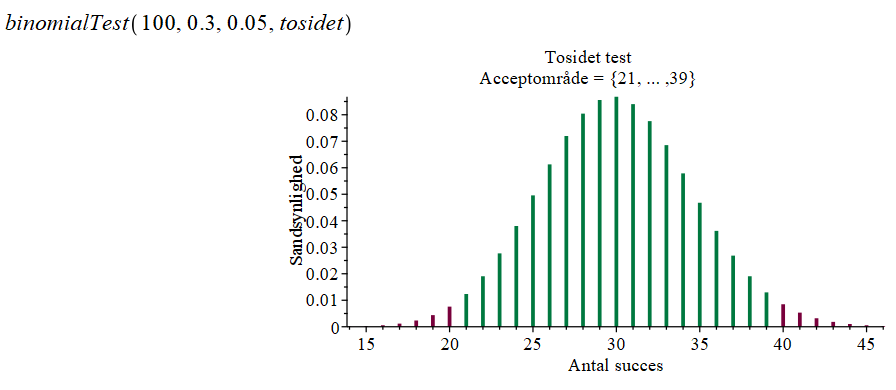
Undersøg hvor god en approksimation normalfordelingen er til binomialfordelingen.   
F.eks. ved at beregne hvis er binomialfordelt med sandsynlighedsparameter og antalsparameter og sammenligne med hvad normalfordelingen giver.

*Tip: det var noget med et sumtegn når man skulle bestemme sandsynligheden for flere udfald med binomialfordelingen.*

### Opgave 2

Nedenfor ses en binomialtest med sandsynlighedsparameter , antalsparameter og et signifikansniveau på .

1. Hvordan er acceptområdet i testen bestemt?
2. Approksimér acceptområdet vha. de normale udfald i normalfordelingen.
3. Find på et eksempel hvor testen kunne bruges.



Lad være binomialfordeling med sandsynlighedsparameter og antalsparameter . Da binomialfordelingen kan approksimeres med normalfordelingen, har vi fra normalfordelingen at

Vi er interesseret i at lave et konfidensinterval for sandsynlighedsparameteren . Hvis vi indsætter middelværdi og spredning, får vi:

Hvis vi har en stikprøve af vores binomialfordelte indfører vi og kan estimere et konfidensinterval for således:

Vi har tidligere set at er den sandsynlighedsparameter hvor sandsynligheden for at få stikprøven er størst, så det giver god mening at den er midten af vores estimat. Intervallet kan omskrives til det nedenstående (som vi har brugt tidligere):

### Eksempel

Megafon laver en opinionsundersøgelse i forhold til næste folketingsvalg hvor de spørger 1000 personer og 480 svarer at de ville stemme på rød blok. Hvis vi antager at de 1000 personer er udvalgt tilfældigt blandt vælgerne, dvs. de udgør en repræsentativ stikprøve, så har vi antallet som svarer at de stemmer på rød blok er binomialfordelt med antalsparameter og en ukendt sandsynlighedsparameter som angiver andelen af vælgerne som svarer at de stemmer på rød blok. Vi har desuden vores stikprøve   
. Dermed har vi estimatet

Som fortæller at vi ud fra stikprøven estimerer at af vælgerne svarer at de stemmer på rød blok. Hvis vi gerne vil vide hvor sikre vi er på dette estimat, så kan vi lave et konfidensinterval for :

Dermed har vi estimeret et konfidensinterval for til mellem og .

### Opgave 3

1. Afgør vha. konfidensintervallet i eksemplet ovenfor om nulhypotesen nedenfor kan forkastes på et signifikansniveau.

*af vælgerne svarer at de vil stemme på rød blok.*

1. Hvordan ville man vha. et konfidensinterval afgøre om en nulhypotese kunne forkastes på et signifikansniveau?