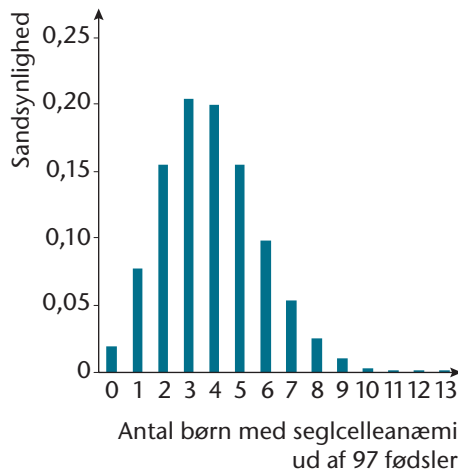


	Raske	Syge
Voksne	341	0
Børn 0-5 år	92	5
Børn 12-17 år	105	1

Figur A. Fordelingen af seglcelleanæmi i forskellige aldersgrupper.



Figur B. Sandsynlighed for at et bestemt antal børn i alderen 0-5 år er født med seglcelleanæmi ved 97 fødsler.

SANDSYNLIGHEDER

Grafisk aflæsning og beregning bør kombineres. *Tosidet test:* Når man i Excel har kolonnen med binomial-sandsynligheder, kan man markere felter fra oven i kolonnen indtil man når 0,025 i 'Sum:' vist i regnearkets nederste kant. Når man er over 0,025, har man det lave tal i acceptområdet. Det samme gøres nedefra, hvorved det høje tal i acceptområdet findes. Området uden for acceptområdet kaldes for det kritiske område. *Ensidet venstretillet test:* Markér celler fra oven indtil man når over 0,05. *Ensidet højrestillet test:* Markér celler fra neden indtil man når over 0,05.

Binomialfordelingstest

Binomialfordelingstesten kan foretages når forudsætningerne for binomialfordelingen er til stede – se boks side 63. Fx kan dominant/recessiv autosomal nedarvning eller en gruppe og dens komplementærgruppe, fx 'farveblinde kvinder' og 'kvinder med normalsyn', testes ved en binomialfordelingstest.

Sygdommen seglcelleanæmi nedarves recessivt og forekommer hos individer med genotypen *ss*. Allelfrekvensen for *s* i en undersøgt befolkningsgruppe er 0,2, og derved er allelfrekvensen for *S* = 0,8. Under antagelse af Hardy-Weinberg ligevægt er sandsynligheden for at et individ har sygdommen derfor $0,2^2 = 0,04$. I figur A ses at 5 ud af 97 børn i alderen 0-5 år er syge. I Excel kan man udregne sandsynligheden for et bestemt antal børn ud af 97 fødes med seglcelleanæmi. Dette gøres ved at man i A-kolonnen laver en talrække fra 0 til 97 så 0 står i A1, og 97 står i A98, svarende til at henholdsvis 0 og 97 børn er syge. I B1 skrives formelen '=BINOMIAL.FORDELING(A1;97;0,04;FALSK)'. Formlen kopieres til cellerne B1 til B98. Resultaterne kan afbilledes som et søjlediagram i Excel, i stil med figur B. Dette uddybes i (3.4).

Sandsynlighedsbetragtninger

Ønsker man med et signifikansniveau på $p = 0,05$ at undersøge om der er Hardy-Weinberg-ligevægt hos gruppen af børn mellem 0 og 5 år, skal man lave en tosidet test og finde de midterste 95% kaldet 'acceptområdet'. Man antager at det kritiske område fordeler sig med 0,025 i hver side af sandsynlighedsfunktionen. Først når antallet af børn med sygdommen ligger udenfor acceptområdets 95%, regner man med at fordelingen ikke er i Hardy-Weinberg-ligevægt. I figur B kan det aflæses at 0,025 i venstre side af figuren først opnås ved 1 barn, og at 0,025 i højre side ind mod grafens midte først opnås ved 8 børn. Der er derfor Hardy-Weinberg-ligevægt hvis der ud af 97 børn er mellem 1 og 8 børn med sygdommen. 5 syge børn betyder derfor at man ikke kan forkaste hypotesen om at fordelingen er i Hardy-Weinberg-ligevægt.

Ensidet binomialfordelingstest

I figur A fremgår det at der er 341 voksne, der alle er raske. Udtrykket '=BINOMIAL.FORDELING(0;341;0,04;FALSK)' giver værdien $9 \cdot 10^{-7}$. Da denne værdi er under 0,025 er gruppen af voksne ikke i Hardy-Weinberg-ligevægt. Individer med genotypen *ss* har en markant lavere forventet levealder end resten af befolkningen. Derfor kan man stille spørgsmålet: Er der færre voksne med seglcelleanæmi end forventet? Svaret på dette spørgsmål kræver en ensidet test. I dette tilfælde skal hele $p = 0,05$ placeres i venstre side af sandsynlighedsfunktionen. Derved kan der være op til 7 voksne med seglcelleanæmi i gruppen og stadig være færre end forventet. Se mere i (3.4).