

Ofte udføres eksperimenter, hvor man forventer et bestemt resultat. Byg, der har været udsat for radioaktiv

beståelse, vil efter spiring fremkomme i to forskellige farver. Man kan undersøge hyppigheden mellem de to farver og sammenligne med det forventede resultat ifølge en biologisk model. Man kan også sammenligne blodtypers fordeling i en gymnastiekklasse, med den kendte blodtypefordeling i Danmark. I begge tilfælde skal man

vurdere, om afvigelsen mellem det observerede og det forventede resultat, er acceptabel eller ej, dvs. om afvigelsen skyldes en tilfældighed, eller om afvigelsen skyldes en virkelig forskel. For at forstå

hvordan afgørelsen træffes, er det vigtig først at udføre metoden "håndholdt" for til sidst at få Excel til at udføre hele arbejdet. Metoden er enkel at udføre, men

svær at fortolke.

Det græske bogstav χ udtales χ , og χ^2 udtales "ki i anden".

Eksempel 1. Test om en mønt er fair.

Hypotese. Mønten er fair, dvs. sandsynlighed for plat er $\frac{1}{2}$ og sandsynligheden for krone er $\frac{1}{2}$.

Observation. Ved 20 kast med en mønt fås 6 plat og 14 kroner.

Udfald	plat	krone
Observeret	6	14
Forventet	10	10

Hypotesen er vores model af virkeligheden og observationen er virkeligheden. For at afgøre om model og virkelighed afviger fra hinanden, skal vi opstille og vurdere et mål for denne afvigelse.

Teststørrelse. Teststørrelsen er et tal, der angiver afvigelsen mellem det forventede (Expected) og det observerede (Observed).

$$\text{Teststørrelse} = \frac{(\text{observeret antal plat} - \text{forventet antal plat})^2}{\text{forventet antal plat}} + \frac{(\text{observeret antal krone} - \text{forventet antal krone})^2}{\text{forventet antal krone}}$$

$$\text{Teststørrelse} = \frac{(O_{\text{plat}} - E_{\text{plat}})^2}{(O_{\text{krone}} - E_{\text{krone}})^2} + \frac{E_{\text{krone}}}{E_{\text{plat}}}$$

$$\text{Teststørrelse} = \frac{(6 - 10)^2}{(14 - 10)^2} + \frac{10}{10} = \frac{16}{16} + 1 = 2$$

Hvis tælleren ikke sættes i anden, ville de to led ophæve hinanden og teststørrelsen ville give 0. Jo større teststørrelsen er, jo mere afviger det observerede antal fra det forventede. Den svære afgørelse er, hvor stor teststørrelsen skal være (kritisk værdi), før afvigelsen er så betydningsfuld (signifikant), så vi vælger at forkaste (afvise) hypotesen om en fair mønt.

Bestemmelse af antallet af frihedsgrader. Hvis vi kaster 20 gange med en mønt, kan vi nøjes med at opskrive antallet af kroner. Hvis der er 12 kroner, ved vi der må være 8 plat. Når man tester en mønt, er antallet af frihedsgrader 1. Når man tester en terning, er antallet af frihedsgrader 5. Der gælder generelt, at antallet af frihedsgrader er 1 mindre end antallet af udfald.

Signifikansniveau. Er det sandsynligt, at en teststørrelse på 3,2 kan være fremkommet ved kast med en fair mønt? Hvis denne sandsynlighed er over 5 % accepteres hypotesen om en fair mønt og hvis sandsynligheden er under 5 % forkastes hypotesen. De 5 % kaldes signifikansniveauet.