# Binomialtest - $p$-værdi

Ved binomialtest betragter vi en observeret værdi, og vi har vurderet den i forhold til kritisk område og acceptområde.

Man kan også afgøre, om en nulhypotese skal forkastes ved at betragte $p$-værdien, i stedet for at bruge kritisk område og acceptområde. Lad os se på, hvad der menes med begrebet $p$-værdi.

<u>Definition af $p$-værdien</u>: Sandsynligheden for at få et udfald, under antagelse af nulhypotesen, der er mindst lige så langt væk fra middelværdien (det forventede udfald), som det observerede udfald.

- Det gælder at den stokastiske variabel $X$ er binomialfordelt, $X \sim b(n, p)$.
- En kendt sandsynlighed (ofte fra en population) er kendt. Den kan kaldes for $p_0$. Da har vi $X \sim b(n, p_0)$, i det tilfælde at nulhypotesen er sand.
- $p$-værdien beregnes med den observerede værdi og vores viden om, at der er tale om et tosidet eller etsidet test. Værdien beregnes som en kumuleret sandsynlighed.
- $p$-værdien sammenlignes med signifikansniveauet når man skal afgøre, om nulhypotesen skal forkastes eller ikke forkastes.
- I dette materiale *antager* vi, at *middelværdien er et helt tal*.
- Man beregner kumulerede sandsynligheder i begge sider, *for et en tilfældig observation er lige langt væk fra middelværdien i begge sider.*
    - Lad $x_{obs}$ være observationen.
    - Lad $\mu$ være middelværdien.
    - Lad $d$ være den absolutte afvigelse fra middelværdien: $d = |\mu - x_{obs}|$
    - $p$-værdi $= P(X \leq \mu - d) + P(X \geq \mu + d)$
    - I Nspire benyttes `binomCdf` til beregning af de kumulerede sandsynligheder.

## Eksempel 1 på beregning og brug af $p$-værdi

En producent hævder, at 40% af deres produkter bliver solgt inden for den første uge.

Nulhypotese: $H_0: p = 0{,}40$

Alternativ hypotese: $H_1: p \neq 0{,}40$

Tosidet binomialtest. Signifikansniveau på 5%.

**Stikprøve** Der udtages 50 tilfældigt valgte produkter og 15 blev solgt inden for den første uge.

Middelværdien er $\mu = 50 \cdot 0{,}40 = 20$

Beregning og brug af $p$-værdien

I dette tilfælde er forskellen mellem observationen og middelværdien $|20 - 15| = 5$.
Vi beregner nu summen af højre og venstre side således:

$$p\text{-værdi} = P(X \leq 20 - 5) + P(X \geq 20 + 5) = P(X \leq 15) + P(X \geq 25) = 0{,}193$$

Nulhypotesen kan ikke forkastes på et 5% signifikansniveau, da $p$-værdien er større end 0,05.

Det svarer til at den observerede værdi ligger i acceptområdet.

## Eksempel 2 på beregning og brug af $p$-værdi

En producent hævder, at 40% af deres produkter bliver solgt inden for den første uge.

Nulhypotese: $H_0: p = 0{,}40$

Alternativ hypotese: $H_1: p \neq 0{,}40$

Tosidet binomialtest. Signifikansniveau på 5%.

**Stikprøve** Der udtages 50 tilfældigt valgte produkter og 12 blev solgt inden for den første uge.

Beregning og brug af $p$-værdien som i det tidligere eksempel, men nu med:

$$p\text{-værdi} = P(X \leq 20 - 8) + P(X \geq 20 + 8) = P(X \leq 12) + P(X \geq 28) = 0{,}0293$$

Nulhypotesen forkastes på et 5% signifikansniveau, da $p$-værdien er mindre end 0,05.

Det svarer til at den observerede værdi ligger i det kritiske område.

## Eksempel 3 – fra Sandsynlighedsregning (Mathematicus)

Dette er en direkte kopi af et eksempel fra bogen, men skrevet lidt anderledes.

Eksempel 3.8 og 3.9

I Vaffelbjerg Kommune fik Protestpartiet ved sidste kommunalvalg 17,2% af stemmerne. Ved en meningsmåling hvor man har spurgt 1000 repræsentativt udvalgte personer, siger 243 personer at de vil stemme på partiet ved næste kommunalvalg.

Hvis man vil vide om partiets stemmeandel har ændret sig, kan man udføre et tosidet binomialtest. Nulhypotesen bliver i dette tilfælde

$\mathrm{H}_0$: Partiets stemmeandel er 17,2%.

Binomialfordelingen har så $n = 1000$ og $p = 17{,}2\%$.

- 243 af de 1000 personer i stikprøve vil stemme på protestpartiet.
- $\mu = n \cdot p = 1000 \cdot 0{,}172 = 172$
- $|\mu - x_{obs}| = |172 - 243| = 71$
- Det skal være et tosidet test, da der er mulighed for at flere eller færre personer vil stemme på Protestpartiet.
- $p$-værdi $= P(X \leq 172 - 71) + P(X \geq 172 + 71) = P(X \leq 101) + P(X \geq 243)$
- Vi beregner $p$-værdien i Nspire:
  $p$-værdi = binomcdf(1000, 0.172, 101) + binomcdf(1000, 0.172, 243, 1000)
  $p$-værdi $= 8{,}37 \cdot 10^{-9}$.
- Det tal er mindre end 0,05. Nulhypotesen forkastes på et 5% signifikansniveau.
- Der er evidens for at tilslutningen til Protestpartiet har ændret sig.