

3. Hvor overraskende (sannsynlig) er det at  $h$  er korrekt, uavhengig av  $e$ ?

**Base rate:** Sannsynligheten for at noe er tilfelle i utgangspunktet, *før* ny evidens kommer inn, for eksempel sannsynligheten for at en hypotese er korrekt *før* den utsettes for en test.

Hvilken tillit vi skal ha til en hypotese, handler ikke bare om sannsynligheten for en observasjon, men også *hvor sannsynlig hypotesen allerede er*. Vi kan illustrere poenget med en kjent problemstilling: *base rate-problematikken*.

Anta at det finnes en svært sjelden sykdom X som er 100 % dødelig. Anta at det finnes en test for sykdom X som er 99 % pålitelig, altså at den gir riktig resultat i 99 % av testtilfellene (merk at det ikke er akkurat dét som menes med «99 % pålitelighet» når man for eksempel snakker om koronatester). La oss si at alle er blitt anbefalt å ta testen, uansett symptomer, og du tilfeldigvis tar den og tester positivt!

Anta så at det finnes en kur som er 100 % effektiv, men som dessverre på egen hånd tar livet av 30 % av pasientene som gjennomgår den. Du har ikke tid til å ta en ny test. Takker du ja til kuren?

Undersøkelser viser at de fleste svarer «ja» på det spørsmålet (Tversky & Kahneman, 1974). Å svare «ja» er galskap.

For hva er sjansen for at du har sykdom X gitt testen? Testen er 99 % pålitelig. Derfor tenker mange at svaret er «99 %». Det er feil. Sannheten er at du ikke har nok informasjon til å avgjøre hva

sannsynligheten er. Men den er antakeligvis langt lavere enn sannsynligheten for at du dør av kuren om du sier «ja». Cluet er «svært sjelden». For å avgjøre hvor sannsynlig det er at du har sykdommen, er det ikke nok å vite hvor pålitelig testen er. Du må også ta hensyn til base rate-sannsynligheten: *Du må vite hvor sannsynlig det er at du har sykdommen i utgangspunktet.* La oss si at «svært sjelden» betyr at blant én million vilkårlig valgte mennesker vil – statistisk sett – én person være syk. Dersom disse én million menneskene tok en test som er 99 % pålitelig, hvor mange vil få et positivt resultat? 1 % av resultatene er gale, så av én million som tar testen og ikke har sykdommen, vil 10 000 få et usant positivt resultat. Sjansen for at du er en av de 10 000 som fikk et usant resultat, er mye høyere enn sjansen for at du er den ene som har sykdommen.

Eller sagt annerledes: Får du et positivt resultat, vet du at noe usannsynlig har skjedd. Enten har du sykdommen, eller så har testen gitt deg galt resultat. Du må avgjøre hvilket av disse scenarioene som er mest sannsynlig:

1. Du har sykdommen (én på en million).
2. Testen ga feil resultat (én på hundre).

Det er langt mer sannsynlig at alternativ 2 er tilfellet enn alternativ 1.

Med prinsipp 3 kan vi fange opp denne dynamikken: Hvor sannsynlig hypotesen at du har sykdommen ( $h$ ) er, gitt en positiv test ( $e$ ), avhenger av hvor sannsynlig (eller overraskende) hypotesen var i utgangspunktet, før du tok testen. Bruker vi de to første prinsippene, ser vi at en positiv test oppfyller begge i stor grad: En positiv test er sannsynlig (99 %) om du har sykdommen, og usannsynlig (1 %)