

- Excel: Bruk funksjonen KJI.TEST, hvor du angir hvilke regnearkceller som inneholder henholdsvis de observerte og de forventede verdiene. Deretter beregnes signifikanssannsynligheten ( $p$ -verdien) for testen. Hvis beregningsresultatet er lavere enn valgt  $\alpha$ , kan nullhypotesen forkastes.

### 8.5.2 Simpsons paradoks

Hvis vi forkaster nullhypotesen, kan vi konkludere med at det er en statistisk sammenheng mellom rad- og kolonnekjennetegnene. Men vi har ikke påvist noen årsakssammenheng. Vi må alltid være oppmerksom på at utelatte variabler kan være årsaken til den sammenhengen som vi observerer. Kanskje fjordingene er sterkere fordi de brukes til tungt skogsarbeid, mens araberhestene er slappe fordi de står det meste av dagen i sine båser? Årsakssammenhengen er i så fall knyttet til hvordan hestene brukes, og ikke til hestens rase.

Våre statistiske konklusjoner må alltid ledsages av sunn fornuft og god kunnskap om det fenomenet vi studerer. I noen tilfeller kan vi faktisk få helt feil bilde dersom vi utelater en viktig variabel, slik at den riktige sammenhengen er motsatt av den vi har oppdaget. Dette kalles Simpsons paradoks og illustreres med det følgende eksemplet. Slike feilkonklusjoner skyldes ofte at data fra ulike kilder (tabeller) er slått sammen i en tabell. Les eksemplet, og husk budskapet neste gang du uttaler deg på bakgrunn av statistiske undersøkelser.

**Eksempel 258** Det hadde vært interessant å sende ut et spørreskjema til bilførere for å kartlegge hvor mange som har fått en bilskade de siste årene. Resultatet av undersøkelsen kunne blitt slik:

	Bilskade	Ikke bilskade	Total
Mann	233	323	556
Kvinne	87	194	281
Total	320	517	837

Her er det en signifikant forskjell på kjønnetenes evne til å kjøre skadefritt. Andelen  $233/556 = 0.42$  av mennene var involvert i en skade, mens kvinnenes andel bare var  $87/281 = 0.31$ . Betyr dette at kvinner er dyktigere sjåførere enn menn?

Løsning: Ved å undersøke spørreskjemaene nærmere kunne vi foreta en ekstra kategorisering, avhengig av hvor stor bil personene kjører. Da kunne resultatene presenteres slik:

	Store biler			Små biler		
	Skade	Ikke skade	Total	Skade	Ikke skade	Total
Mann	150	35	185	82	288	370
Kvinne	16	2	18	72	192	264
Total	166	37	203	154	480	634

Nå er bildet snudd på hodet. For store biler var 88 % av kvinnene mot 81 % av mennene involvert i en skade. For små biler var 27 % av kvinnene mot 22 % av mennene involvert. Kvinnene har høyest skadeandel både for store og små biler!

### 8.5.3 Modelltest (goodness of fit)

Istedenfor å sammenligne radene i en krystabell med hverandre kan vi velge å sammenligne en tabellrad med en slags "hypotetisk fasit". Dette er en av de statistiske testene som benyttes mest. Testen er svært anvendelig for å analysere data fra spørreundersøkelser, som vi skal se i neste eksempel.