

Baggrundsnotat: Signifikans

Begrebet ”signifikansniveau” bliver ofte misforstået. Denne artikel hjælper dig med at forstå begrebet statistisk signifikans samt de tal The Survey System producerer.

Artiklen er delt i to dele. Den første del simplificerer statistisk signifikans så meget som muligt, så konceptet også kan bruges af nybegyndere til analysen af data. Den anden del indeholder en mere fyldestgørende diskussion af statistisk signifikans til erfarne brugere af statistiske begreber.

Del 1: Statistisk signifikans for begyndere

I daglig tale betyder ”signifikant” vigtigt, men i forhold til statistik betyder ”signifikant” sandsynligvis sandt (altså ikke påvirket af tilfældigheder). Et resultat i undersøgelsen kan være sandt uden at være vigtigt. Når statistikere siger, at et resultat i høj grad er signifikant, betyder det, at det sandsynligvis er sandt. De mener ikke nødvendigvis, at resultatet er meget vigtigt.

Prøv at se på nedenstående tabel. Chi² –tallene nederst i tabellen viser to rækker af tal. Tallene i den øverste række, 0.07 og 24.4, er de statistiske Chi² –tal. I forhold til formålet med artiklen kan betydningen af tallene ignoreres. Den næste række indeholder tallene .795 og .001. Tallene er signifikansniveauerne og vil blive gennemgået under tabellen.

Køber du brændstof hos X firma							
		Område		Køretøjets type			
	Total	Centrum	Forstad	Bil	Lastbil	Bus	Varebil
Base	713	361	352	247	150	44	180
Ja	428 60%	215 60%	213 61%	131 53%	74 49%	29 66%	131 73%
Nej	285 40%	146 40%	139 40%	116 47%	76 51%	15 34%	49 27%
Chi Square			0.07 .795				24.4 .001

Signifikansniveau

Signifikansniveauerne viser, hvor sandsynligt det er at et resultat skyldes tilfældigheder. Det mest brugte niveau for at et resultat er troværdigt er 0,95. Det betyder, at der er 95% sandsynlighed for at resultatet er sandt. Men værdien bruges også på en misvisende måde. Der er ikke noget statistik program, som vil vise 95% eller 0,95 som en indikator for niveauet. I stedet vil programmet vise dig 0,05, hvilket betyder at der er 5% sandsynlighed for at resultatet ikke er sandt. De 5% er det omvendte af, at der er 95% sandsynlighed for, at resultatet er sandt. For at finde signifikansniveauet til resultatet skal du trække det viste tal fra 1. For eksempel betyder tallet 0,01, at der er 99% sandsynlighed ($1 - 0,01 = 0,99$), for at resultatet er sandt. I ovenstående tabel har det sandsynligvis ingen betydning for købet af benzin X, om man bor i centrum af byen eller i forstæderne, fordi sandsynligheden er 0,795 (dvs. at der kun er 20,5% sandsynlighed for at forskellen er sand). Derimod har biltype et højt signifikansniveau (0,001 eller 99,9%), hvilket indikerer, at biltype næsten helt sikkert har betydning for, om bilejerne i populationen, som stikprøven er taget fra, køber benzin X.

Signifikansniveauer i The Survey System

The Survey System bruger signifikansniveauer ved adskillige statistiske tests. I alle tilfælde er det p-værdien, der viser sandsynligheden for, at resultatet ikke er sandt. Hvis en Chi² –test viser en sandsynlighed på 0,04, betyder det, at der er 96% ($1 - 0,04 = 96\%$) sandsynlighed for, at svarene givet af forskellige grupper af svarpersoner i et banner virkelig er forskellige. Hvis en t-test viser en sandsynlighed på 0,07, betyder det, at der er 93% sandsynlighed for at de to gennemsnit, som sammenlignes, også ville være forskellige, hvis hele populationen blev undersøgt.

Del 2: Signifikans for de erfarne

Mange personer mener, at 95% niveauet er helligt, når det drejer sig om signifikansniveauer. Hvis en test viser 0,06 sandsynlighed, betyder det, at der er 94% sandsynlighed for at resultatet gælder for hele populationen. Du kan selvfølgelig ikke være ligeså sikker, som hvis sandsynligheden var 95%, men oddsene er stadig, at resultatet er gældende. 95%-niveauet stammer fra akademiske publikationer, hvor teorien i reglen er, at der skal være 95% sandsynlighed for at et resultat er sandt, før det er værd at offentliggøre. Hvis der er 90% sandsynlighed, for at noget er rigtigt i erhvervslivet (sandsynlighed = .1), kan det ikke endegyldigt bevises, men det er formentlig bedst at handle efter at det er rigtigt og ikke forkert.

Risiko ved mange tests

Hvis du foretager et stort antal tests, kan urigtige signifikansresultater være et problem. Husk at 95% chance for at et resultat er rigtigt, betyder at der er 5% risiko for at det er forkert. Det medfører at for hver 100 tests der viser resultater med et 95% signifikansniveau, er oddsene at fem af dem er forkerte. Hvis du bruger et helt tilfældigt meningsløst datasæt til at lave 100

signifikanstests, er oddsene at fem af testene vil vise at resultaterne er signifikante uden at de er det. Jo flere tests du laver, jo større vil problemet med de falske tests være. Du kan ikke se, hvilke tests der er falske, du ved bare, at der er en sandsynlighed for, at nogle af dem viser, at resultaterne er signifikante, uden at de er det.

Opnår større sikkerhed ved gentagelse af tests

Du kan mindske problemet ved at begrænse antallet af tests til en lille gruppe, du har udvalgt inden dataindsamlingen. Hvis det ikke er muligt, er der andre løsninger. Fra et statistisk synspunkt er den bedste fremgangsmåde at gentage undersøgelsen for at se om, du får de samme resultater. Hvis et resultat viser sig at være statistisk signifikant i to undersøgelser, er det formentlig rigtigt. I praksis er det ofte ikke muligt at gentage en undersøgelse, men du kan bruge ”halverings”-teknikken, hvor du tilfældigt deler din stikprøve i to dele og laver tests på begge. Hvis noget er signifikant i begge dele, er det formentlig rigtigt. Det største problem ved denne teknik er, at når du halverer stikprøvestørrelsen, så skal en forskel være større for at være statistisk signifikant.

Problemer med tilfældighed

Den sidste mest almindelige fejl er også vigtig. De fleste signifikanstest går ud fra at du har en ægte tilfældig stikprøve. Hvis din stikprøve ikke er helt tilfældig, kan en signifikanstest muligvis overdrive nøjagtigheden af resultaterne, fordi testen kun tager højde for tilfældige fejl. Testen kan ikke tage højde for skævheder, der stammer fra ikke-tilfældige fejl, for eksempel en dårlig udvalgt stikprøve.

Opsamling:

- Signifikant behøver ikke nødvendigvis at betyde vigtigt i statistiske begreber
- Sandsynlighedsværdier skal læses omvendt ($1-p$)
- For mange signifikanstests vil vise nogle forkerte signifikante forhold
- Kontroller din stikprøveprocedure for at undgå skævheder
- Den præcise betydning af statistiske signifikans-tal

Diskussionen i det foregående afsnit anbefaler at sandsynlighedsværdierne læses omvendt ($1-p$). Fremgangsmåden vil almindeligvis føre til korrekt beslutningstagen, men det er noget af en oversimplificering set fra den tekniske side. En mere kompleks, teknisk korrekt diskussion præsenteres her.

Detaljeret fortolkning af signifikansværdier

Uheldigvis fortæller statistiske signifikansværdier ikke direkte præcis, hvad vi vil vide. De fortæller os, hvor sandsynligt det er at finde forskelle mellem grupper, der er lige så store eller større end dem vi har i stikprøven, hvis der ingen forskelle er mellem grupperne i populationen, som er repræsenteret i stikprøven. Med andre ord fortæller signifikansværdierne os, hvor sandsynlige vores data er, givet at der ingen forskelle er i populationen. Det vi gerne vil vide er, hvor sandsynligt det er, at der er forskelle i populationen givet vores data.

Logisk set vil det sige, at hvis det er tilstrækkeligt usandsynligt at finde en forskel i stikprøven, hvis der ikke er forskel i populationen, så er det sandsynligt at der er en forskel i populationen. Vi brugte denne logik i den første del af denne artikel, da vi sagde, at du kan tolke signifikansværdier ved at anse 1-p som sandsynlighed for at der er en forskel i populationen (p er signifikansværdien beregnet af programmet). Et eksempel er et signifikansniveau på 0,05, hvor du kan anse sandsynligheden for, at der er en forskel i populationen, som 95% ($1 - 0,05$).

Mens denne logik består den sunde fornufts prøve, så garanterer matematikken bag statistisk signifikans ikke at 1-p giver den eksakte sandsynlighed for at der er en forskel i populationen. Alligevel bruger mange forskere 1-p som sandsynlighed af to grunde. Den ene grund er, at ingen har opfundet et mere formålstjenstligt mål. Den anden grund er, at beregningen som regel vil føre til en brugbar tolkning af de statistiske signifikansværdier.

På nogle forskningsområder, hvor surveys ikke bruges, vil muligheden for at 1-p ikke er den eksakte sandsynlighed være mere vigtig. På disse områder vil brugen af statistiske signifikansværdier være kontroversiel.

Copyright © 1998 Creative Research Systems. Oversættelse og bearbejdelse Aspekt R&D.