

Smit Consult biedt Consultancy in Statistics

Wij bieden de beste ondersteuning en assistentie in statistische consultancy op verschillende gebieden:

1. **Het opzetten van experimenten.** Op welke wijze kunnen wij experimenten opzetten, die aan een aantal eisen moeten voldoen? Deze eisen zijn o.a. een hoge D-efficiency, het behouden van de orthogonaliteit en het verkrijgen van zoveel als mogelijk informatie van de responses. Daarnaast kan het wenselijk zijn om te toetsen of het model voor de responsevariabelen lineair is; men zal derhalve vooraf, bij het opzetten van het experiment, daar rekening mee moeten houden. Het aantal experimenten (bepaald door de instellingen van de experimentele factoren) is afhankelijk van de vereiste sensitiviteit van het uiteindelijke model, maar ook van de kosten per experiment en de aanwezige mogelijkheden (resources) om deze experimenten eenvoudig te kunnen uitvoeren. Het is van groot belang om de vooraf aanwezige informatie (1-priori) te gebruiken om het domein van de experimentele ruimte te beperken tot zinvolle experimenten. Experimenten zullen in random volgorde moeten worden uitgevoerd om effecten van volgorde te vermijden. Daarnaast zal rekening moeten worden gehouden met experimenten die onder dezelfde condities worden uitgevoerd. Het is verstandig dat als aan sommige vereisten niet kan worden voldaan, dit te verwerken in het model. Er zijn vele modellen, zoals een split plot model, of het willekeurig toekennen van experimenten aan groepen van ruwe materialen.

2. **Het analyseren van de gegevens om een model te verkrijgen.** Dit vereist een aantal technieken, waarbij uiteindelijk het model de gegevens het beste beschrijft. In het algemeen geldt dat uiteindelijk de verschillen tussen model en gegevens, de residuen, normaal verdeeld moeten worden en dat de spreiding ervan constant is voor alle mogelijke variabelen en voor elk gebied van de factoren, of onafhankelijke variabelen. De gegevens zijn tegenwoordig multivariate van aard. Soms geldt dat alle gegevens in een model kunnen worden geplaatst, soms geldt een afhankelijke en causale relatie tussen een aantal variabelen en een aantal andere variabelen. Residuen zijn derhalve ook multivariate, wat geen enkel probleem hoeft te zijn. De responsevariabelen kunnen hoog gecorreleerd zijn, waardoor technieken gebruikt kunnen worden om de dimensionaliteit van het model te reduceren. Eigenschappen van het model zijn, dat het een hoge mate van sensitiviteit heeft, maar ook dat het onderscheid met intrinsieke spreiding van metingen heeft. En dat een kleine aanpassing van de factoren die niet aanwezig zijn in het model, niet tot grote schommelingen leidt in de responses, wat een zogeheten robuust model genereert. Modellen kunnen worden vergeleken met andere modellen, om uiteindelijk het meest gewenste te verkrijgen, welke overigens wel gevalideerd moet worden voor het gewenste domein.

3. Wij hebben veel ervaring met **het presenteren van de gegevens en tonen van informatie die er echt toe doen**, zonder dat het de waarheid geweld aandoet. Wij hebben veel ervaring met multivariate gegevensanalyse en –verwerking en kunnen, samen met de materie-experts, relaties vinden en presenteren die met gewone technieken niet gevonden kunnen worden.

4. **Literatuur** is de drijvende kracht van het gebruik van statistiek, niet de software, maar ook degene die statistische technieken toepast. Vele technieken worden met vele softwareproducten toegepast, en er zijn hoegenaamd nauwelijks verschillen tussen deze producten. Er zijn wel nuances tussen deze producten en er zijn verschillende manieren om ermee te werken. Wij begonnen met SPSS in de 80-er jaren, maar hebben nu ervaring met vele producten zoals **JMP, STATGRAPHICS, SigmaPlot, EViews, GAUSS, MODDE, the Unscrambler** en vele anderen. Veel analyses worden tegenwoordige uitgevoerd met Excel, zelfs Maximum Likelihood schatters kunnen ermee worden berekend, maar er zijn ook toepassingen voor bootstrapping en simulaties. Maar specifieke analyses vereisen soms specifieke producten. Wij beheren de meeste. De grootte van de gegevensbestanden kan soms een probleem zijn, voor ons is dat slechts een uitdaging.

5. **Multivariate technieken** zijn niet gemakkelijk te gebruiken. Wij zullen de geijkte paden daarbij moeten bewandelen bij de analyse en presentatie, zoals het toepassen van een autocorrelatie-term, bij aanwezigheid ervan. Er zijn vele technieken die toegepast kunnen worden, teveel om op te noemen, zoals het toepassen van wegingen voor de rijen en kolommen, het gebruik van transformaties en terug- transformaties om uiteindelijk residuen te verkrijgen die normaal verdeeld zijn. Het gebruik van de Lack Of Fit kan een maat zijn voor de gevonden spreiding in het model.

In Excel hebben we een reeks statistische technieken opgezet zoals:

- Een aantal niet parametrische toetsen voor het vergelijken van de mediaan (gepaarde, groepen, etc.)
- Exacte toetsen (o.a. kansen, 2*2 tabellen op onafhankelijkheid)
- Een of meer factor Fixed Effects ANOVA
- Regressie analyses
- Gelijkheid van kansen
- Exacte betrouwbaarheidsintervallen van een kans (totaal en aantal successen bekend)

- Multivariate regressie
- Multivariate Logistische regressie
- PLS and PCA
- Vele smoothing technieken en classificatie analyses.

Alle analyses bevatten de volgende aspecten:

- General statistische grootheden van de data
- Test van de hypothesen
- Sensitiviteit van de toets
- Toetsen om na te gaan of
- Model specificatie
- toevoeging : bootstrapping statistische grootheden, simulaties en andere technieken

Drs Ing. J.A.C.M. (Jan) Smit is zeer numeriek ingesteld en tevens service gericht. Elke consultancy gaat samen met een beschrijving van het werk wat gedaan zal worden, van de analyses, de conclusies en de adviezen. Jan Smit is een wiskundige, met specialisatie Statistiek. Een korte CV:

- 1983 afgestudeerd aan de VU, Wiskunde
- 1983-1989 werkzaam bij SPSS Europe, Benelux
- 1990 start van bedrijf Smit Consult, Adviesbureau voor Gegevensanalyse
- 1990-1993 consultancy bij verschillende bedrijven en instellingen
- 1993 begonnen met distributie van software voor statistische analyses
- 1999 uitbrengen van software producten SC-Biplot en SC-graph voor portfolio analyses en multivariate analyses
- 2001 overname van activiteiten van Oasis BV te Nieuwegein
- 2002 overname van Cosinus Computing BV (vanaf toen had ons bedrijf deze naam)
- 2002-2012 ondersteuning software voor statistische analyses
- Augustus 2013 aanbidding van consultancy

Onze missie: werken met de industrie en instellingen om te komen tot verbeterde of nieuwe producten, nieuwe inzichten in de materie, door middel van statistische analyses. Met het gebruik van statistiek kan men solide en gevoelige statistische informatie verkrijgen die een uitgangspunt kunnen zijn voor het beleid, keuze van aanpassingen, keuze voor toekennen, enzovoort. Met behulp van de statistiek kunnen wij modelleren, wij kunnen nagaan wat wij kunnen verwachten, wij kunnen de effecten van maatregelen berekenen. Met de statistiek kunnen wij de structuur van de beschikbare informatie en de toegevoegde marktinformatie, of andere informatie omzetten in een model. Statistiek is ondersteunend. Met statistiek kunnen wij nagaan wat te doen om risico's te verminderen, en producten produceren onder gelijke kwaliteit.

Ons tarief is afhankelijk van de te verwachten uren, de duur van het project en de afstand tot de instelling. Wij kunnen de opdracht op ons kantoor uitvoeren. Onderdeel van het project kan een cursus zijn in het gebruik van de statistiek technieken. Het project kan veelal alleen goed worden uitgevoerd samen met de experts van de materie zelf.

Wij zien er naar uit u van dienst te zijn. (Voor meer informatie neemt u contact met ons op :
tel: +31 (0) 416 378125 of per e-mail: jan.smit@smitconsult.nl.