Tentamen Regressie en Variantie-Analyse (191530440)
Vrijdag 12 april 2013 van 8:45 tot 11:45 uur

Dit tentamen bestaat uit 6 vragen en 1 bonuspuntvraag.

Als er gevraagd wordt naar een toets, geef dan aan (voor zover nog niet gegeven) :
(1) De modelveronderstellingen (het "statistisch model"),
(2) De formulering van de nulhypothese en alternatieve hypothese,
(3) De formule voor de toetsingsgrootheid, indien er meerdere formules mogelijk zijn kies dan voor de formule die rekentechnisch het eenvoudigst is.
(4) De verdeling van de toetsingsgrootheid onder de nulhypothese,
(5) Geef aan voor welke waarden van de toetsingsgrootheid de nulhypothese verworpen moet worden. (Als T de toetsingsgrootheid is, verwerp je dan de nulhypothese voor $T \geq c$? Of voor $T \leq -c$? Of voor zowel $T \geq c$ als $T \leq -c$?)

Vraag 1
Uitgaande van het algemene regressiemodel met normale verdelingen, is de schatter $\hat{\beta}$ $N(\beta, \sigma^2 (X^T X)^{-1})$-verdeeld. Toon dit aan.

Vraag 2
Toon aan dat $a^T \hat{\beta}$ de BLUE is van $a^T \beta$ als de Gauss-Markov-voorwaarden van het regressiemodel vervuld zijn.

Vraag 3
Geef aan hoe we het model van de variantie-analyse met 2 factoren **met** interactie in de vorm "$Y = X\beta + \varepsilon$" kunnen schrijven, en pas algemene toetsingstheorie van regressiemodellen toe om te toetsen of de interactietermen γ_{ij} in het model thuishoren. (Bij het beantwoorden van deze vraag worden voor kansverdelingen en formules van toetsingsgrootheden **geen bewijzen** verwacht.)

Vraag 4
Beschouw het algemene regressiemodel met normale verdelingen: $Y \sim N(X\beta, \sigma^2 I_n)$, met X een $n \times (k+1)$ matrix van rang $k+1$. Beschrijf de constructie van betrouwbaarheidsintervallen voor β_k. Geef daarbij ook aan wat de geschatte standaardafwijking ('standard error') van de betrokken schatter is.

Vraag 5
Beschouw het regressiemodel met modelvergelijking
$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i$
waar de storingen ε_i voldoen aan een AR(1)-model. Laat zien dat $cov(\varepsilon) = \sigma^2 \Omega$ geldt met Ω een matrix die alleen van een correlatiecoëfficiënt ρ afhangt. Geef de formule voor de 'estimated generalized least squares estimator' van de vector β.

Vraag 6
Ga uit van het algemene regressiemodel (k verklarende variabelen, model met constante β_0) met de Gauss-Markov-voorwaarden. Toon aan dat
$S^2 = \sum_{i=1}^{n} E_i^2 / (n-k-1)$ een zuivere schatter is van σ^2.

BONUSPUNTVRAAG

Vraag B
Beschouw het algemene regressiemodel met normale verdelingen:
$Y \sim N(X\beta, \sigma^2 I_n)$,
waar X een $n \times (k+1)$ matrix is van rang $k+1$ ($k+1 < n$). We toetsen de nulhypothese $H_0 : (\beta_{r+1}, \ldots, \beta_k) = 0$ tegen de alternatieve hypothese $H_1 : (\beta_{r+1}, \ldots, \beta_k) \neq 0$. Bewijs dat de algemene toetsingsgrootheid F ook te schrijven is als een functie van twee 'residual sums of squares'.

Normering:
De vragen 1 t/m 6 worden beoordeeld met een "rapportcijfer" (cijfer tussen 1 en 10). Tentamencijfer is hiervan het gemiddelde, eventueel verhoogd met een bonuspunt verdiend met de bonuspuntvraag.