Tentamen Regressie en Variantie-Analyse (191530440)
Woensdag 3 juli 2013 van 13:45 tot 16:45 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
We gaan uit van het algemene regressiemodel met k verklarende variabelen. Toon aan dat de som van de residuen altijd nul is als het model de constante β_0 bevat.

Vraag 2
Stel we hebben onderling onafhankelijke stochastische variabelen Y_i die verdeeld zijn volgens:

$$Y_i \sim N(\beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}, \sigma^2),$$

waar $\beta_1, \beta_2, \beta_3$ en σ^2 onbekende parameters zijn en de x_{ij}'s gegeven getallen. Geef formules voor de zuivere schatters van de onbekende parameters. Geef ook de verdelingen aan van de schatters. Zijn de schatters onafhankelijk van elkaar? Geef aan hoe we de nulhypothese $H_0 : (\beta_1, \beta_2) = 0$ kunnen toetsen tegen de alternatieve hypothese $H_1 : (\beta_1, \beta_2) \neq 0$. (Bij het beantwoorden van deze vraag worden voor kansverdelingen en formules van toetsingsgrootheden **geen bewijzen** verwacht.)

Vraag 3
Geef aan hoe we het model van de variantie-analyse met 1 factor in de vorm "$Y = X\beta + \varepsilon$" kunnen schrijven, en pas algemene toetsingstheorie van regressiemodellen toe om de nulhypothese $H_0 : (\alpha_1, ..., \alpha_k) = 0$ te toetsen tegen de alternatieve hypothese $H_1 : (\alpha_1, ..., \alpha_k) \neq 0$. (Bij het beantwoorden van deze vraag worden voor kansverdelingen en formules van toetsingsgrootheden **geen bewijzen** verwacht.)

Vraag 4
Beschouw enkelvoudige lineaire regressie zonder constante β_0:

$$Y_i = \beta_1 x_i + \varepsilon_i \qquad (i = 1,2,\dots,n),$$

waarbij we aannemen dat de storingen ε_i onderling onafhankelijk en $N(0,\sigma^2)$-verdeeld zijn. Beschrijf de constructie van betrouwbaarheidsintervallen voor β_1. Geef daarbij ook aan wat de geschatte standaardafwijking ('standard error') van de betrokken schatter is.

Vraag 5
We gaan uit van het algemene regressiemodel met k verklarende variabelen. Toon aan dat de vector van residuen E en de vector $\hat{\beta}$ (schatter van β) ongecorreleerd zijn als de Gauss-Markov-voorwaarden vervuld zijn, en onderling onafhankelijk zijn als $Y \sim N(X\beta, \sigma^2 I)$.

Vraag 6
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 $k+1$ $(k+1 < n)$. Toon aan dat

$$\frac{(n-k-1)S^2}{\sigma^2} = \frac{\sum_i E_i^2}{\sigma^2}$$ een χ^2_{n-k-1}-verdeling heeft.

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 $H_0 : (\beta_{r+1},\dots,\beta_k) = 0$ tegen de alternatieve hypothese $H_1 : (\beta_{r+1},\dots,\beta_k) \neq 0$. 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.