Herentamen Inleiding Statistiek

maart 2010

Bij het tentamen mogen boek en aantekeningen gebruikt worden.

Een vriend heeft me een merkwaardige dobbelsteen gegeven. Op basis van de vorm van de dobbelsteen krijg ik het idee dat de twee kansen of een "1" resp. op een "6", en dat de vier kansen op een "2", "3", "4" resp. op een "5", aan elkaar gelijk zouden kunnen zijn. Maar zeker weet ik niks. Ik besluit om de dobbelsteen n maal te worpen, en om de uitkomsten – een rij van n getallen uit de verzameling $\{1,2,3,4,5,6\}$ – te beschouwen als onafhankelijke en identiek verdeelde trekkingen uit een willekeurige discrete verdeling op $\{1,2,3,4,5,6\}$. Definieer p_j als de kans dat een van de n worpen in een "j" resulteert, $j = 1, ..., 6$, met uiteraard $p_j \geq 0$ en $\sum_{j=1}^{6} p_j = 1$, en definieer de random variabele N_j als het aantal keer dat we een "j" zien in de n worpen.

(1) Gegegeven de data $N_j = n_j, j = 1, ..., 6$, wat is de "log likelihood functie" voor de parameters $p_1, ...p_6$?

(2) Leid de meest-aannemelijke schatters van $p_1, ..., p_6$ af, (i) zonder enige beperking op deze parameters, (ii), onder de veronderstelling dat $p_1 = p_6$ én $p_2 = p_3 = p_4 = p_5$.

(3) Ik besluit om de nul-hypothese H_0: "$p_1 = p_6$ én $p_2 = p_3 = p_4 = p_5$" te toetsen, tegen het alternatief H_A dat H_0 niet waar is. Leid een "generalized likelihood ratio toets" af voor dit toetsingsprobleem, en geef aan hoe U met grote steekproef theorie een toets van niveau α ongeveer gelijk aan 0.05 zou uitvoeren.

(4) Geef een alternatieve maar bij benadering equivalente vorm van de toets met behulp van de relatie tussen zogenaamde chi-kwadraat toetsen voor hypothesen over kansen in kruistabellen, en de generalized likelihood ratio toets

(5) Onder de nul-hypothese kunnen we een nieuwe parameter $\phi = p_1 + p_6$ invoeren, waarvoor dan geldt $p_1 = p_6 = \phi/2$, $p_2 = p_3 = p_4 = p_5 = (1 - \phi)/4$. Gebruik de grote steekproef theorie van meest-aannemelijke schatters om een benaderend 95% betrouwbaarheidsinterval voor ϕ af te leiden, onder de aanname dat de nul-hypothese waar is, en vergelijk deze met het resultaat van een voor de hand liggende interval gebaseerd op de binomiale verdeling (met parameters n en ϕ), van $N_1 + N_6$.