

STATISTIEK (WIS/FYS)

(16/06/2008)

- 1 In een groep van 23 studenten spreken er 17 Frans en 21 Engels. Elke student spreekt minstens één taal. Op hoeveel manieren kan je dan 2 studenten uit deze groep kiezen zodat er eentje Frans kan praten en eentje Engels kan spreken?
- 2 De vraag gaat over volgend artikel.

Vrouwen met een calorierijk dieet krijgen vaker een jongetje.



Aspirant-ouders die vurig hopen op een jongetje, hoeven zich niet langer uitsluitend met gebeden tot een hogere macht te wenden. Nee, wie de kansen op een blauwe kinderkamer wil vergroten, kan ook zelf iets doen. Brits onderzoek wijst namelijk uit dat vrouwen die voorafgaand aan de conceptie een calorierijk ontbijt eten, vaker een jongetje baren.

Het gaat om een studie van de universiteiten van Exeter en Oxford. Wetenschappers daar volgden 740 vrouwen vóór, tijdens en na hun zwangerschap. Ze ontdekten dat 56% van de vrouwen met een energierijk dieet een jongetje kreeg, tegenover 45% van de vrouwen met caloriearm dieet. Het eten van muesli vergrootte de verschillen nog verder: van de vrouwen die dagelijks een kommetje ontbijtgranen naar binnen werkten, schonk maar liefst 59% het leven aan een jongen, tegenover 43% van de vrouwen die dat minder dan één keer per week deden.

In het artikel wordt niets vermeld over hoeveel vrouwen uit het onderzoek een calorierijk dieet volgden.

- (a) Stel dat 300 vrouwen calorierijk aten. Ga je dan akkoord met de titel van het artikel? Gebruik $\alpha = 0.05$.
- (b) Verantwoord je keuze van H_0 en H_1 .
- (c) Stel een 95%-betrouwbaarheidsinterval op voor de populatieproportie vrouwen die een calorierijk dieet volgen en een jongetje krijgen. Neem hierbij weer aan dat 300 vrouwen dat dieet volgden.
- (d) Wat is de betekenis van dit betrouwbaarheidsinterval?
- (e) Wat gebeurt er met het betrouwbaarheidsinterval als de steekproefgrootte vergroot? Interpreteer je resultaat.
- 3 Beschouw het volgend kansspelletje. Ik gooi een dobbelsteen op. Als het aantal ogen 1, 2 of 3 is, dan win ik 1 euro. Anders verlies ik 1 euro.
- (a) Bepaal de gemiddelde winst na n worpen.
- (b) Bepaal de variantie van de winst na n worpen. Interpreteer dit resultaat.
- (c) Stel dat ik begin met een startkapitaal van 5 euro en mijn tegenstander begint met 10 euro. Als ik win, krijg ik 1 euro van hem. Als ik verlies, moet ik hem 1 euro. Deze regels gelden ook voor hem, en de spelers mogen schulden hebben. Bepaal de kans dat na 25 beurten geen enkele speler schulden heeft.
- 4 (Schets van de vraag.) Het blijkt dat supernova's zich verwijderen van ons. Hubble vermoedde dat de verwijderingssnelheid v en de afstand d van de supernova t.o.v. ons gerelateerd zijn via de vergelijking $v = Hd$, met H de Hubble constante. Tegenwoordig vermoedt men dat $H = 50$. Om dit te testen, heeft men v en d bepaald voor 24 supernova's. De ANOVA-tabel (enkel de Model en Error rijen) en een tabel met

standard errors en schatters e.d. van het intercept en richtingscoëfficiënt zijn gegeven. Verder mag je veronderstellen dat de modelveronderstellingen voor lineaire regressie voldaan zijn.

- (a) Bepaal R^2 .
- (b) Ga na of het model van Hubble zinvol is aan de hand van volgende vraagjes. Gebruik steeds $\alpha = 0.05$:
 - Test of het intercept 0 is.
 - Test of de richtingscoëfficiënt significant verschilt van 0.
 - Test of de schatting van de Hubble constante significant verschilt van 50.
- (c) Schat de verwijderingssnelheid van een supernova die op 10 Megaparsec (eenheid ingevoerd in de vraagstelling) van ons ligt.

5 We willen de hypotheseset

$$H_0 : \mu = \mu_0 \quad \longleftrightarrow \quad H_1 : \mu > \mu_0$$

uitvoeren op het populatiegemiddelde van een normaal verdeelde variabele waarvan σ^2 niet gekend is. Hiervoor berekenen we de bij de t -test behorende p -waarde. Omdat we dit kunnen berekenen voor elke steekproef, kunnen we de p -waarde beschouwen als een variabele. Toon aan dat, onder H_0 , de dichtheid van de p -waarde uniform is op $[0, 1]$.

Hint: Toon aan dat de verdelingsfunctie gelijk is aan de verdeling van de uniforme verdeling.