

Fasit til STK1000 eksamen V06

1 Oppgave 1

- a) Krav til binomisk fordeling: 1) n uavhengige enkelt utfall 2) suksess / ikke suksess 3) samme sannsynlighet for suksess i alle utfall 4) uavhengighet mellom enkelt utfallene. Dvs. Binomisk fordeling såsant kjønn i ulike fødsler er uavhengig (bør se bort fra flerfødsler)
- b) i) 0.2461 ii) 0.3770
- c) testobservator $z = 2.86$ og $p = 0.004$

2 Oppgave 2

- a) $H_0 : \mu_1 = \mu_2$ mot $H_a : \mu_1 \neq \mu_2$
Forkaster H_0 : siden $p < 0.0005$, dvs. signifikant forskjell i vekt mellom bamser og binner.
- b) Testen med ulike standardavvik ble benyttet. Dette er fornuftig siden standardavvik for bamser (120) var nesten dobbelt så stort som for binner (64.5)
- c) $\bar{x} \pm t^* SE$ der $t^* = 97.5$ persentil i $t_{135} (= 1.978)$ og $SE = \sqrt{(s_1^2/99 + s_2^2/44)}$.
Fortolkning 95% KI: Vil ha med faktisk verdi av differansen i forventninger i 95% av tilfellene hvis tilsvarende studie ble gjentatt mange ganger, dvs. vi kan føle oss 95% sikre på at verdien er med i dette aktuelle intervallet.
- d) R^2 er forklart andel variasjon, $R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$
Her er $R^2 = 93.3\%$, forklarer altså ganske mye. Dessuten residualt standardavvik = 28.7 som er tilnærmet lik standardavvik til en predikert verdi.
- e) Residual $e_i = y_i - \hat{y}_i$. Forskjellige typer plott kan avsløre 1) ikke-linearitet, 2) ikke-konstant varians, 3) avhengighet, 4) avvik fra normalfordelt feilledd og 5) outliere
Her ligger residualene langs en kurve (ikke rett linje rundt $e_i = 0$). Tyder på en kurvatur mellom x-er og y-er (som også er mulig å observere fra plottet i punkt d)).
- f) Modell: $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$ der ϵ_i er $N(0, \sigma)$ og uavhengige
Estimat for β_0 : $b_0 = 12.65$, for β_1 : $b_1 = -3.166$, for β_2 : $b_2 = 0.21247$ og for σ : $S = 23.33$.
Nå er det ikke lenger mulig å se noen ikke-linearitet, men det virker som om variasjonen i residualene vokser her med verdien av x
- g) Modell: $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_{i3} + \epsilon_i$ der β_3 angir forskjellen mellom binner og bamser, dvs. $H_0 : \beta = 0$ med p-verdi lik 0.066. Ikke-signifikant forskjell! i motsetning til signifikant forskjell i punkt a). Kjønnforskjellen i punkt a) kan nok forklares med at bamsene har større brystomkrets enn binnene.