NTNU
Kunnskap for ei betre verd

Lineær regresjon

Torstein Fjeldstad
Institutt for matematiske fag, NTNU
08.11.2018

I dag

- Praktisk informasjon
- Fullføre kap. 10 (hypotesetest)
- Lineær regresjon

Informasjon

Informasjon I

— Sara vil vere vikar tysdag 20.11 (repetisjon blokk 1)
— Du er sjølve ansvarleg for at du har godkjend øvingsopplegg
— Eksamen onsdag 28.11
— Det vil bli spørjetimer/statistikklab før eksamen. Tidspunkt vil bli lagt ut på heimesida snart

Informasjon II

Tentativ plan

Dato	Plan
08.11	Avslutning av hypotesetest og introduksjon til lineær regresjon
13.11	Eigenskapar til estimatorar (lineær regresjon)
15.11	Prediksjonsintervall og residualplott
20.11	Oppsummering blokk 1
22.11	Oppsummering blokk 2

Informasjon III

Kræsjkurs onsdag 21.11 klokka 16:15 i F1

1. Send forslag til oppgåver/tema til torstein.fjeldstad@ntnu.no innan onsdag 14.11

NOKUT undersøking

NOKUT undersøking

Du kan vinne gåvekort på 5 000 kr

Hypotesetest to-utval

Eksempel meiningsmåling

Meiningsmåling med kun to alternativ ja/nei.
Utfører meiningsmålinga to gonger og vil avgjere om andelen ¨ja¨ har endra seg.

Eksempel meiningsmåling

Meiningsmåling med kun to alternativ ja/nei.
Utfører meiningsmålinga to gonger og vil avgjere om andelen ¨ja¨ har endra seg.

- 1. meiningsmåling: spør n personar, X svarar ¨ja¨, dvs. $X \sim b(x; n, p_1)$
- 2. meiningsmåling: spør m personar, Y svarar ¨ja¨, dvs. $Y \sim b(x; m, p_2)$

Eksempel meiningsmåling

Meiningsmåling med kun to alternativ ja/nei.
Utfører meiningsmålinga to gonger og vil avgjere om andelen ¨ja¨ har endra seg.

- 1. meiningsmåling: spør n personar, X svarar ¨ja¨, dvs. $X \sim b(x; n, p_1)$
- 2. meiningsmåling: spør m personar, Y svarar ¨ja¨, dvs. $Y \sim b(x; m, p_2)$

Har andelen ¨ja¨ endra seg?

Enkel lineær regresjon

Enkel lineær regresjon

Situasjon: har observert $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$.

— $x_1, x_2, \ldots, x_n$ er kjende tal

— $y_1, y_2, \ldots, y_n$ er realisasjonar frå uavhengige stokastiske variablar $Y_1, Y_2, \ldots, Y_n$ med

$$Y_i \sim n(y_i; \alpha + \beta x_i, \sigma)$$

Mål (i dag): estimere α, β og σ^2

— Minste kvadraters metode

— Sannsynsmaksimeringsprinsippet

SME enkel lineær regresjon I

Maksimer log-rimelighetsfunksjon

$$l(\alpha, \beta, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \alpha - \beta x_i)^2$$

Får føgljande likningssystem:

$$n\alpha + \beta \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \tag{1}$$

$$\alpha \sum_{i=1}^{n} x_i + \beta \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \tag{2}$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \alpha - \beta x_i)^2 = n \tag{3}$$

SME enkel lineær regresjon II

SME

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \qquad \widehat{\alpha} = \bar{Y} - \widehat{\beta}\bar{x} \qquad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{\alpha} - \widehat{\beta}x_i)^2$$

Tysdag

— Fortsetjing lineær regresjon