



**LUNDS UNIVERSITET**  
Lunds Tekniska Högskola

*Kursplan för*

## **Linjär och logistisk regression med datainsamling**

### **Linear and Logistic Regression with Data Gathering**

**FMSN40, 9 högskolepoäng, A (Avancerad nivå)**

**Gäller för:** Läsåret 2014/15

**Beslutad av:** Utbildningsnämnd B

**Beslutsdatum:** 2014-04-08

#### **Allmänna uppgifter**

**Huvudområde:** Teknik.

**Alternativobligatorisk för:** I3

**Undervisningsspråk:** Kursen ges på begäran på engelska

#### **Syfte**

Regression handlar om att modellera hur en egenskap (längd, vikt, pris, koncentration, etc) samvarierar med en eller flera andra egenskaper (kön, boyta, omsättning, temperatur, etc). Linjär regression introduceras i grundkursen i matematisk statistik men här fyller vi på med, bl.a., "hur kontrollerar jag om modellen passar till data", "vad gör jag om den inte passar", "hur osäker är den" och "hur använder jag den för att dra slutsatser om verkligheten".

När man t.ex. gjort en enkätundersökning där folk kan svara ja/nej eller lite/lagom/mycket eller bil/cykel/buss eller något annat kategoriindelad kan man inte använda linjär regression. Då behövs logistisk regression istället. Andra halvan av kursen handlar om detta.

Som del i kursen ska man själv konstruera en enkät eller försöksplan för en valfri frågeställning, samla in data och analysera dem med lämplig regressionsmodell.

#### **Mål**

*Kunskap och förståelse*

För godkänd kurs skall studenten

- Beskriva skillnaderna mellan kontinuerliga och diskreta data och vilka konsekvenser detta får för valet av statistisk modell,
- Redogöra för principerna bakom olika skattningsprinciper,
- Beskriva de statistiska egenskaperna hos sådana skattningar som förekommer i regressionsanalys,
- Tolka regressions samband i termer av betingade fördelningar,
- Förklara begreppen odds och oddskvot och beskriva deras samband med sannolikheter och med logistisk regression.

#### *Färdighet och förmåga*

För godkänd kurs skall studenten

- Ställa upp en multipel linjär regressionsmodell för ett konkret problem,
- Ställa upp en multipel logistisk regressionsmodell för ett konkret problem,
- Skatta parametrarna i regressionsmodellen och tolka dem,
- Validera modellens giltighet och göra lämpliga modifieringar av modellen,
- Använda den framtagna modellen för prediktion,
- Använda något statistiskt datorprogram för att analysera regressionsdata och tolka resultaten,
- Redovisa analys av och slutsatser från ett praktiskt problem i en skriftlig rapport samt presentera dessa muntligt.
- Konstruera en mall för datainsamling som kan användas för att besvara en konkret frågeställning.

#### *Värderingsförmåga och förhållningsätt*

För godkänd kurs skall studenten

- Alltid kontrollera förutsättningarna innan han/hon ansätter en regressionsmodell,
- Värdera rimligheten i en genomförd studie,
- Reflektera över den valda modellens och skattningsmetodens begränsningar samt möjliga alternativa lösningsmetoder.

## **Kursinnehåll**

Minsta-kvadrat- och maximum-likelihood-metoden; Oddskvoter; Multipel linjär och logistisk regression; Matrisformulering; Metoder för modellvalidering, residualer, outliers, inflytelserika observationer, multikolinjäritet, variabeltransformationer; Val av regressorer, F-test, likelihood-kvot-test; Konfidensintervall och prediktion. Något om korrelerade fel, poissonregression samt multinomial och ordinal logistisk regression. Enkätkonstruktion och försöksplanering.

## **Kursens examination**

**Betygsskala:** TH

**Prestationsbedömning:** Examination sker skriftligt och muntligt i form av projektredovisning, skriftlig och muntlig opposition samt muntlig tentamen.

**Delmoment**

**Kod:** 0114. **Benämning:** Tentamen.

**Antal högskolepoäng:** 5,5. **Betygsskala:** TH. **Prestationsbedömning:** Muntlig tentamen

**Kod:** 0214. **Benämning:** Projekt.

**Antal högskolepoäng:** 3,5. **Betygsskala:** UG. **Prestationsbedömning:** Insamling av, eller beskrivning av befintligt, eget datamaterial, skriftlig projektrapport med muntlig presentation samt opponering/granskning av

annat projekt.

## Antagningsuppgifter

### Förkunskapskrav:

- FMS012 Matematisk statistik, allmän kurs eller FMS032 Matematisk statistik, allmän kurs eller FMS035 Matematisk statistik, allmän kurs eller FMS086 Matematisk statistik eller FMS140 Matematisk statistik, allmän kurs

**Begränsat antal platser:** Nej

**Kursen överlappar följande kurser:** FMSN30, MASM22

## Kurslitteratur

- Rawlings, J.O., Pantula, S.G., Dickey, D.A.: Applied Regression Analysis - A Research Tool, 2ed. Springer, 1998, ISBN: 0-387-98454-2. Finns som e-bok.
- Alan Agresti: An introduction to categorical data analysis, 2nd ed. Wiley, 2007. Finns som e-bok.

## Kontaktinfo och övrigt

**Studierektor:** Anna Lindgren, [studierektor@matstat.lu.se](mailto:studierektor@matstat.lu.se)

**Hemsida:** <http://www.maths.lth.se/matstat/kurser/fmsn30/>