

*Kursplan för*

# Linjär och logistisk regression med datainsamling Linear and Logistic Regression with Data Gathering

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

**Gäller för:** Läsåret 2022/23

**Fakultet:** Lunds tekniska högskola

**Beslutad av:** Programledning I

**Beslutsdatum:** 2022-04-11

## Allmänna uppgifter

**Huvudområde:** Teknik.

**Alternativobligatorisk för:** I3

**Undervisningsspråk:** Kursen ges 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 förhållningssätt*

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, Ftest, likelihood-kvot-test; Konfidensintervall och prediktion. Något om korrelerade fel, poissonregression samt multinomial och ordinal logistisk regression. Enkätkonstruktion och försöksplanering.

## Kursens examination

**Betygsskala:** TH - (U,3,4,5) - (Underkänd, Tre, Fyra, Fem)

**Prestationsbedömning:** Skriftlig och muntlig projektredovisning, kamratgranskning och muntlig tentamen.

Om så krävs för att en student med varaktig funktionsnedsättning ska ges ett likvärdigt examinationsalternativ jämfört med en student utan funktionsnedsättning, så kan examinator efter samråd med universitetets avdelning för pedagogiskt stöd fatta beslut om alternativ examinationsform för berörd student.

### **Delmoment**

**Kod:** 0117. **Benämning:** Tentamen.

**Antal högskolepoäng:** 3. **Betygsskala:** TH. **Prestationsbedömning:** Muntlig tentamen

**Kod:** 0217. **Benämning:** Projekt 1.

**Antal högskolepoäng:** 1,5. **Betygsskala:** UG. **Prestationsbedömning:** Skriftlig projektrapport och kamratgranskning **Delmomentet omfattar:** Linjär regression

**Kod:** 0317. **Benämning:** Projekt 2.

**Antal högskolepoäng:** 1,5. **Betygsskala:** UG. **Prestationsbedömning:** Skriftlig projektrapport och kamratgranskning **Delmomentet omfattar:** Logistisk regression

**Kod:** 0417. **Benämning:** Projekt 3.

**Antal högskolepoäng:** 2,5. **Betygsskala:** UG. **Prestationsbedömning:** Skriftlig projektplan, datainsamling och muntlig projektpresentation. **Delmomentet omfattar:** Eget regressionsproblem

**Kod:** 0517. **Benämning:** Laborationer.

**Antal högskolepoäng:** 0,5. **Betygsskala:** UG. **Prestationsbedömning:** Datorlaborationer

## **Antagningsuppgifter**

### **Förkunskapskrav:**

- FMSF20 Matematisk statistik, allmän kurs eller FMSF25 Matematisk statistik - kompletterande projekt eller FMSF32 Matematisk statistik eller FMSF45 Matematisk statistik, allmän kurs eller FMSF50 Matematisk statistik, allmän kurs eller FMSF55 Matematisk statistik, allmän kurs eller FMSF70 Matematisk statistik eller FMSF75 Matematisk statistik, allmän kurs eller FMSF80 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, ISBN: 978-0-471-22618-5. Finns som e-bok.

## **Kontaktinfo och övrigt**

**Studierektor:** Johan Lindström, [studierektor@matstat.lu.se](mailto:studierektor@matstat.lu.se)

**Hemsida:** <http://www.ctr.maths.lu.se/course/MASM22/>

**Övrig information:** Endast en av kurserna FMSN30 och FMSN40 får ingå i examen.