

Kursplan för

Linjär och logistisk regression

Linear and Logistic Regression

**FMSN30, 7,5 högskolepoäng, A
(Avancerad nivå)**

Gäller för: Läsåret 2020/21

Beslutad av: Programledning I

Beslutsdatum: 2020-04-03

Allmänna uppgifter

Valfri för: BME4, D4, F4, F4-fm, F4-mai, I4, L4-fe, Pi4-fm, Pi4-biek, Pi4-bam, MMSR1

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.

Om man har ett eget datamaterial som lämpar sig för linjär eller logistisk regression kan man få analysera detta som en del av projektet.

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.

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, multikolinjaritet, variabeltransformationer; Val av regressorer, Ftest, likelihood-kvot-test; Konfidensintervall och prediktion. Något om korrelerade fel, poissonregression samt multinomial och ordinal logistisk regression.

Kursens examination

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

Prestationsbedömning: Skriftlig och muntlig projektredovisning, kamratgranskning av projekt samt 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: 1. **Betygsskala:** UG. **Prestationsbedömning:** Muntlig projektredovisning **Delmomentet omfattar:** Andra regressionsmodeller

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 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

Begränsat antal platser: Nej

Kursen överlappar följande kurser: MASM22, FMSN40

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

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

Övrig information: Kursen ges även på naturvetenskaplig fakultet med kurskoden MASM22. Endast en av kurserna FMSN30 och FMSN40 får ingå i examen.