

TNIU66: Statistik och sannolikhetslära

Kursinformation 2018

Kunskaper i statistik och sannolikhetslära är väsentliga för nästan alla människor i ett modernt samhälle. Vi översköljs av påståenden baserade på statistik av mer eller mindre vederhäftig karaktär, och behöver förmågan att kunna skilja det värdefulla från det rena tramset. Som universitetsutbildad inom ett kvantitativt ämne (vare sig det är logistik i någon form eller som ingenjör) har du ett särskilt ansvar att se till att den statistik som används främjar ett gott samhälle och klargör samband mer än förvirrar. Den här kursen bär inte hela den vägen, men den är en god start för ett kreativt användande av starka matematiska verktyg i resten av din utbildning, och förhoppningsvis även i livet efter studierna.

Kursens mål, innehåll och förväntade läranderesultat

Kursens mål är att ge en introduktion till matematisk modellering av slumpmässiga försök och till statistiska metoder och begrepp. Den skall även visa på tillämpningar relevanta för de program för vilken den ges.

Innehållet och de förväntade läranderesultaten är tagna ur kursplanen, och återges här för fullständighetens skull. Före kursens början är dessa svåra att läsa, men jag rekommenderar att du återvänder till dem flera gånger under kursens gång.

Innehåll

Sannolikhetslära: Sannolikhetsbegreppen. Utfallsrum och händelse. Mängdlära och kombinatorik. Betingade sannolikheter, med Bayes sats, och begreppet oberoende händelser. Diskreta och kontinuerliga stokastiska variabler med fördelningsmått såsom väntevärde, varians/standardavvikelse, kovarians och korrelation. De vanligaste sannolikhetsmodellerna, inkluderande likformig-, binomial-, poisson-, exponential- och normalfördelningsmodellen. Centrala gränsvärdessatsen.

Statistik: Deskriptiv statistik med begreppen medelvärde, median och standardavvikelse samt visualiseringar. Punkt- och intervallskattningar av väntevärde (med och utan känd standardavvikelse), sannolikhet och intensitet. Punktskattning av varians och standardavvikelse. Hypotesprövning, inklusive p -värde och styrkefunktion. Korrelation. Linjär regression för två variabler med kontroll av förutsättningar samt konfidens- och prediktionsintervall.

Förväntade läranderesultat

Efter genomförd kurs skall du kunna:

1. analysera fördelningen hos en datamängd avseende centralvärde och spridning, såsom medelvärde och median samt standardavvikelse, samt visualisera detta.
2. redogöra för olika synsätt på begreppet sannolikhet.
3. beräkna sannolikheter för händelser, med användning av begrepp och verktyg som oberoende, betingning, oförenlighet, komplementhändelse, union, snitt, kombinatorik, trädidiagram.
4. formulera en sannolikhetsmodell med hjälp av stokastiska variabler, även med centrala gränsvärdesatsen, och använda den för att bestämma egenskaper hos dess fördelning samt beräkna sannolikheter.
5. beräkna punktskattningar av väntevärde, varians, standardavvikelse, sannolikhet och intensitet samt bedöma deras lämplighet.
6. beräkna konfidensintervall för väntevärde (med och utan känd standardavvikelse), sannolikhet och intensitet samt tolka resultatet.
7. formulera och genomföra en hypotesprövning, och däri kunna tolka begreppen styrkefunktion och p -värde.
8. genomföra en korrelationsanalys och tolka resultatet.
9. ställa upp och tolka en linjär regressionsmodell med två variabler, avgöra om en linjär modell är tillämplig, samt bedöma tillförlitligheten hos skattningar av såväl väntevärden som enskilda observationer.
10. använda datorstöd för alla beräkningar där det är relevant.

Dessa tio punkter kopplar direkt till kursmomenten och till examinationen, enligt nedan.

Förkunskaper

En högskolekurs inom matematik, exempelvis matematisk grundkurs för högskoleingenjörstudenter eller matematik för SL och FTL. Gymnasiematematik, kurserna Ma1 – Ma3 (eller motsvarande). I den mån repetition är nödvändig är det *ditt* ansvar att se till detta (givetvis kan du fråga dina lärare om detaljer, men de saknar tiden att repriser hela kurser).

Litteratur, webbstöd, programvara och miniräknare

Kurslitteratur är boken Statistik – metoder och tillämpningar, av Gunnar G. Løvås, utgiven av Liber AB, Malmö, 2006. Bokakademien är meddelad, och även universitetsbiblioteket har fått besked om att den här kursboken används.

På webben finns mycket stöd för den som önskar ta del av sådant. Dessvärre finns även sådant som är av mer tvivelaktig karaktär (läs exempelvis ”statistik” på svenska wikipedia), varför det krävs ett kritiskt granskande av avsändaren för att kunna bedöma trovärdigheten. Likaså kan beteckningar och konventioner skilja sig mellan det som hittas på webben och de som vi använder i den här kursen. Jag kommer att lägga upp länkar till Khan Academy som inspelad

föreläsning och som stöd för för den som vill höra ytterligare en muntlig förklaring utöver under-tecknads (eller istället för), men poängterar samtidigt att även dessa enbart är ett komplement till kurslitteraturen.

Den programvara som används vid laborationerna är Excel. Orsaken är att jag vill att du skall ha lätt att göra statistiska analyser även efter avslutad kurs. Excel är ett program som du med stor sannolikhet kommer att ha tillgång till även efter avslutad utbildning.

Kursens webbsida ligger under Lisam. **Du får tillgång till kursidan när du registrerar dig på kursen.**¹ Där kan du hitta det som hör till kursen, bl.a. denna kursinformation, alla bilder som visas på föreläsningarna, exempelsamling, laborationsuppgifter samt schema. Där finns också länkar till relevanta delar av Khan Academy. Inlämning av laborationsrapporter sker enbart via Lisam, liksom återkoppling. Detta av den praktiska orsaken att det kommer att lämnas in drygt hundra rapporter i kursen vilka skall bedömas och återkopplas på ett säkert sätt.

I kursen behövs en miniräknare. En grafritande sådan är att föredra, då den oftast innehåller även de vanligaste fördelningsfunktionerna (framför allt normalfördelningen och t -fördelningen), samt stöd för enklare statistiska beräkningar. Du kan dock klara dig även med en mycket enkel räknare med i princip enbart de fyra räknesätten, men måste då vara beredd på att kunna använda de tabeller som finns längs bak i kurslitteraturen samt en hel del skriv- och bokföringsarbete. *Du förväntas själv ta reda på hur just din miniräknare fungerar och vilka funktioner den har.*

Lärare och kontaktytor

Examinator, kursansvarig och föreläsare	Michael Hörnquist
Lektioner och laborationer BI1	Vivianne Deniz
Lektioner och laborationer FTL1	Vivianne Deniz
Lektioner och laborationer SL1	Michael Hörnquist

Bägge lärarna träffas säkrast i samband med undervisningen. Enklare frågor av administrativ eller allmän karaktär kan även ställas till kursansvarig via e-post till michael.hornquist@liu.se. Matematiket är en resurs i samtliga matematikkurser, så även i denna.

Organisation och planering av undervisningen

Kursen är uppdelad i fyra moduler, var och en syftande till några av lärandemålen. Avsikten är att det skall bli enklare för dig att se vad du skall kunna efteråt.

¹Enligt uppgift kan det dröja upp till 24 timmar från registreringen. För dig som läser om kursen krävs att du skickar epost till kursansvarig och meddelar att du önskar tillgång till kursrummet, så läggs du till manuellt. Obs, detta gäller även om du omregistrerar dig

A Deskriptiv statistik och datorstöd

Denna modul svarar mot lärandemålen 1 och 10, och ger grunderna i deskriptiv (=beskrivande) statistik. Laborationen syftar primärt till att komma igång med Excel. Uppgifterna är hämtade från exempelsamlingen på Lisam (anges som x.y) samt ur kurslitteraturen Løvås (anges som Lx.y).

Fö A. Løvås: kap. 2.1 – 5 samt 2.8.

Le A. Före och under lektion: 1.1, 1.2, 2.1, 2.4, 2.5, 2.6

Efter lektion: 1.3, 1.4, L2.7, L2.8, L2.9, L2.15, L2.16, 2.2, 2.3, 2.8

Lab A. Finns på Lisam (inlämnas senast fredagen 26 januari 2018 kl. 07.00, alternativt/komplettering måndagen 12 februari 2018 kl. 07.00)

B Sannolikhetslära

Denna modul svarar mot lärandemålen 2 – 4. Det här är den mest teoretiska delen av kursen, och ger många av de verktyg som behövs för de mer tillämpade statistikdelarna.

Fö B1. Løvås: kap. 3.1 – 6

Le B1. Före och under: 3.1a, 3.1b, 3.2, 3.3, 3.4, 3.5. Efter: L3.21, L3.16, L3.26, L3.7

Fö B2. Løvås: kap. 4.1 – 3

Le B2. Före och under: 4.1, 4.2ab, 4.2cdef, 4.3. Efter: L4.1, L4.12, L4.14, 4.4, 4.5

Fö B3. Løvås: kap. 5.1 – 2, 5.5 – 8

Le B3. Före och under: 5.2a, 5.2bcde, 5.3, 5.4, 5.5. Efter: 5.1, L5.1, L5.18, L5.19, L5.26, L5.28

Le B4. Repetition av hela modulen.

Extra. Gästföreläsning av Valentin Polishchuk, främst för FTL men alla är välkomna

C Statistik

Denna modul svarar mot lärandemålen 5 – 7 och 10. Normalfördelningen och Centrala gränsvärdessatsen är väsentliga *förkunskaper* för den här modulen.

Fö C1. Løvås kap. 5.9.2, 6.1, 6.2.1 – 5, 6.3.1 – 3, 6.3.5 – 6

Le C1. Före och under: 6.1, 6.2, 6.3, 6.4a, 6.4bc. Efter: L6.10, L6.11, L6.14

Lab C. Finns på Lisam (inlämnas senast fredagen 23 februari 2018 kl. 07.00)

Fö C2. Løvås kap. 6.4 – 5

Le C2. Före och under: 6.5, 6.6, 6.7, 6.8a, 6.8b. Efter: L6.20, L6.23, 6.9

D Sambandsanalys

Denna modul svarar mot lärandemålen 8 – 10. Största delen tas upp av så kallad enkel linjär regression, men även korrelation ingår.

Fö D. Løvås kap. 7.1 – 3 (förutom 7.3.9)

Le D. Före och under: 7.1, 7.2, 7.4, 7.6. Efter: 7.3, L7.4, L7.5, 7.5

Lab D. Finns på Lisam (inlämnas senast fredagen 9 mars 2018 kl. 07.00)

Postludium

En lektion, sist i läsperioden, där inget nytt stoff presenteras. Det är era frågor som är stommen.

Om undervisningen, inklusive examinationen

Föreläsningar: Föreläsningarna genomförs på konventionellt sätt. De *introducerar* begrepp och ideer, men *förutsätter* att du själv studerar de relevanta avsnitten i kurslitteraturen. Om du har kikat på stoffet före föreläsningen kan jag nästan lova att din behållning kommer att öka avsevärt. Alla bilder som visas på föreläsningarna kan laddas ned från Lisam senast vid lunch dagen före. Föreläsningarna kommer *inte* att filmas, bland annat för att inte göra det svårare för den student som önskar avbryta med en fråga (vilket alltså uppmuntras).

Laborationer, LAB1: Det finns tre laborationer i kursen. Det är obligatorisk närvaro vid laborationspassen, och om du av någon orsak inte kan närvara vid ditt pass uppmanas du i första hand att bevista någon annan grupps pass.² Till varje laboration finns på Lisam de uppgifter du ska göra. **Några av dessa är förberedelseuppgifter, och dessa ska vara lösta innan du påbörjar laborationen.**³ Redovisning av laborationen sker genom att du, tillsammans med en till två andra studenter, skriver en redogörelse och lämnar in samt att du innan du lämnar datorsalen tillsammans med din grupp kortfattat berättar för läraren vad ni har gjort och vad som eventuellt återstår innan ni är klara. **Det är viktigt att ni också skriver upp namnen på gruppmedlemmarna på den lista som läraren tillhandahåller.** Beräkningarna görs i Excel och grupperna formerar ni själva. Det går bra att variera grupper mellan laborationerna. Alla arbeten lämnas in via Lisam. Om du av någon orsak inte skulle lämna in i tid, eller inte får redovisningen godkänd, erbjuds ytterligare fyra (4) möjligheter till inlämning under året (motvarande omtentamina).⁴ Det är kl. 07.00 den 19 mars, 9 april, 9 juni samt 3 september, 2018, som är sista tidpunkt. Notera att vi *inte* rättar inlämnade rapporter kontinuerligt, utan endast efter respektive sista inlämningsdag.⁵ Den student som inte är godkänd på alla tre laborationer efter sista inlämningstillfället, 3 september 2018, behöver normalt göra om hela momentet LAB1 nästa år. Samarbete mellan grupperna är tillåtet, men inte att skriva av varandra. Varje gruppmedlem ansvarar fullt ut för den lösning gruppen lämnar in, och läraren kan begära muntlig förklaring av vad som gjorts för att säkerställa enskilda students insats.⁶

Lektioner: Avsikten med lektionerna är att dina frågor som student skall stå i centrum. Den mindre gruppen (jämfört med föreläsningarna) uppmuntrar förhoppningsvis dialog och samtal. Vid sex av lektionpassen finns på undervisningsplanen en inramad uppgift, dessa är studentredovisningsuppgifter som kan ge bonuspoäng till tentamen (syns i kursplanen som UPG1). Till dessa lektioner skall du, om du önskar bonuspoängen, ha med dig en handskriven lösning till den inramade uppgiften, att lämna in till läraren, som lottar bland de inlämnade lösningarna vem som skall få redovisa vid tavlan. Den inlämnade lösningen får givetvis användas som stöd vid tavelredovisningen. Alla som har lämnat in får *en halv* (1/2) bonuspoäng, även den som inte redovisar vid tavlan, och bonusen gäller till och med omtentamen i augusti 2018. Totalt kan alltså upp till tre (3) bonuspoäng erhållas. Lösningen som redovisas behöver inte vara perfekt, men den skall vara ett ärligt försök, och du som redovisar måste visa att du har förstått uppgiften. Om det är uppenbart att du enbart har skrivit av någon annan, utan att själv sätta dig in i uppgiften, tappas samtliga bonuspoäng som du har samlat ihop under kursen så långt. De schemalagda tomma passen efter föreläsningarna är avsedda att ge tid både till att efterbearbeta föreläsningen och till att förbereda kommande lektion. De är inlagda så ofta övriga schemat har tillåtit det. Uppgifterna angivna som "efter" är lämpliga som repetition, och är i första hand tagna ur kurslitteraturen för att göra dem mer lämpade att studera på egen hand då de ansluter tätare till Løvås framställning.

²Det finns även två uppsamlingstillfällen, dels måndagen 5 mars 2017 kl. 15.15 – 17, dels ett tillfälle i omtentamensperioden i juni (exakt dag och tid meddelas senare), vilka kan användas av den som har missat ordinarie tillfälle, exempelvis på grund av sjukdom.

³Läraren kommer att kontrollera detta och avvisa från salen den som inte kan visa upp egna lösningar, eller åtminstone lösningsförsök, till uppgifterna.

⁴Fem (5) möjligheter för första laborationen, även 12 februari 2018 kl. 07.00.

⁵Motsvarande att ni inte får tentor rättade direkt för att kunna lämna in igen vid samma tillfälle.

⁶Att ha sitt namn med på en laborationsrapport där man inte kan förklara det som står är att försöka vilseleda examinator att prestationen är ens egen, utan att den är det. Det är inte acceptabelt, och sådant beteende kommer att anmälas till Disciplinnämnden som grundad misstanke om fusk.

Notera gärna språkbruket, *uppgifterna* är hämtade ur *exempelsamlingen*. De uppgifter som löses är alltid *exempel på något*, det kan vara en princip eller ett begrepp inom sannolikhetsläran, en sats i statistikteorin, eller något annat. Men var och en är värd att stanna upp inför och fundera över vad du kan lära av just det exemplet. Det är orsaken till att det är tämligen få uppgifter per modul, kvalitet skall ersätta kvantitet.

Tentamen, TEN1: Tentamen består av sex uppgifter à sex poäng. Till det kommer upp till tre bonuspoäng du kan ha med dig från lektionsdelen av kursen. För betyget n räcker $6n - 1$ poäng.

Uppgifterna på tentamen kopplar direkt till de fyra modulerna (modulen Sannolikhetslära och modulen Statistik kopplar till två uppgifter vardera) och därmed även till lärandemålen. Notera dock att ämnet är kumulativt i sin uppbyggnad, delarna bygger på varandra, varför senare avsnitt ofta bygger på sådant du har tillgodogjort dig i tidigare avsnitt.

Tillåtna hjälpmedel på tentamen är kurslitteraturen, dvs Løvås,⁷ samt valfri miniräknare.⁸ Det får finnas normala anteckningar i boken, sådana man ofta gör i en lärobok som studeras.⁹ Likaså är det ok med små (ett par centimeter) påklitrade sidmarkeringar, ”pagemarkeringar”.

Välkommen till kursen!
Michael Hörnquist

⁷Som *alternativ* kan valfri annan bok i ämnet användas, dock enbart *en* (1) bok totalt. Examinator tar emellertid inget ansvar för att någon annan bok än Løvås täcker det som behövs för att klara tentamen, exempelvis att den skall innehålla nödvändiga tabeller.

⁸Miniräknaren i mobiltelefonen är dock *inte* tillåten, då den ger möjlighet till kontakter med omvärlden.

⁹Ingen kommer att anmälas till Disciplinnämnden för att det finns för mycket skrivet i boken.