

Statistikens grunder  
(an, 7,5 hsp)

Tatjana Nahtman

Statistiska institutionen, SU

# KURSENS INNEHÅLL

Statistiken ger en empirisk grund för ekonomin.

I denna kurs betonas statistikens idémässiga bakgrund och dess tillämpning inom empiriska undersökningar, speciellt behandlas beskrivande statistik och statistisk slutledning.

Kursen ger en orientering om statistikens roll inom de ekonomiska vetenskaperna. Modellbegreppet diskuteras utförligt med speciell tonvikt på sannolikhetsmodeller och deras tillämpningar inom ekonometrin.

Kursens innehåll ger kunskaper som är av stor nytta vid studier i nationalekonomi men även vid tillämpningar av statistiska metoder inom en mängd andra områden.

## De begrepp som behandlas i kursen är:

- Kunskapsbyggnad. Modeller, speciellt sannolikhetsmodeller.
- Grundläggande sannolikhetslära.
- Diskreta och kontinuerliga stokastiska variabler och deras sannolikhetsfördelningar.
- Datainsamling.
- Beskrivande statistik i form av tabeller och diagram.
- Index.
- Samplingfördelningar och centrala gränsvärdesatsen.
- Punktskattning. Intervallskattning.
- Hypotesprövning.
- Anpassningstest och oberoendetest.
- Regression.
- Statistiska undersökningar.
- Beslutsteori.

# Lärandemål

Efter att ha genomgått kursen förväntas studenten kunna:

- kritiskt granska en statistisk undersökning
- formulera en modell för ett elementärt ekonomiskt problem
- lösa elementära problem enligt kursens innehåll
- lösa elementära problem vid punktskattning, konfidensintervall och hypotesprövning i ett antal elementära och vanligt förekommande fall
- genomföra enkla dataanalyser med hjälp av statistisk programvara samt presentera resultaten

# Schema

| Datum | Tid | Innehåll |
|-------|-----|----------|
|-------|-----|----------|

---

|       |       |                                                    |
|-------|-------|----------------------------------------------------|
| 21.01 | 15-18 | F1: Introduktion till kursen (F.A. 1)              |
| 22.01 | 13-15 | F2: Vetenskap (F.A.2)                              |
| 23.01 | 13-15 | F3: Modeller, kunskapsbyggnadsprocessen (F.A. 3,4) |
| 29.01 | 10-12 | F4: Repetition av matematik                        |
| 30.01 | 08-10 | F5: Beräkning av sannolikheter (F.A. 5)            |
| 31.01 | 10-12 | F6: Stokastiska variabler (F.A. 6)                 |
| 04.02 | 10-12 | F7: Diskreta stokastiska variabler (F.A. 7)        |
| 06.02 | 10-12 | F8: Konnuerliga stokastiska variabler (F.A. 8)     |
| 12.02 | 10-12 | F9: Databildning, Statistisk inferens (F.A. 9, 10) |
| 13.02 | 10-12 | F10: Deskription, en variabel (F.A.11)             |
| 14.02 | 13-15 | F11: Deskription, flera variabler (F.A.12)         |
| 19.02 | 10-12 | F12: Deskription, tidsserier (F.A. 13)             |
| 22.02 | 09-14 | Brunnsvikssalen: Skriftlig tentamen                |
| 18.03 | 09-14 | Brunnsvikssalen: Skriftlig omtentamen              |

OBS! F.A. står för Föreläsningsanteckningar.

## Vad är statistik?

Prediktion av framtiden har alltid varit önskad. Sierskor försökte att göra prognoser men lyckades inte alltid.

Det finns ingen statistisk metod som kan eliminera eller förklara alla osäkerheter i verkligheten, men statistik kan användas för att kvantifiera denna osäkerhet.

## Vad är statistik?

Ofta hör vi på nyheterna eller läser i tidningar uppgifter som bygger på statistik.

- Väljarundersökningar
- Arbetslöshetsundersökningar
- Konsumtionsundersökningar
- Medicinska forskningsresultat, etc.

Kan vi lita på alla dessa siffror?

Vilka slutsatser kan vi dra?

Hur säkra är vi?

# Tidsserier.

Tidsserier förekommer inom många områden:

- Ekonomiska data
  - Arbetslöshetssiffror
  - Försäljningsvärden
  - Konsumentprisindex och andra index
  - Export- och importmängder
- Miljödata
  - Halten av olika föroreningar i vattendrag
  - Ozonhalten i luften i större tätorter
- Samhällsdata
- Efterfrågan på sjukvård
- Födelsetalen i olika länder



## Vad är statistik?

Man har säkert gjort några statistiska påståenden i vardagslivet eller tänkandet.

Påståenden som

- *"I genomsnitt spenderar jag 100 kronor per månad på choklad"*
- *"Det är mera troligt att du blir godkänd på tentamen om du börjar förbereda materialet redan nu"*

är egentligen statistiska till sin natur.

## Vad är statistik?

Statistik är ett ämne som kan användas för att:

- planera olika försök och andra typer av datainsamling,
- sammanfatta information för att bättre förstå något fenomen (naturföreteelse),
- dra slutsatser från data,
- uppskatta nuet och prediktera framtid.

## Vad är statistik?

För att prediktera använder statistik ett annat ämne - sannolikhetssteori som modellerar slumpen matematiskt som möjliggör beräkningen av chansen(slumpen) i olika komplicerade situationer.

Statistik är viktigt verktyg som används inom olika områden: medicin, psykologi, utbildning, sociologi, ekonomi, ja m.m.

Statistik är vetenskapen som lär oss att sälla ut information ur data, d.v.s. fundera kring den informationen som finns i data.

Resultat från statistiska undersökningar utgör ofta en del av beslutsunderlag.

För att kunna besluta i en viktig fråga är det viktigt att man har förståelse för den informationen som en statistisk rapport ger.

## Vad är statistik?

I praktiken ger ett datorprogram våra beräkningar, men datorn gör inga tolkningar åt oss.

Det viktigaste är då inte att kunna stoppa in siffrorna på rätt ställe i formeln, utan att kritiskt kunna granska de uppgifter som presenteras för oss.

För att kunna göra det måste vi förstå de statistiska metoderna och lära oss ett statistiskt tänkesätt.

## Var det apelsinjuice?

Exempel. I augusti 1991 spred sig en epidemi av matförgiftning på inrikesflyg i Australien.

Det tydade på att apelsinjuice var som problem. Men var det verkligen den? Hur kan man vara säkert?

Antagg att nästan alla som drack apelsinjuice blev sjuka och majoriteten av passagerarna som inte drack juicen insjuknade inte. Bevisar det något?

Svaret är Nej.

## **Var det apelsinjuice?, forts.**

Viktingt: De passagerarna som drack apelsinjuice har kanske ätit bröd eller något annat.

Man behöver sortera alla typer av dryck och mat som passagerarna hade med eller fick på flygplan.

Fråga 1: bland de passagerarna som drack och åt samma har de som drack också apelsinjuice blivit sjukare än de som inte drack juicen?

Fråga 2: om det var apelsinjuice som orsakade förgiftningen varför inte alla som drack den har blivit sjuka?

## Solsken och hudcancer.

Det första övertygande beviset om att det finns sambandet mellan exponering av solsken och hudcancer gjordes av australisk statistiker Oliver Lancaster.

Han observerade att andel hudcancer i Australien bland kaukasier korrelerar starkt med latitud, d.v.s. med solsken: andelen är större i norra delar av riket än i södra delar.

Sådan observation blev möjligt med hjälp av noggrann datainsamling och statistisk analys av cancer kvot.

## Vad är statistik?

Ämnet statistik omfattar många områden. Ett av dessa områden handlar om hur vi kan samla in data.

Det är viktigt att planera sin undersökning nogga och tänka igenom hur datainsamlingen skall gå till för att vi ska ha en möjlighet att tolka resultaten.

Ett annat område handlar om hur vi sammanställer och analyserar våra insamlade data.

Hur vi presenterar data i tabeller och grafer, och hur vi INTE ska göra.

Vad ska vi tänka på när vi ser diagram som andra presenterar för oss?

Det är viktigt vi kritiskt kan granska gjorda undersökningar.



## Vad är statistik?

En del av denna kurs handlar om slump och sannolikheter.

Ett viktigt område inom ämnet är statistisk inferens, dvs. att dra slutsatser från data och generalisera dem till en större grupp än den undersökta.

Hur säkra kan vi vara på att resultatet vi fått i vårt stickprov går att generalisera till hela populationen?

Statistiska metoder används idag inom många områden och alla borde av åtminstone allmänbildande anledning bekanta sig med de statistiska metoderna och det statistiska tankesättet.

# Deskriptiv statistik

Deskriptiv (beskrivande) statistik handlar att organisera, summera och presentera sina data på ett överskådligt och informativt sätt.

För att göra detta används olika statistiska mått, tabeller och grafer.

Exempel på sådana är medelvärde, korstabeller och histogram.

## Statistisk inferens

Ofta görs statistiska undersökningar i form av ett urval från en större grupp (population).

Om populationen är stor blir det ofta alltför kostsamt och tar för lång tid att undersöka hela populationen (totalundersökning).

Ibland medför undersökningen att den undersökta delen förstörs. Om du till exempel tillverkar och säljer pizzor vill du att dina pizzor ska smaka gott och gör då kvalitetskontroll av produktionen. Men du kan ju inte gärna sälja de pizzor du just provsmakat. Det är då nödvändigt att begränsa sin undersökning till en delmängd av populationen.

## Statistisk inferens, forts.

Men vid urvalsundersökningar är man ju inte enbart intresserad av den information du får om stickprovet.

Det är av intresse att kunna generalisera till hela populationen.

(även de pizzor du inte provsmakat bör ju helst smaka gott.)

Metoder för att kunna dra slutsatser om en population utifrån den begränsade information du har att tillgå i form av stickprovdata behandlas under avsnitten för statistisk inferens.

## Vad är statistik?

De metoder som används för att sammanfatta samlade erfarenheter, göra förutsägelser, dra slutsatser och fatta beslut då informationen är ofullständig brukar kallas statistiska metoder.

Inom vetenskapen statistik studeras dessa metoder. Man försöker också konstruera nya metoder för nya situationer och problemställningar som uppkommer i det moderna samhället.

Inom den statistiska vetenskapen försöker man också vidareutveckla och förbättra gamla metoder så att t ex förutsägelser om framtiden kan göras med större precision.