

# Studiehandledning

Python programmering för statistisk dataanalys och maskininläring

45 Yhp

Uppdaterad 2022-08-21

## Syfte

Syftet med kursen är att den studerande ska få de kunskaper, färdigheter och kompetenser inom Python-programmering som krävs för att kunna arbeta med kvalificerad dataanalys. Den studerande kommer att få bekanta sig med bibliotek så som Numpy, Pandas och Matplotlib. De studerande kommer även att arbeta med grunderna i linjär algebra och statistik innefattande sannolikhetsteori och inferensteori samt implementera detta i Python.

## Innehåll

Kursen omfattar följande moment:

- Användning av Pythons inbyggda typer
- Organisera och strukturera kod enligt kodstandard med relevanta kommentarer
- Implementera klasser och objekt
- Versionshantering med Git
- Grunderna i sannolikhetsteori och statistisk inferensteori
- Grunderna i linjär algebra
- Paket för dataanalys såsom, NumPy, Pandas, Matplotlib
- Utföra grundläggande beräkningar och visualisera data
- Dokumentation av kod

## Mål/Läranderesultat

Målet med kursen är att den studerande ska behärska pythonprogrammering och de i branschen vanligt förekommande biblioteken. De studerande ska även behärska grunderna i sannolikhetsteori och statistisk inferensteori. De studerande ska ha även ha grundläggande kunskaper om linjär algebra i sin verktygslåda. Kunskaper och färdigheter från denna kurs är nödvändiga för att utföra dataanalys och implementera maskininlärningsmodeller.

Efter genomförd kurs ska den studerande kunna:

**Kunskaper:**

1. Redogöra för pythons inbyggda datatyper, kontrollstrukturer, funktioner och grundläggande objektorientering samt de i branschen vanligt förekommande biblioteken
2. Förklara grunderna i sannolikhetsteori såsom att förstå diskreta och kontinuerliga

fördelningar

3. Redogöra för grunderna i statistisk inferensteori såsom konfidentintervall och hypotesprövning
4. Redogöra för grunderna i linjär algebra såsom matriser och vektorer

#### Färdigheter:

5. Tillämpa grundläggande pythonprogrammering för att skriva program som utför beräkningar och visualisering av given data
6. Använda grundläggande sannolikhetsteori och statistisk inferensteori som verktyg för att analysera och dra slutsatser i olika beslutssituationer
7. Använda de i branschen vanliga biblioteken för dataanalys och maskininlärning på ett effektivt sätt

#### Kompetenser:

Inga kompetenser

#### Kunskapskontroll och betygsättning

Kursen kommer att genomföras med blended learning med inspelningar och aktiva lektioner. Under kursens gång erbjuds även s.k. Open Office Hours, där studenterna har ytterligare möjlighet att få hjälp av kursledarna genom att ställa frågor.

Examination kommer att ske genom:

- 2 inlämningsuppgifter (IG/G/VG)
- 1 grupprojeckt som redovisas muntligt och skriftligt (IG/G)

Läranderesultat	Inlämningsuppgift 1 (G)	Inlämningsuppgift 2 (G/VG)	Projekt + rapport (G)
1	X		
2		X	
3		X	
4		X	
5	X		X
6		X	X
7		x	X

#### För betyget Godkänd ska den studerande

- Ha nått samtliga läranderesultat för kursen

#### För betyget Väl Godkänd ska den studerande:

- Uppnått kraven för betyget Godkänd
- Skriftligen i en inlämningsuppgift redogöra för hur en har löst ett mer avancerat

programmeringsproblem avseende komplexitet och kodkvalitet.

**Ikke Godkânt ges till studerande som har fullföljt kursen men inte nått alla mål för kursen.**

## Kunskapskontroller

**Resultat och betyg registreras senast inom 10 arbetsdagar från genomförd kunskapskontroll.**

**Om man inte kan utföra examinationen på utsatt deadline (till exempel på grund av allvarlig sjukdom) eller behöver utföra en komplettering så är "andra" examinationstillfället senast en vecka efter att den studerande fått sin uppgift rättad. Du kan lämna in examinationen när du vill under denna period.**

**Det "tredje" (och sista) examinationstillfället är senast tre veckor efter att den studerande fått sin uppgift rättad vid senaste tillfället.**

### Kunskapskontroll 1

Den första examinationen består av två delar och har deadline 28 augusti kl: 23.59. Endast betyget G kan erhållas. Bedömning sker via "kamratbedömning" vilket innebär att facit kommer finnas uppladdat och ni skall i grupper av 2 personer (eller fler om ni föredrar) kolla igenom varandras lösningar och diskutera det ni tycker behöver diskuteras. När ni är klara skickar ni in uppgiften och i lösningen så skriver ni vilken kurskamrat som har "godkânt" er. Detta är bra träning inför framtida jobb. Det är OK att rätta sig själv utifrån givet facit om man av olika anledningar absolut måste det (t.ex. på grund av sjukdom). Alla hjälpmedel är tillåtna att använda (diskussion med kurskollegor, internet, böcker, ...) så länge ni förstår vad ni gjort och kan förklara koden ni skrivit. Försök att göra klart alla uppgifter innan ni kollar i facit då detta gynnar er inlärning.

### Kunskapskontroll 2

Den andra examinationen består av fyra inlämnings delar. För VG så måste man klara samtliga VG uppgifter på Del 1 – Del 3 i examinationen och 75% rätt på Del 4 och Del 5.

Del 1 – Del 3: Deadline 13 september kl: 23.59. Samma regler för G gäller här som i Kunskapskontroll 1, d.v.s. "kamratbedömning" skall göras. För de som gör VG uppgifterna så kommer undertecknad (Antonio) att betygsätta dem uppgifterna. Alla hjälpmedel är tillåtna att använda (diskussion med kurskollegor, internet, böcker, ...) så länge ni förstår vad ni gjort och kan förklara koden ni skrivit. Försök att göra klart alla uppgifter innan ni kollar i facit då detta gynnar er inlärning.

Del 4: Hemtentamen i sannolikhets teori, deadline 5 oktober. Man har 6 timmar på sig att slutföra examinationen och för G krävs det att man har minst 60% av totalt antal poäng och för VG krävs 75%. Hemtentamen skall göras helt individuellt och endast formelsamlingen samt miniräknare (eller göra beräkningarna via t.ex. Python) får användas som hjälpmedel. Lösningarna skannas / fotas och skickas in. Skriv tydligt och ta tydliga foton (t.ex. via skanningsverktyg på mobilen).

~~Del 5: Hemtentamen i statistisk slutledning (inferensteori), deadline 21 oktober. Man har 6 timmar på sig att slutföra examinationen och för G krävs det att man har minst 60% av totalt antal poäng och för VG krävs 75%. Hemtentamen skall göras helt individuellt och endast formelsamlingen samt miniräknare (eller göra beräkningarna via t.ex. Python) får användas som hjälpmedel. Lösningarna skannas / fotas och skickas in. Skriv tydligt och ta tydliga foton (t.ex. via skanningsverktyg på mobilen).~~ **Examinationen är struken.**

### Kunskapskontroll 3

Den tredje examinationen är en gruppuppgift som består av två delar och skall göras i grupper av 3-5 personer. Uppgiften är ”mindre” sett till omfattningen. Alla hjälpmedel är tillåtna att använda (diskussion med kurskollegor, internet, böcker, ...) så länge ni förstår vad ni gjort och kan förklara vad ni gjort.

Uppgiften presenteras muntligt. Endast betyget G kan erhållas och för det krävs det att man har bra helhetsförståelse över det ni presenterar.

Del 1: Berör simulering och statistiska fördelningar. Deadline: 30 september.

Del 2: Berör Statistisk Slutledning. Deadline: 20 oktober.

### Utbildare

Namn: Antonio Prgomet

E-post: antonio.prgomet@ecutbildning.se

Tfn: 0720 – 16 08 81

Tillgänglighet: Möjligheten att ställa frågor och diskutera med utbildaren sker på lektionstid och vid behov via mejl funktionaliteten på Omniway. Försök att nyttja lektionstiden framför mejl för att kontakta mig. Vid något akut så smsa vad det gäller så kan jag ringa upp dig eller besvara sms:et (du kan prova ringa men jag har i majoriteten av tiden ljudet avstängt så ser inte inkommande samtal).

### Gästföreläsare

Namn, titel, företag

### Kursmaterial

Typ av material	Kommentar
Python Tutorial: <a href="https://docs.python.org/3/tutorial/index.html">https://docs.python.org/3/tutorial/index.html</a> .	För att läsa enligt god studieteknik ( <i>Kolla igenom rubriker, skumma igenom texten och sedan djupläsning</i> ) så behöver jag personligen oftast ha dokument utskrivna. Boken finns tillgänglig som pdf under kurstillbehör i Omniway.
Matrix Algebra for Engineers, Chasnov: <a href="https://www.math.hkust.edu.hk/~machas/matrix-algebra-for-engineers.pdf">https://www.math.hkust.edu.hk/~machas/matrix-algebra-for-engineers.pdf</a>	Varje kapitel innehåller länk för videoföreläsningar.
Körner, S. & Wahlgren, L. (2015). Statistisk dataanalys. (5. uppl.). Lund: Studentlitteratur.	ISBN: 9789144108704 Kan köpas t.ex. här: <a href="https://www.adlibris.com/se/bok/statistisk-dataanalys-9789144108704">https://www.adlibris.com/se/bok/statistisk-dataanalys-9789144108704</a>  OBS: Införskaffas inför kursstart.
Körner, S. & Wahlgren, L. (2016). Tabeller och formler för statistiska beräkningar. (3. uppl.). Lund: Studentlitteratur.	ISBN: 9789144114545. Kan köpas t.ex. här: <a href="https://www.adlibris.com/se/bok/tabeller-och-formler-for-statistiska-berakningar-9789144114545">https://www.adlibris.com/se/bok/tabeller-och-formler-for-statistiska-berakningar-9789144114545</a>  OBS: Införskaffas inför kursstart.

## Kursplanering / Veckoplanering – Vad skall jag göra varje arbetsdag?

I detta avsnitt så framgår vad som skall göras under varje kursvecka. Det är viktigt att vara i fas då man annars lätt hamnar efter och riskerar att inte klara examinationerna.

## Upplägg på Föreläsningar / Lektioner

Föreläsningarna / Lektionerna inleds med en kort genomgång på sådant som ni har frågor på eller tyckte var utmanande. Ni kommer även presentera uppgifter ni jobbat med under veckan som vi sen diskuterar. Slutligen så kan ni jobba gruppvis eller individuellt med möjligheten att ställa frågor. Lektionerna är mellan 09.00 – 16.30 och försök att undvika distraktioner (t.ex. mobiltelefoner) så du kan jobba effektivt för bästa möjliga inlärningsresultat.

Generellt så är lektionerna upplagda enligt följande (vi är flexibla så vi anpassar oss för det vi tycker är bäst):

1. Lektionen startar 09.00 men först 09.15 startar föreläsning / frågestund så ni hinner tänka igenom t.ex. frågor ni har.
2. Lektionerna inleds med repetition av det som gjorts tidigare.
3. Ca 12.00 – 13.00 Lunch.
4. På fredagar så gör vi en genomgång på vad som skall göras nästa vecka.
5. 15.45 Gemensamt avslut.

Mellan dessa ”hållpunkter” så är det eget arbete eller samarbete med kurskamrater där ni kan ställa frågor till mig (Antonio). Försök att *innan* lektionerna skriva ned sådant det uppstått frågor om under veckans gång, annars glömmer du det lätt.

## Schema:

Tisdagar 09.00 – 16.30 i Malmö och via Teams.

Fredagar: 09.00– 16.30 i Helsingborg och via Teams.

## Vecka 1: Python Programmering.

Denna vecka skall vi lära oss grunderna i Python programmering. Under veckan kommer ni jobba med er första examination som är utformad som en inlämningsuppgift och består av två delar. Kursmaterial under veckan är:

- Inspelad Python kurs: <https://www.youtube.com/watch?v=M7bnYJyCx0Q>
- Läsmaterial: <https://docs.python.org/3/tutorial/>
- **Inlämningsuppgift 1 som består av två delar och filerna heter "assignment\_1\_part\_1" och "assignment\_1\_part\_2".**

Måndag (Introduktionsdag):

1. Du skall ha laddat ned Anaconda innan introduktionsdagen, se separat dokument "*inför\_start*".
2. Genomgång av studiehandledningen inklusive "*Kursplanering/Veckoplanering – Vad skall jag göra varje arbetsdag?*" för att få en överblick av kursen. Notera att redan dagen efter, på tisdag den 23 augusti så har ni arbetsuppgifter som skall utföras.
3. Innan vi avslutar dagen, kontrollera att alla har Anaconda fungerande på sin dator.
4. Kolla igenom utbildningsplattformen Omniway.

Tisdag - Onsdag

1. Kolla på inspelad föreläsning:  
<https://www.youtube.com/watch?v=M7bnYJyCx0Q&t=4478s> från 0:00 till 1:02:20.
2. Innan du kollar på videon om mängder ("Sets" på engelska) vid tidpunkt [39:38] så läs igenom texten och kolla videorna i följande länk:  
<https://www.matteboken.se/lektioner/matte-5/mangdlara> .
3. **Gör samtliga uppgifter i examination 1 del 1 "assignment\_1\_part\_1".**
4. Läs kapitel 1, 3, 4, 5, 7.1 i "*The Python Tutorial*": <https://docs.python.org/3/tutorial/> .
5. Om du hinner påbörja gärna Torsdagens/Fredagens arbete.
6. **Extrainsatt frågestund kl: 19.30 – 20.30 på tisdag för de som önskar, har du inga frågor så behöver du inte närvara.**

Torsdag - Fredag

1. Kolla på inspelad föreläsning:  
<https://www.youtube.com/watch?v=M7bnYJyCx0Q&t=4478s> från 1:02:20 till 1:14:59.
2. **Gör samtliga uppgifter i examination 1 del 2 "assignment\_1\_part\_2".**
3. Läs kapitel 8 och 9 (skumma igenom kap 9.1 och 9.2) i "*The Python Tutorial*".
4. **Inlämningsuppgift 1 bestående av "assignment\_1\_part\_1" och "assignment\_1\_part\_2" har deadline 28 augusti kl: 23.59.**

## Vecka 2: Linjär Algebra

Denna vecka är fokus på Linjär algebra där vi bland annat kommer lära oss om Matriser och Vektorer som är en central del i programmering och maskininläring. Vi kommer tillämpa det vi lärt oss nästa vecka genom NumPy där matriser och vektorer kallas för "arrays" som är ett lite mer generellt begrepp (2 dimensionell "array" = matris, 1 dimensionell "array" = vektor).

Kursmaterial under veckan är:

- Matrix Algebra for Engineers (Chasnov):  
<https://www.math.hkust.edu.hk/~machas/matrix-algebra-for-engineers.pdf> [Notera, boken lär ut allmän matristeori så oroa dig inte över att det står "for Engineers"].
- Inspelade föreläsningar:  
<https://www.youtube.com/watch?v=IZcyZHomFQc&list=PLkZjai-2Jcxlg-Z1roB0pUwFU-P58tvOx&index=3&t=16s> som är kopplade till boken "Matrix Algebra for Engineers (Chasnov)".

Rekommenderad arbetsgång: Skumma igenom varje kapitel i boken innan du kollar på videon. Efter att du kollat på videon, djup-läs det kapitel som videon gick igenom. Lös slutligen uppgifterna som skall göras enligt planeringen nedan.

Måndag:

1. Lektion 1-5 i boken.
2. Innan lektion 2: Kolla igenom följande video som förklarar summa notationen:  
<https://www.youtube.com/watch?v=KbvD6FIJGU>.
3. Efter lektion 2, kolla igenom följande video som bl.a. vid tidpunkten 1.10 visar en bra grafisk minnesregel på när man kan utföra matrismultiplikation:  
<https://www.youtube.com/watch?v=nSNebx6C5Vg>
4. Uppgifter:
  - Kap 1: Uppgift 1
  - Kap 2: Uppgift 1-3
  - Kap 3: Uppgift 1, 2 (bara första delen), 3 (bara första delen).
  - Quiz på sid.11, uppgift 1,2.
  - Kap 4: Inga uppgifter.
  - Kap 5: Uppgift 1.

Tisdag:

1. Lektion 6 i boken.
2. Lektion 7-9 i boken, inte lika viktigt som lektion 1-6.
3. Uppgifter:
  - Kap 6: Uppgift 1, 2 (uppgift 2 bygger på ett trick som är lätt när man väl gjort det en gång. Annars är det en svårare uppgift).
  - Quiz på sid.19, uppgift 1,3.

Onsdag:

1. Lektion 10-12
2. Uppgifter:
  - Kap 10: Uppgift 1.
  - Kap 11: Uppgift 1.
  - Kap 12: Uppgift 1.
  - Quiz på sid.41, uppgift 1-3.

Torsdag:

1. Lektion 13-15.
2. Uppgifter:

- Kap 13: Uppgift 1.
- Kap 14: Uppgift 1.
- Kap 15: Uppgift 1.
- Quiz på sid.51, uppgift 1-3.

Fredag:

1. Repetition av det som gjorts under veckan.



## Vecka 3: NumPy, Matplotlib, Pandas

Denna vecka är fokus på NumPy, Matplotlib och Pandas som är centrala paket för data science. Varje delområde kommer examineras och för de vars ambitionsnivå är VG finns det extra uppgifter. Examinationerna som skall göras denna vecka är Kunskapskontroll 2, Del 1 – Del 3.

Måndag:

1. Kolla igenom föreläsningarna kopplat till NumPy:  
<https://www.youtube.com/watch?v=M7bnYJyCx0Q> [1:14:59 till 1:17:30].
2. Läs länkarna nedan och experimentera med koden:
  - <https://numpy.org/doc/stable/user/quickstart.html>
  - <https://numpy.org/doc/stable/user/basics.broadcasting.html>
3. Gör inlämningsuppgift 2.1: NumPy.

Tisdag - Onsdag:

1. Kolla igenom föreläsningarna kopplat till Matplotlib:  
<https://www.youtube.com/watch?v=M7bnYJyCx0Q> [1:17:30 till 1:23:31].
2. Läs länkarna nedan och experimentera med koden:
  - <https://matplotlib.org/stable/tutorials/introductory/usage.html#sphx-glr-tutorials-introductory-usage-py>
  - <https://matplotlib.org/stable/tutorials/introductory/pyplot.html#sphx-glr-tutorials-introductory-pyplot-py>
3. Gör inlämningsuppgift 2.2: Matplotlib.

**Excel och GitHub är inget som examineras i kursen men det är väldigt bra att kunna.**

4. Kolla på Excel kursen som är tillgänglig här: <https://youtu.be/4UMLFC1SoHM>
  - Kapitel 1-6 bör alla kolla på.
  - Kapitel 7-9 är för den ambitiösa och som har tid.
  - **Som Data Scientist förväntas du kunna Excel, i princip alla använder det i någon form.**
5. GitHub.
  - Översiktsvideo av GitHub: [https://www.youtube.com/watch?v=Q5fb\\_tsadMQ](https://www.youtube.com/watch?v=Q5fb_tsadMQ).
  - Man kan använda Git lokalt på sin dator och sen "Pusha" sina filer / "Commits" online till GitHub. Följande video går igenom Git och GitHub:  
[https://www.youtube.com/watch?v=SWYqp7iY\\_Tc&t=1577s](https://www.youtube.com/watch?v=SWYqp7iY_Tc&t=1577s). Bra att åtminstone ha en grundläggande förståelse då många företag använder GitHub. Oroa dig inte om du inte förstår alla detaljer. I den inspelade Python kursen jag (Antonio) spelat in så laddade jag upp filerna i GitHub på följande länk:  
<https://github.com/AntonioPrgomet/Learn-Python-Programming-for-Beginners>. Det är trevligt då man bland annat kan följa ändrings-historiken och enkelt dela filerna med andra via nätet.

Torsdag:

1. Kolla igenom föreläsningarna kopplat till Pandas:  
<https://www.youtube.com/watch?v=M7bnYJyCx0Q> [1:23:31 till 1:29:25].
2. Läs länken nedan och experimentera med koden:  
[https://pandas.pydata.org/docs/user\\_guide/10min.html](https://pandas.pydata.org/docs/user_guide/10min.html)
3. Gör inlämningsuppgift 2.3: Pandas.

Fredag:

1. Repetition av det som gjorts under veckan.

## Vecka 4: Sannolikhetssteori

Under vecka 4-6 är fokus på sannolikhetssteori och examinationen sker via en inlämningsuppgift som är utformad som en hemtentamen. Uppgifterna kommer vara av samma karaktär och svårighetsgrad som övningsuppgifterna i boken av Körner & Wahlgren (2015).

Många kapitel avslutas med ett avsnitt som heter "Blandade Uppgifter" (se t.ex. sid 71), dessa ingår inte i kursen.

**Se dem två sista sidorna för att se vilka videor du skall kolla på inför varje kapitel under vecka 4-9.**

Måndag:

1. Läs och gör tillhörande uppgifter: 1.1 – 1.5.
  - Skippa exempel 2 på sid. 17 och skippa uppgift 108.

Tisdag:

1. Läs och gör tillhörande uppgifter: 1.5 - 1.6.
2. **Kamratbedömning av Del 1 – Del 3 i kunskapskontroll 2.**

Onsdag:

1. Läs och gör tillhörande uppgifter: 2.1 – 2.2.

Torsdag:

1. Läs och gör tillhörande uppgifter: 2.3 - 2.4.

Fredag:

1. Läs och gör tillhörande uppgifter: 2.5 – 2.6.

## Vecka 5: Sannolikhetssteori

Måndag:

1. Läs och gör tillhörande uppgifter: 3.1 – 3.3.

Tisdag:

1. Läs och gör tillhörande uppgifter: 3.4 – 3.5.

Onsdag:

1. Repetition av det som gjorts hitintills.

Torsdag:

1. 3.6 – 3.8.

Fredag:

1. Repetition av det som gjorts hitintills.

## Vecka 6: Sannolikhetssteori

Måndag:

1. Läs och gör tillhörande uppgifter: 4.1 – 4.6.

Tisdag:

1. Läs och gör tillhörande uppgifter: 4.1 – 4.6.
2. Om du hinner kolla gärna på avsnitt 5.1 – 5.2.

Onsdag:

1. Läs och gör tillhörande uppgifter: 5.1 – 5.2.

Torsdag:

1. Läs och gör tillhörande uppgifter: 5.3 – 5.4.
2. **Kolla igenom gruppuppgift del 1 som skall presenteras imorgon (på fredag).**

Fredag:

1. **Presentation av gruppuppgift del 1 (Kunskapskontroll 3 del 1).**
2. Repetition av det som gjorts hitintills.

## Vecka 7: Sannolikhets teori & Inferens teori

Denna vecka är det hemtentamen i sannolikhets teori. Därefter är fokus fokus på statistisk slutledning (Inferens teori) och examinationen sker via en inlämningsuppgift som är formad som en hemtentamen. Uppgifterna kommer vara av samma karaktär och svårighetsgrad som övningsuppgifterna i boken av Körner & Wahlgren (2015).

Många kapitel avslutas med ett avsnitt som heter "Blandade Uppgifter", dessa ingår inte i kursen.

**Se dem två sista sidorna för att se vilka videor du skall kolla på inför varje kapitel under vecka 4-9.**

Måndag:

1. Repetition inför hemtentamen.

Tisdag:

1. Repetition inför hemtentamen.

Onsdag:

- 1. Hemtentamen i Sannolikhets teori. Kunskapskontroll 2, del 4.**

Torsdag:

1. Läs och gör tillhörande uppgifter: 6.1 – 6.3.

Fredag:

1. Läs och gör tillhörande uppgifter: 6.4 – 6.6.
  - Gör uppgift 610 om du har tid, annars hoppa den.
- 2. Utdelning av gruppuppgift del 2 (Kunskapskontroll 3 del 2). Deadline 20 Oktober.**

## Vecka 8: Inferensteori

Måndag:

1. Läs och gör tillhörande uppgifter: 7.1 – 7.2.

Tisdag:

1. Repetition av det som gjorts hitintills.

Onsdag:

1. Läs och gör tillhörande uppgifter: 7.3 – 7.6.

Torsdag:

1. Läs och gör tillhörande uppgifter: 8.0.

Fredag:

1. Läs och gör tillhörande uppgifter: 8.1 – 8.2.

## Vecka 9: Inferensteori

Måndag:

1. Läs och gör tillhörande uppgifter: 8.3 – 8.5.

Tisdag:

1. Läs och gör tillhörande uppgifter: 8.6 – 8.7.

Onsdag:

1. Repetition.

Torsdag:

1. Deadline gruppuppgift del 2.

Fredag:

- ~~1. Hemtentamen i Statistisk Slutledning / Inferensteori, kunskapskontroll 2 del 5.~~  
**Examinationen är struken.**

## Videos Kopplat till v.4 – v.9

Kapitel hänvisningarna nedan refererar till boken av Körner & Wahlgren (2015).

### Kapitel 1-2

För kapitel 1-2, kolla följande spellista som en ”introduktion”:

[https://www.youtube.com/watch?v=pV3nZAsJxl0&list=PLvxOuBpazmsOGOursPoofaHyz\\_1NpxbhA&index=5](https://www.youtube.com/watch?v=pV3nZAsJxl0&list=PLvxOuBpazmsOGOursPoofaHyz_1NpxbhA&index=5)

3. Video 5: *What Does Independence Look Like on a Venn Diagram?* Kan du skippa.
4. En extra video om Kombinatorik (kap 1.6):  
<https://www.youtube.com/watch?v=XJnIdRXUi7A&t=79s>

### Kapitel 3:

För kapitel 3 kolla igenom följande spellista innehållande 13 videos:

<https://www.youtube.com/watch?v=oHcrna8Fkl8&list=PLvxOuBpazmsNIHP5cz37oOPZx0JKyNszN>

5. Video 3 som handlar om Bernoulli fördelningen är det boken kallar för ”Tvåpunktsfördelad variabel”.
6. På s.91 i boken så framgår det att man kan approximera binomialfördelningen med Poissonfördelningen, för detta se video 9 och 10, video 10 bevisar det och är lite ”överkurs” men det är ett vackert bevis så kolla gärna på det.
7. Video 6 handlar om ”Geometric Distribution” som boken kallar för ffg-fördelning.
8. Video 11 kopplat till ”Negative Binomial Distribution” ingår inte i kursen och kan släppas.
9. Video 12 kopplat till ”Multinomial Distribution”, är likt trinomial fördelningen som dyker upp först i kapitel 4.6 i boken. Förstår man den ena så förstår man den andra.

Hur man använder Tabell 1: Binomialfördelningen i formelsamlingen kan ses här:

<https://www.youtube.com/watch?v=gfDWDujLtfM&t=71s>

### Kapitel 4:

#### 4.4

- <https://www.youtube.com/watch?v=KDw3hC2YNFc>
- <https://www.youtube.com/watch?v=85Ilb-89sjk>

#### 4.5

- Härledning av räkneregler för varians:  
<https://www.youtube.com/watch?v=zdhkXWyy0K0> , i videon används den generella egenskapen av förväntans värdet att:  
 $E[aX + bY + c] = aE[X] + bE[Y] + c$ , dvs. förväntans värdet är en linjär operation.

#### 4.6

Multinomial fördelningen är väldigt lik trinomial fördelningen, förstår man den ena så förstår man den andra. I följande video förklaras multinomial fördelningen:

<https://www.youtube.com/watch?v=syVW7DgvUaY&t=382s>

### Kapitel 5:

Kolla först på video 1-7 i följande spellista om kontinuerliga sannolikhetsfördelningar:

[https://www.youtube.com/watch?v=OWSOhpS00\\_s&list=PLvxOuBpazmsPDZGwqhjhjE3KkLWnTD34R0](https://www.youtube.com/watch?v=OWSOhpS00_s&list=PLvxOuBpazmsPDZGwqhjhjE3KkLWnTD34R0)



- Vi är hittills vana vid diskreta fördelningar och vi vet t.ex. att  $\sum_i P(x_i) = 1$  eftersom summan av alla sannolikheter är 1. I det kontinuerliga fallet så använder man integraler istället för summor, motsvarande formel blir:  $\int_{-\infty}^{\infty} f(t) dt = 1$  där vi integrerar från minus oändligheten till oändligheten. I det diskreta hade vi en "Probability mass function"  $p(x_i)$  medan vi i det kontinuerliga fallet hade en "probability density function"  $f(t)$ .
- På liknande sätt vet vi att väntevärdet för diskreta fördelningar kan räknas ut enligt följande formel:  $E[X] = \sum_i x_i p(x_i)$  (d.v.s. utfall multiplicerat med sannolikhet). För kontinuerliga fördelningar får vi:  $E[X] = \int_{-\infty}^{\infty} t f(t) dt$ .  
Notera att vi kan använda t, x eller någon annan integrationsvariabel, det spelar ingen roll:  $\int_{-\infty}^{\infty} t f(t) dt = \int_{-\infty}^{\infty} x f(x) dx$ .
- Kontinuitetskorrektur (se sid. 137 i boken) går igenom i video 7 vid tiden: 5:42. Bra förklarat på något som kan upplevas förvirrande i början.

Hur man använder normalfördelningstabellen (tabell 3a i formelsamlingen):

[https://www.youtube.com/watch?v=p\\_KApjpyBHE](https://www.youtube.com/watch?v=p_KApjpyBHE).

För kapitel 5.3, se följande spellista: [https://www.youtube.com/watch?v=Zbw-YvELsaM&list=PLvxOuBpazmsP7UN00cNZX64N1o\\_8635ds](https://www.youtube.com/watch?v=Zbw-YvELsaM&list=PLvxOuBpazmsP7UN00cNZX64N1o_8635ds)

## Kapitel 6:

6.3: <https://www.youtube.com/watch?v=xJlwSkyeP0k>

## Kapitel 7

Se hela spellistan som är kopplad till kapitel 7:

<https://www.youtube.com/watch?v=27iSnzss2wM&list=PLvxOuBpazmsMdPBRxBTvwLv5Lhuk0tuXh>

## Kapitel 8:

Se hela spellistan som är kopplat till kapitel 8:

<https://www.youtube.com/watch?v=tTeMYuS87oU&list=PLvxOuBpazmsNo893xlpXNfMzVpRBjDH67>