
Oblig 2a - Grunnleggende sannsynlighet og fordelinger

Levering: 1 PDF, i rett mappe på Canvas. Lever eventuell **R**/**MatLab**/**Wolfram**-kode som kildefil *i tillegg*.

Førstefrist: 8. oktober, 18:00

Sistefrist: 15. oktober, 18:00

Fristene er absolutte. Grupper på 3 eller færre må ha godkjent 60%.

Grupper med 4 deltagere må ha godkjent 75%.

Grupper med 5 deltagere må ha godkjent 85%.

Større grupper: Kontakt faglærer for å få tildelt tilleggsoppgaver. / Se Canvas

Samarbeidende grupper vil bli regnet som én gruppe.

Fjorårets studenter anbefaler: *Spar tid, bruk formelheftet aktivt under arbeidet med oppgavene.*

1. Bokoppgaver fra kapittel 7 (30%)

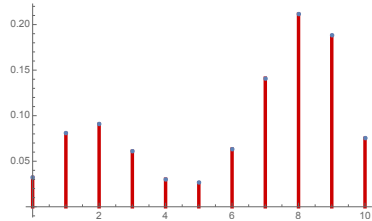
- (a) 2b
- (b) 5d

2. Grafisk vurdering av sannsynlighetsfordelinger (30%)

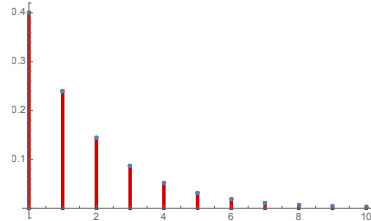
- (a) Eksamen 2020.12.11, oppgave 1.c.ii
- (b) Under vil dere finne 3×7 grafer. Først 7 sannsynlighetsfordelinger $f(x)$, eller *pdf*-er. Deretter 7 kumulative sannsynlighetsfordelinger $F(x)$, eller *CDF*-er. Så til slutt 7 inverst kumulative sannsynlighetsfordelinger $F^{-1}(p)$, eller *iCDF*-er. Dere skal matche dem sammen, i 7 grupper på 3 og 3. Gi gruppen navn etter pdf-ens bokstav.
- (c) Finn *modus*, X_{max} , for alle de gruppene hvor dette er mulig. Begrunn, der dette ikke er mulig. Hvilken av de 3 grafene er lettest å bruke for å finne modus?
- (d) Finn *medianen* \tilde{X} for alle gruppene, og markér x -verdien. Hvilken av de 3 grafene er lettest å bruke for å finne median?
- (e) Gjør et anslag på *forventningsverdien* $E[X] = \mu_X$ for alle gruppene, og markér x -verdien. Hvordan tenkte du for å finne $E[X]$?
- (f) Finn $P(X \leq 6)$ for alle gruppene. Hvilken av de 3 grafene er lettest å bruke for dette?
- (g) Finn 5. og 95. prosentil for hver av de gruppene, og markér det på riktig måte på pdf, CDF og på iCDF. Dette er et symmetrisk 90% intervallestimat, og blir regnet som en avansert øvelse. Dere klarte det nok ganske lett. Hvordan føles det?

pdf-ene

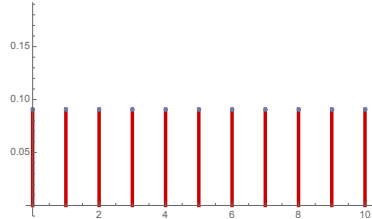
a)



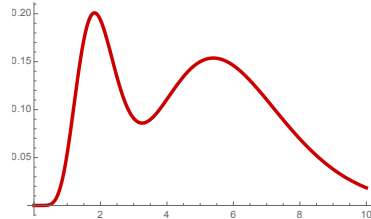
b)



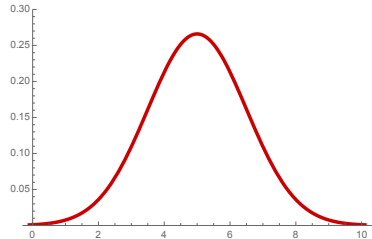
c)



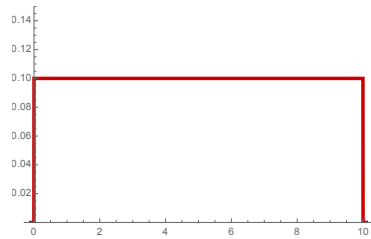
d)



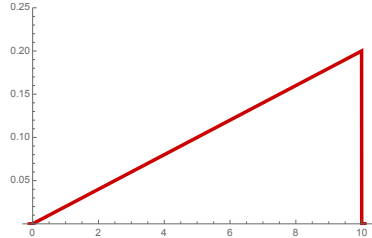
e)



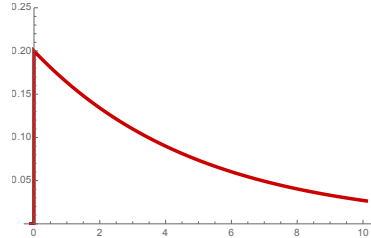
f)



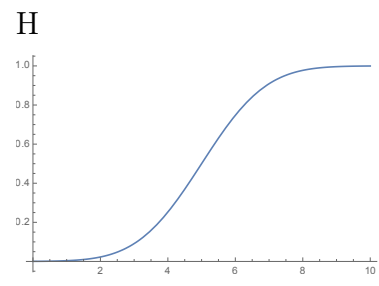
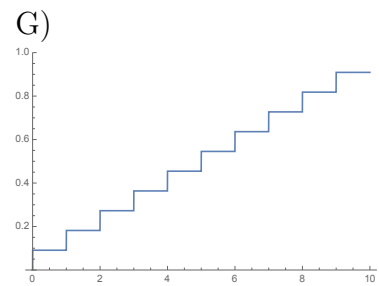
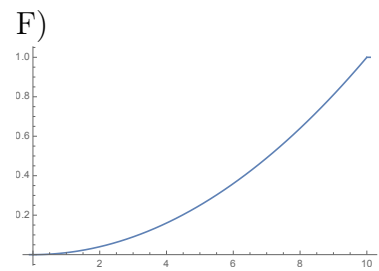
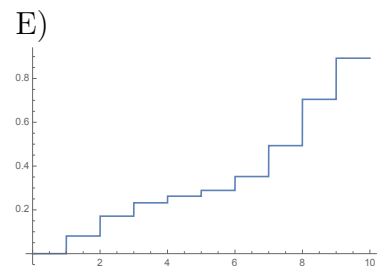
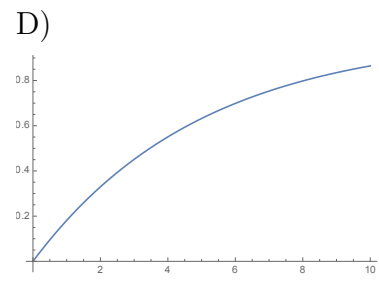
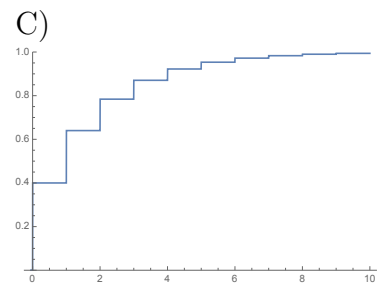
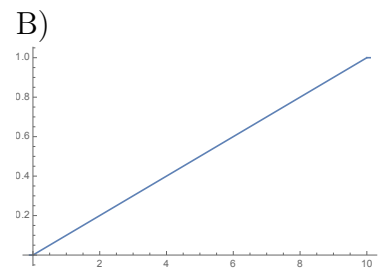
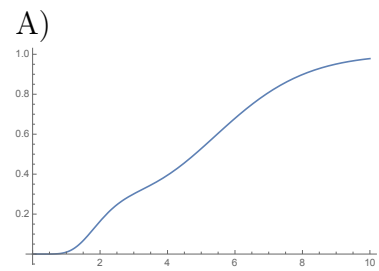
g)



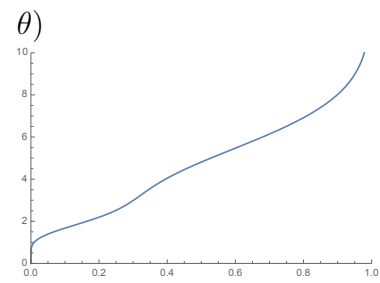
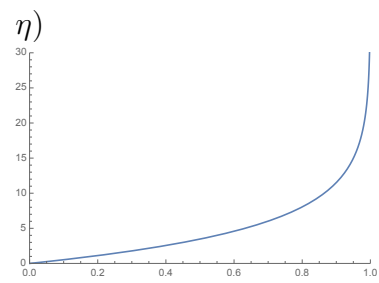
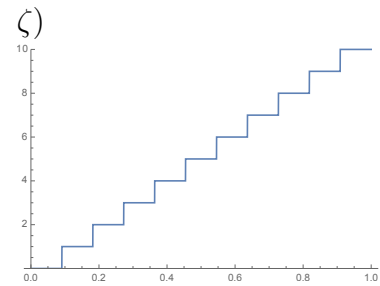
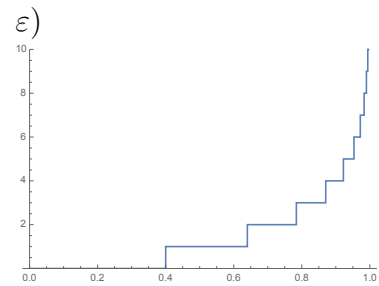
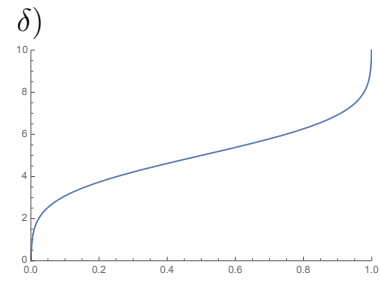
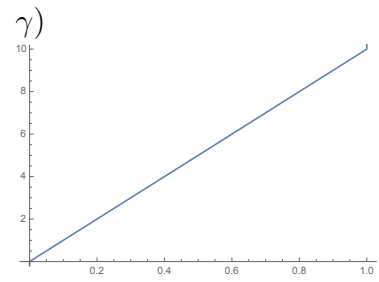
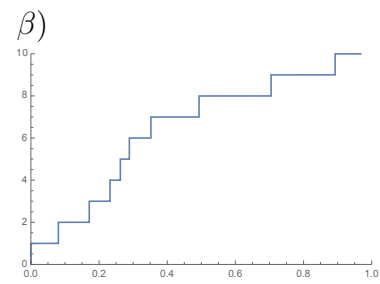
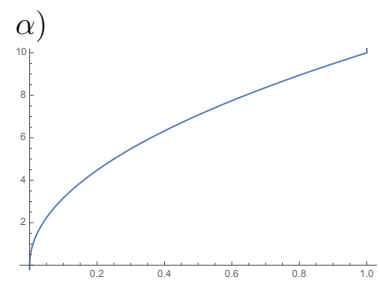
h)



CDF-ene



CDF-ene



3. Basics for diskrete sannsynlighetsfordelinger: (16%)

- (a) Kan du bruke $P(X \leq a)$ for å finne $P(X > a)$? Hvordan?
- (b) Hvorfor kan vi ikke regne med at $P(X < c) = P(X \leq c)$ når X er en diskret sannsynlighetsfordeling?
(Hvorfor vil de for det meste være forskjellige?)
- (c) Hvorfor er $P(X \geq 5) = 1 - P(X \leq 4)$ for diskrete X , og ikke $P(X \geq 5) = 1 - P(X \leq 5)$?
- (d) Hva er "pdf" for en *diskret* sannsynlighetsfordeling?
- (e) Hvorfor kan verdiene for pdf-en til en *diskret* sannsynlighetsfordeling aldri bli større enn 1?
- (f) Hva er "CDF" for en *diskret* sannsynlighetsfordeling?
- (g) Hva er største og minste mulige verdi for en diskret "CDF(x)"?

4. Basics for kontinuerlige sannsynlighetsfordelinger: (24%)

- (a) Kan du bruke $P(X \geq a)$ for å finne $P(X < a)$? Hvordan?
- (b) Hvorfor er alltid $P(X = c) = 0$ når X er kontinuerlig?
- (c) For diskrete sannsynlighetsfordelinger er $P(X \geq 5) = 1 - P(X \leq 4)$, men for kontinuerlige sannsynlighetsfordelinger er $P(X \geq 5) = 1 - P(X \leq 5)$. Hvorfor har vi disse to forskjellige reglene for diskrete og for kontinuerlige sannsynlighetsfordelinger?
- (d) Når X er kontinuerlig: I uttrykket $P(a \leq X \leq b)$ kan \leq og $<$ byttes med hverandre uten at du får forskjellige verdier. Hvorfor det?
- (e) Hva er "pdf" for en kontinuerlig sannsynlighetsfordeling? Hvordan skiller den seg fra "pdf" for en diskret sannsynlighetsfordeling?
- (f) Hvorfor kan verdiene for en pdf bli større enn 1 for kontinuerlige sannsynlighetsfordelinger når de ikke kan det for diskrete sannsynlighetsfordelinger?
- (g) Hva er "CDF" for en kontinuerlig sannsynlighetsfordeling?
- (h) Hva er største og minste mulige verdi for en kontinuerlig "CDF(x)"?
- (i) Hva er "iCDF" for en kontinuerlig sannsynlighetsfordeling?
- (j) Hvilke begrensninger finnes det på en iCDF?