# Eksempler på eksamensoppgaver

Pernille Kjeilen Fauskanger

December 1, 2021

## Oppgave 1

La $A$, $B$, $C$ og $D$ være hendelser definert i utfallsrommet $\mathcal{S}$, der

- $(A \cup B \cup C) \cap D = \emptyset$ og $D \neq \emptyset$
- $C \subset (A \cup B)$ og $C \cap A \cap B = A \cap B$
- $A \cap B \neq \emptyset$, $A \cap C \neq \emptyset$ og $B \cap C \neq \emptyset$
- $P(C|A) = P(C|B)$

**a)** Tegn hendelsene $A$, $B$, $C$ og $D$ inn i et venndiagram. Forklar hvorfor $D$ og $E = A \cup B \cup C$ er disjunkte hendelser.

**b)** Vis at $P(C) = P(A \cup B)P(C|A) - P(A \cap B)P(C'|A)$.

**c)** Vis at $C$ er avhengig av $A$ og $B$.

**Bonus (Bonus = ikke gjennomgått på kurset)**

**d)** La $P(C|A) = P(C|B) = 0.5$ og $P(A) = 0.15$ og $P(B) = 0.30$. Bruk dette til å regne ut $P(C \cup D)$.

## Oppgave 2

La $X \sim \text{Bin}(30, 0.60)$ og $f(X) = X^2$.

**a)** Regn ut sannsynlighetene $P(X - 1 = 30)$ og $P(29 \geq X)$.

**b)** Beregn forventning og varians til $X$ og $f(X)$. Du kan bruke at $E[X^4] = 1.19 \cdot 10^5$.

**Bonus**

**c)** Beregn sannsynligheten for at $X + \sqrt{f(X)}$ ligger mindre enn ett standardavvik unna dens forventning. (Hint: Lag en grovskisse av sannsynlighetsfunksjonen til $X + \sqrt{f(X)}$ og marker hendelsen langs x-aksen.)

**d)** Forklar hvordan vi kunne gått frem for å beregne en tilnærmet sannsynligheten til $\mathrm{P}(f(X) \leq 300)$. Hvordan ville du endret $p$ for å gjøre denne tilnærmingen mest mulig presis?

## Oppgave 3

På en gjennomsnittlig dag og på et helt tilfeldig sykehus som vi kaller for Haukeland blir det innlagt 20 pasienter med COVID-19. Vi antar at antall pasienter innlagt i løpet av $T$ døgn følger en poisson-fordeling med forventning $20T$.

**a)** Hvor mange pasienter kan vi forvente blir innlagt fra klokken 00:00 til 12:00 på en tilfeldig dag? Hva med antall som blir innlagt i løpet av resten av den samme dagen? Hva er sannsynligheten for at det blir innlagt færre pasienter enn forventet i begge disse tidsrommene i løpet av hele døgnet?

**Etter vi er ferdig med kjente fordelinger**

**b)** Vis at summen av de to første ventetidene (tid til første pasient og tid mellom første og andre pasient) er gammafordelt.

## Oppgave 4

En utvalgt person blant en populasjon der alle har hjerteinfarkt ser vi på konsentrasjonen til et spesifikt mineral i kroppen. La oss kalle mineralkonsentrasjonen hos en tilfeldig valgt person fra denne populasjonen for $K$. I følge en lege på et helt tilfeldig sykehus i Bergen, så følger $K$ en gammafordeling med $\alpha = 1$ og $\beta = 0.4$. Sannsynlighetsfunksjonen til en gammafordelt variabel for vilkårlige $\alpha$ og $\beta$ er gitt ved

$$f_K(k) = \frac{1}{\beta^\alpha \Gamma(\alpha)} k^{\alpha-1} \exp(-k\beta^{-1}) \quad k > 0. \tag{1}$$

Vi får oppgitt av the $m$'te momentet til $K$ er gitt ved

$$\mathrm{E}[K^m] = \beta^m \cdot \frac{\Gamma(\alpha + m)}{\Gamma(\alpha)}. \tag{2}$$

**a)** Tegn en grovskisse av fordelingen ved å bruke det du kan om gamma-fordelingen. (Hint: se på verdien til form-parameteren $\alpha$)

**b)** Hvis du har gjort oppgave a) riktig, så kan du se at fordelingen likner på en eksponentiell fordeling. Bruk verdiene til $\alpha$, $\beta$ sammen med sannsynlighetsfordelingen til $K$ for å vise at dette faktisk er en eksponentiell fordeling med intensitet $\lambda = \frac{1}{\beta}$. Betyr dette at legen hadde feil?

**c)** Regn ut forventning og varians til $K$. Pasienter med $K < 0.05$ blir diagnosert med sykdommen XD. Beregn sannsynligheten for at en tilfeldig valgt person blant populasjonen ikke blir diagnosert med XD.

**Etter vi er ferdig med punktestimatorer**

La oss anta at $\frac{1}{\beta}$ er en ukjent parameter som vi vil estimere fra data ($\alpha$ er forsatt 1). Vi samler inn $n$ iid. observasjoner av $K$: $\{K_i\}_{i=1}^n$.

**d)** Finn MLE-estimatoren til $\frac{1}{\beta}$, og undersøk om MLE-estimatoren er forventnigsrett.

## Oppgave 5

La $X_i$ være verdien til en aksje på dag $i$. en helt tilfeldig person mener at for å undersøke om den forventede verdien til aksjen har økt i løpet av $n + 1 = 21$ dager kan vi se på de $n = 20$ dagvise differansene

$$D_i = X_{i+1} - X_i \quad i \in \{1, 2, \ldots, 19, 20\}. \tag{3}$$

Vi antar at verdien til aksjen på dag $i + 1$ er avhengig av verdien til aksjen på dag $i$, men ikke avhengig av eventuelle tidligere dager. Verdien på aksjen på dag 1 er $a$. Merk: selvom aksjeverdiene for påfølgende dager er avhengige, så kan vi forsatt si at de dagvise differansene er normalfordelte og uavhengige.

**a)** Anta at $X_{i+1} \sim N(X_i + a + \mu_D, \sigma^2)\ \forall\ i$. Vis at variansen til $D_i$ er mindre enn $2\sigma^2$ hvis $X_{i+1}$ og $X_i$ positivt korrelerte. Forklar hvorfor det er vanskelig å finne forventningen her.

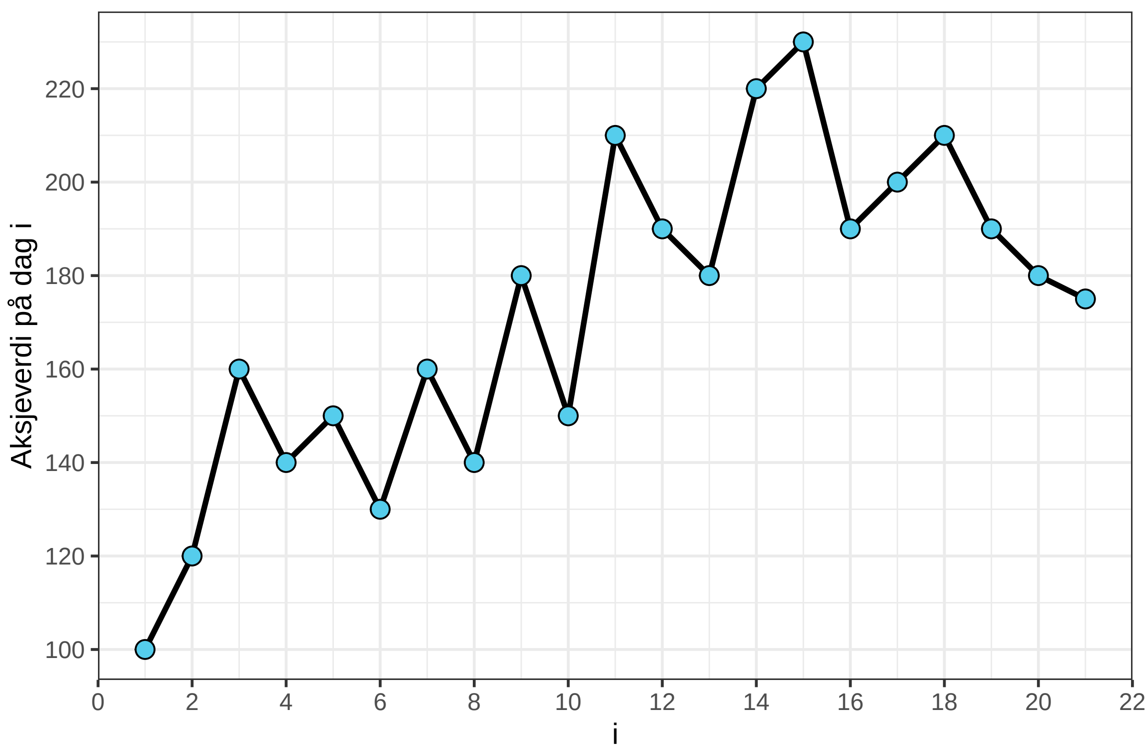
**b)** Vi henter ut observasjoner som er iid. $\{d_1, d_2, \ldots, d_{20}\}$. Du får vite at $\overline{d} = 3.75$ og $S_D = 27.67$. Konstruer et 95% kondifensintervall for $\mu_D$. Undersøk om 0 inngår i konfidensintervallet.

**Etter vi er ferdige med hypotesetesting**

**c)** Formuler to hypoteser som kan benyttes for å undersøke om forventet aksjeverdi har økt i løpet av disse $n = 21$ dagene ved å se på $\mu_D$.

**d)** Gjennomfør en hypotesetest med signifikansnivå $\alpha = 0.05$ for å undersøke om vi har nok empiri for å forkaste $H_0$ formulert i c). Hva er konklusjonen til testen? Kan 0 være i konfidensintervallet konstruert i b) samtidig som vi forkaster $H_0$ her?.

**e)** Finn den største verdien av $S_D$ som resulterer i forkastning av $H_0$. I figuren under kan du se kurven som viser aksjeverdien i løpet av disse 21 dagene. Er du enig med test-konklusjonen i d)?




**f)** Fra figuren kan vi regne ut at $\overline{x} = 171.67$ og $S_X = 34.25$. Vi formulerer hypotesene

$$H_0 : \mu_X \leq a \tag{4}$$
$$H_1 : \mu_X > a \tag{5}$$

Hva er $a$ i dette tilfellet? Gjennomfør en t-test og konkluder ved å bruke p-verdi om vi kan forkaste $h_0$ gitt valgt $\alpha = 0.05$ og $\{x_i\}_{i=1}^{n+1}$. Hvorfor kan vi ikke stole på resultatet fra denne testen?

**Bonus - for de spesielt interesserte**

**g)** En kan vise at en $t$-fordelt variabel med $\nu$ frihetsgrader er definert ved

$$T = Z \cdot \sqrt{\frac{\nu}{V}} \sim t_\nu. \tag{6}$$

Her er $Z \sim N(0,1)$ og $V \sim \chi^2_\nu$. Bruk denne sammenhengen til å vise at

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}. \tag{7}$$

Her er $\{X_i\}_{i=1}^n$ iid. der $X_i \sim N(\mu, \sigma^2)$.