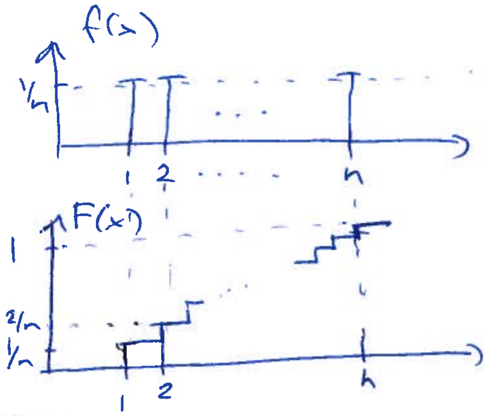
SUGGESTED SOLUTION TO:

TMA 4300 EXAM, 2013.

1

a)



$$f(x) = \frac{1}{n}, \quad x = 1, \dots, n$$

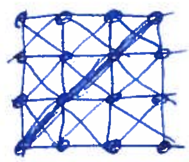
$$F(x) = \sum_{y=1}^{x} \frac{1}{n} = \frac{x}{n}$$

Trekk $U \sim U(0,1)$

sett $\underline{X = F^{-1}(U)}$, dvs $\underline{\text{inversjon}}$.

Dvs
$X = 1$ hvis $U \in (0, 1/n)$
$X = 2$ hvis $U \in (1/n, 2/n)$
$\vdots$
$X = n$ hvis $U \in (n-1/n, 1)$

b)



$$P(T = (m-1)\sqrt{2}) = P(\text{robot diagonal vei})$$

$$= 1/3 \cdot \underbrace{1/8 \cdot 1/8 \cdots 1/8}_{m-2 \text{ ledd}}$$

$$= \underline{1/3 \cdot (1/8)^{m-2}}$$

$\underline{\text{Pseudokoden}}$ baserer seg på å trekke en uniform retning i hver node, inntil den kommer til (m,m).
På hver iterasjon adderes distansen.

$T = 0$, $pos = (pos_1, pos_2) = (1,1)$
while $(pos \neq (m,m))$
  # Trekk tilfeldig retning, $X = F^{-1}(u)$, der $F(x) = \frac{x}{8}$, $x = 1, \dots, 8$
  (kantnoder får kun 3 eller 5 retninger)
  if (retninger horisontal eller vertikal)
    $T = T + 1$; $pos = posnew$; 4 options: $pos = (pos_1, pos_2 - 1)$, $pos = (pos_1, pos_2 + 1)$, $pos = (pos_1 - 1, pos_2)$, $pos = (pos_1 + 1, pos_2)$
  else # diagonal retning
    $T = T + \sqrt{2}$; $pos = posnew$; 4 options: $pos = (pos_1 - 1, pos_2 - 1)$, $pos = (pos_1 - 1, pos_2 + 1)$, $pos = (pos_1 + 1, pos_2 - 1)$, $pos = (pos_1 + 1, pos_2 + 1)$
end

i)

① c) $\hat{p} = \dfrac{\text{Antall run med } T^b > 100}{10000} \approx \dfrac{(10000-9000)}{10\,0000} = \underline{0.1}$

{ Ca 1000 av MC kjøringene har distanse lengre enn 100. $\hat{p} = \dfrac{\text{gunstige}}{\text{mulige}} \approx \dfrac{1000}{10000} = 0{,}1$ }

$\underline{Y = \hat{p} \cdot B \sim Bin(p, B)}$ siden hver kjøring enten blir lengre/kortere enn 100, kjøringene er uavhengige, og alle har samme sannsynlighetsmodell, med lik p.

$$Var(Y) = B \cdot p \cdot (1-p)$$

$$Var(\hat{p}) = Var\left(\frac{Y}{B}\right) = \frac{p(1-p)}{B} \approx \frac{\hat{p}(1-\hat{p})}{B}$$

$$\approx \frac{0.1 \cdot 0.9}{10000}$$

$$= \underline{(0.003)^2}$$

Asymptotisk er

$$\frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{B}}} \approx N(0,1)$$

Et 90 % konfidensintervall er

$$p \in \left(\hat{p} \pm 1.645 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{B}} \right)$$

Lengden er $2 \cdot 1.645 \cdot \sqrt{\dfrac{\hat{p}(1-\hat{p})}{B}} = 0{,}001$

Da kreves $B = \dfrac{2^2 \cdot 1.645^2 \cdot \hat{p} \cdot (1-\hat{p})}{(0.001)^2}$

$$= \frac{4 \cdot 1.645^2 \cdot 0{,}1 \cdot 0{,}9}{(0.001)^2} \approx \cancel{392\,200}$$

$\underline{974.170}$

::)

1 c) Forts.

Fra Figur 2 ser vi at $P(T>300)$ er veldig liten. Slike 'rare events' gir ustabil estimering. Spesielt er Coefficient of variation (CV)

$$CV^2 = \left(\frac{Var(\hat{p})}{E(\hat{p})^2}\right) \approx \frac{p(1-p)}{p^2 B} = \frac{1-p}{p \cdot B}.$$

Siden p nær 0, må B bli veldig stor for å relativt varians (sammenlignet med $E(\hat{p})^2$) skal bli liten.

Ide: Bruke importance sampling til å fremskaffe flere $T<300$.

I utgangspunktet er fordelingen til T basert på uniforme retninger, $1/8$.

I forslagsfordelingen, $q()$, kan vi gi mindre sannsynlighet for å gå opover, mot høyre.

For eksempel

$q(\text{posnew} = \text{pos}(1,1)) = \frac{1}{2}$.

$q(\text{posnew} \neq \text{pos}(1,1)) = \frac{1}{2}/7$, for alle 7 andre retninger.

De konstruerte $T^1, \ldots, T^B$ vil oftere bli større enn 300. Importance sampling kan kompensere for å ha brukt $q()$ istedet for $\pi = 1/8$ for alle 8 retninger.

$$\hat{I}_{IS} = \sum_{b=1}^{B} I(T^b > 300) \cdot \frac{\prod_{i \text{ steg}} \pi(\cdot)}{\prod_{i \text{ steg}} q(\cdot)},$$

HER BLE ULIKE KREATIVE FORSLAG VURDERT OG BELØNNET!

iii)

2

a) $p(\mu|y_1,\dots y_n,q) \propto p(\mu)\cdot p(q)\cdot \prod_{i=1}^{n} p(y_i|\mu,q)$

$$\propto p(\mu)\cdot \prod_{i=1}^{n} p(y_i|\mu,q)$$

$$\propto \exp\left(-\frac{r}{2}(\mu-\eta)^2 - \frac{q}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$$

$$\propto \exp\left(-\frac{(r+n\cdot q)}{2}\cdot\mu^2 + \left(r\cdot\eta + q\sum_{i=1}^{n} y_i\right)\cdot\mu\right)$$

Dette har kvadratisk funksjon i eksponenten, og må være en normalfordeling. Normalfordelingen har $X\sim N(\mu, 1/r)$

$$p(x) = \frac{1}{\sqrt{2\pi}}\cdot\sqrt{r}\cdot \exp\left(-\frac{1}{2}\cdot r\cdot(x-\mu)^2\right)$$

$$\propto \exp\left(-\frac{r}{2}\cdot x^2 + r\cdot\mu\cdot x\right)$$

Vi gjenkjenner dermed at $E(\mu|y_1,\dots y_n,q) = \frac{r\cdot\eta + (\sum_{i=1}^{n} y_i)q}{r+n\cdot q}$

$$Var(\mu|y_1,\dots,y_n,q) = \frac{1}{r+n\cdot q}$$

$$p(q|y_1,\dots y_n,\mu) \propto p(q)\cdot p(\mu)\cdot \prod_{i=1}^{n} p(y_i|\mu,q)$$

$$\propto p(q)\prod_{i=1}^{n} p(y_i|\mu,q)$$

$$\propto q^{\alpha-1} e^{-\beta q}\cdot \prod_{i=1}^{n}\sqrt{q}\cdot e^{-\frac{q}{2}(y_i-\mu)^2}$$

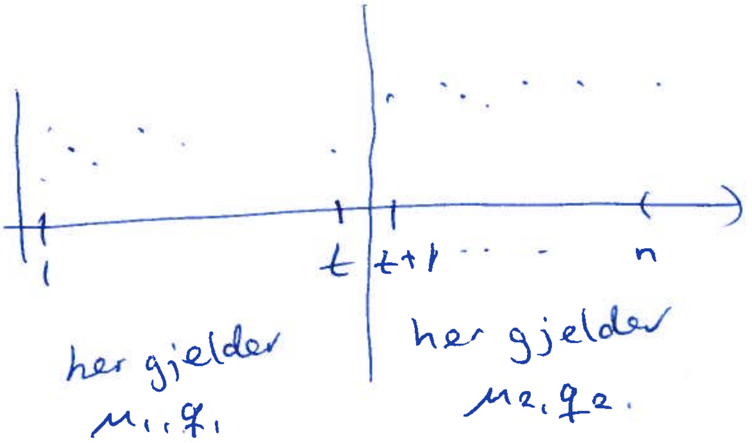
$$= q^{\alpha-1} e^{-\beta q}\cdot q^{n/2}\cdot e^{-\frac{q}{2}\sum_{i=1}^{n}(y_i-\mu)^2}$$

$$= q^{\alpha+n/2-1} e^{-q\left(\beta+\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right)}$$

Vi gjenkjenner formen til gamma-fordelingen. Nye parametre, gitt alt annet, dvs fulle betinget er $\alpha + n/2$ og $\beta + 1/2\sum_{i=1}^{n}(y_i-\mu)^2$.

iv)

b)



$$p(\mu_1 \mid \mu_2, q_1, q_2, y_1, \dots, y_n) \propto p(\mu_1)\cdot p(\mu_2)\cdot p(q_1) p(q_2) \cdot \prod_{i=1}^{t} p(y_i \mid q_1, \mu_1) \cdot \prod_{i=t+1}^{n} p(y_i \mid q_2, \mu_2)$$

$$\propto p(\mu_1)\cdot \prod_{i=1}^{t} p(y_i \mid q_1, \mu_1).$$

Dette er som før med t istedet for n, og $y_1, \dots y_t$ istedet for $y_1, \dots, y_n$.

$$p(\mu_1 \mid \mu_2, q_1, q_2, y_1, \dots, y_n) = p(\mu_1 \mid q_1, y_1, \dots y_t) \sim N\left(\frac{r\cdot\eta + q\sum_{i=1}^{t} y_i}{r + t\cdot q}, \frac{1}{r + t\cdot q}\right)$$

Tilsvarende

$$p(\mu_2 \mid \mu_1, q_1, q_2, y_1 \dots, y_n) = p(\mu_2 \mid q_2, y_{t+1}, \dots y_n) \sim N\left(\frac{r\cdot\eta + q\sum_{i=t+1}^{n} y_i}{r + (n-t)\cdot q}, \frac{1}{r + (n-t) q}\right)$$

Tilsvarende

$$p(q_1 \mid q_2, \mu_1, \mu_2, y_1 \dots y_n) \propto p(q_1)\cdot \prod_{i=1}^{t} p(y_i \mid q_1, \mu_1)$$

$$\propto q_1^{\alpha-1} e^{-\beta\cdot q_1} \cdot q_1^{t/2} e^{-\frac{q_1}{2}\sum_{i=1}^{t}(y_i-\mu_1)^2}$$

$$\sim \text{Gamma}\left(\alpha + t/2,\ \beta + \frac{1}{2}\sum_{i=1}^{t}(y_i-\mu)^2\right)$$

$$p(q_2 \mid q_1, \mu_1, \mu_2, y_1 \dots y_n) = p(q_2 \mid \mu_2, y_{t+1}, \dots y_n)$$

$$\sim \text{Gamma}\left(\alpha + (n-t)/2,\ \beta + \frac{1}{2}\sum_{i=t+1}^{n}(y_i-\mu)^2\right)$$

c)

2 c)

$$p(t) = \frac{1}{n} \qquad t = 1, \ldots, n.$$

$$q(t^{new} \mid t^b) = \frac{1}{2h+1}, \quad t^{new} = t^b - h, \ldots \; t^b + h$$

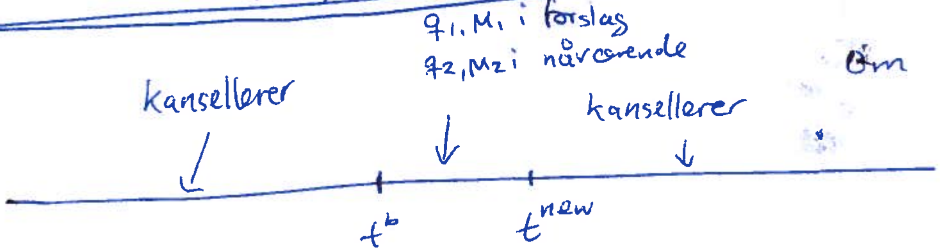
Accept-rate:

$$\min\left(1, \frac{\prod_{i=1}^{t^{new}} p(y_i \mid q_1, \mu_1) \cdot \prod_{i=t^{new}+1}^{n} p(y_i \mid q_2, \mu_2) \cdot p(\mu_1) \cdot p(\mu_2) \cdot p(q_1) p(q_2) p(t^{new}) \; q(t^b \mid t^{new})}{\prod_{i=1}^{t^b} p(y_i \mid q_1, \mu_1) \cdot \prod_{i=t^b+1}^{n} p(y_i \mid q_2, \mu_2) p(\mu_1) p(\mu_2) \cdot p(q_1) p(q_2) p(t^b) \; q(t^{new} \mid t^b)}\right)$$

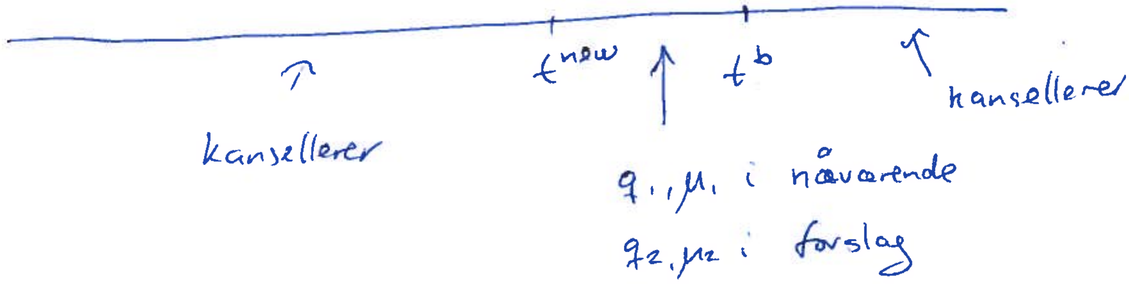
irrelevant   uniform   uniform, unless at edges, then 0.

$$= \min\left(1, \frac{\prod_{i=\min(t^b, t^{new})+1}^{t^{new}} p(y_i \mid q_1, \mu_1) \cdot \prod_{i=t^{new}+1}^{\max(t^b, t^{new})} p(y_i \mid q_2, \mu_2)}{\prod_{i=\min(t^b, t^{new})+1}^{t^b} p(y_i \mid q_1, \mu_1) \prod_{i=t^b+1}^{\max(t^b, t^{new})} p(y_i \mid q_2, \mu_2)}\right)$$

Om $t^b < t^{new}$



om $t^b > t^{new}$



vi)

Forts (2) c)

Anta $t^b < t^{new}$. Da forenkles akseptraten til

$$\min\left(1, \frac{e^{-\frac{q_1}{2}\sum_{i=t^b+1}^{t^{new}}(y_i-\mu_1)^2}}{e^{-\frac{q_2}{2}\sum_{i=t^b+1}^{t^{new}}(y_i-\mu_2)^2}}\right)$$

Motsatt, om $t^b > t^{new}$

$$\min\left(1, \frac{e^{-\frac{q_2}{2}\sum_{i=t^b+1}^{t^{new}}(y_i-\mu_2)^2}}{e^{-\frac{q_1}{2}\sum_{i=t^b+1}^{t^{new}}(y_i-\mu_1)^2}}\right)$$

d) Burn-in ser ut til å være ca 500 iterasjoner. Shift-punktet t konvergerer senest. Etter burn-in er Markovkjeden i en stasjonær fase, uten trender.

En større h ville trolig gitt kortere burn-in, vi ville gjenkjent $t \approx 60$ raskere. I den stasjonære fasen ville derimot 'mixing' vært dårlig siden for mange forslag ble forkastet, iallefall med mye større h.

e) Oppdatering av t med stor bredde h i forslags-fordelingen gir liten akseptrate siden μ_1, μ_2, q_1, q_2 ikke endres i samme runde. En simultan forslagsfordeling vil kunne gi bedre 'miksing':

HER BLE ULIKE KREATIVE FORSLAG VURDERT OG BELØNNET! ÉN LØSNING SKISSERES HER:

Forslagsfordelingen bygges som følger:

$$q(t^{new} \mid t^b) = \frac{1}{2h+1}, \quad t^{new} = t^b - h, \dots, t^b + h.$$

$$q(\mu_1^{new} \mid t^{new}, y_1, \dots, y_{t^{new}}, q_1^b) \sim N\left(\frac{r \cdot \eta + q_1^b \sum_{i=1}^{t^{new}} y_i}{r + q_1^b \cdot t^{new}}, \frac{1}{r + q_1^b t^{new}} \right)$$

$$q(\mu_2^{new} \mid t^{new}, y_{t^{new}+1}, \dots y_n, q_2^b) \sim N\left(\frac{r \cdot \eta + q_2^b \sum_{i=t^{new}+1}^{n} y_i}{r + q_2^b (n - t^{new})}, \frac{\ }{r + q_2^b (n - t^{new})} \right)$$

$$q(q_1^{new} \mid t^{new}, y_1, \dots y_{t^{new}}, \mu_1^{new}) \sim \text{Gamma}\left(\alpha + \frac{t_{new}}{2}, \beta + \frac{1}{2} \sum_{i=1}^{t^{new}} (y_i - \mu_1^{new})^2 \right)$$

$$q(q_2^{new} \mid t^{new}, y_{t^{new}+1}, \dots y_n, \mu_2^{new}) \sim \text{Gamma}\left(\alpha + \frac{(n - t_{new})}{2}, \beta + \frac{1}{2} \sum_{i=t^{new}+1}^{n} (y_i - \mu_2^{new})^2 \right)$$

Akseptraten blir:

$$\min\left(1, \frac{\prod_{i=1}^{t^{new}} p(y_i \mid \mu_1^{new}, q_1^{new}) \cdot \prod_{i=t^{new}+1}^{n} p(y_i \mid \mu_2^{new}, q_2^{new}) \cdot p(\mu_1^{new}) p(\mu_2^{new}) \, p(q_1^{new}) p(q_2^{new}) \, \cancel{p(t^{new})} \, q(\circ^{b} \mid \circ^{new})}{\prod_{i=1}^{t^b} p(y_i \mid \mu_1^b, q_1^b) \cdot \prod_{i=t^b+1}^{n} p(y_i \mid \mu_2^b, q_2^b) \, p(\mu_1^b) p(\mu_2^b) \cdot p(q_1^b) p(q_2^b) \, \cancel{p(t^b)} \, q(\circ^{new} \mid \circ^{b})} \right)$$

uniforme

Merk at forslaget for q_1 og q_2 er basert på de fulle betingede fordelingene:

$$q(q_1^{new} \mid t^{new}, y_1, \dots y_{t^{new}}, \mu_1^{new}) \propto p(q_1^{new}) \cdot \prod_{i=1}^{t^{new}} p(y_i \mid \mu_1^{new}, q_1^{new})$$

$$q(q_2^{new} \mid t^{new}, y_{t^{new}+1}, \dots, y_n, \mu_2^{new}) \propto p(q_2^{new}) \cdot \prod_{i=t^{new}+1}^{n} p(y_i \mid \mu_2^{new}, q_2^{new})$$

Dette gjør at en hel del kanselleres i aksept-raten.

Et alternativ er å foreslå q_1^{new}, q_2^{new} for μ_1^{new}, μ_2^{new},

Da vil på lignende vis

$$q(\mu_1^{new} \mid t^{new}, y_1, \ldots y_{t^{new}}, q_1^{new}) \propto p(\mu_1^{new}) \cdot \prod_{i=1}^{t^{new}} p(y_i \mid \mu_1^{new}, q_1^{new}),$$

Forskjellen er en betinging på q_1^{new} istedet for q_1^{b}.

Tilsvarende for μ_2^{new} sitt forslag.

Denne simultane forslagsfordelingen vil ha større kostnad per beregning, men trolig vil 'mixing' bli bedre, og det gjør at MCMC algoritmen kan kjøres over færre iterasjoner.

3

a) Bootstrap vil lage replikate verdier av $\bar{y}$, ved å trekke fra <u>empirisk fordeling</u>.

$$(y_1, \dots y_n) \xrightarrow{1} (y_1^{*1}, \dots y_n^{*1}), \quad \bar{y}^{*1} = \frac{1}{n}\sum_{i=1}^{n} y_i^{*1}$$

$$\vdots$$

$$(y_1, \dots y_n) \xrightarrow{B} (y_1^{*B}, \dots y_n^{*B}), \quad \bar{y}^{*B} = \frac{1}{n}\sum_{i=1}^{n} y_i^{*B}$$

Her trekkes hver y_i^{*b} blant $(y_1, \dots y_n)$. Hver av $y_1, \dots y_n$ har sannsynlighet $1/n$, og trekningene gjøres <u>med</u> tilbakelegging.

Ideen er å bruke empirisk fordeling, uten noen form for antakelser om den underliggende modellen.

Pseudokode:

```
For b = 1, ..., B
    for i = 1, ..., n
        Trekk uniformt tall j ∈ {1, ..., n}
        Sett y_i^{*b} = y_j
    end
    Sett ȳ^{*b} = (1/n) Σ_{i=1}^{n} y_i^{*b}
end
```

Fordelingen eller utvalgsfordelingen til $\bar{y}$ kan studeres utfra $\bar{y}^{*1}, \dots, \bar{y}^{*B}$.

x1

3) b) Et 90% konfidensintervall for $E(\bar{y})$ gjøres ved å plukke nedre 5% percentil og øvre 95% percentil av $\bar{y}^{*1}, \dots \bar{y}^{*B}$.

Sortering gir
$\bar{y}^{*(5)} = 4.41$, $\bar{y}^{*(95)} = 14.89$

konfidensintervall $\approx$ (4.41, 14.89).

HER FINNS MER SOFISTIKERTE ALTERNATIV!
DE HAR IKKE BLITT GJENNOMGÅTT I KURSET.

De klassiske konfidensintervallene antar at underliggende fordeling for data er normalfordelt. $y_1, \dots y_n$ er et uavhengig utvalg og $y_i \sim N(\mu, \sigma^2)$.
s^2 er et forventningsrett estimat på σ^2.

Normalfordelingen, med $z_{0.95}$ percentil, er også gyldig når $n \to \infty$ under antakelsene knyttet til sentralgrenseteoremet

$$\frac{\bar{y} - \mu}{\sigma / \sqrt{n}} \xrightarrow[n \to \infty]{} N(0,1)$$