

STK1000: Løsningsforslag Uke 40

2019

Oppgave 5.27

a)

```
x = c(98, 63, 137, 210, 52, 88, 151, 133, 105, 168)
mu = mean(x)
mu
```

```
## [1] 120.5
```

b)

Vi trekker tre tilfeldige individer og regner ut gjennomsnittet av de tre.

```
sub = sample(x, 3)
sub
```

```
## [1] 63 88 52
```

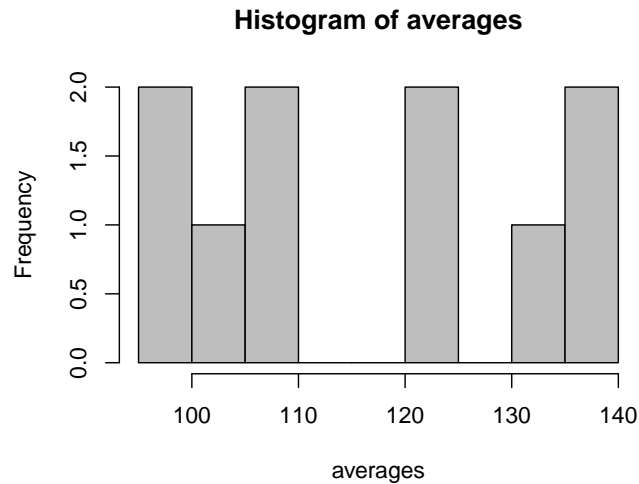
```
mean(sub)
```

```
## [1] 67.66667
```

c)

Vi repeterer b) 10 ganger med en for-løkke

```
n = 10
averages = numeric(n) # lager en liste med 10 nullere (som vi skal fylle inn senere)
for (i in 1:n) { # repeterer 10 ganger
  sub = sample(x, 3) # trekker 3 tilfeldige individer
  avg = mean(sub) # regner ut gjennomsnitt
  averages[i] = avg # fyller inn plass "i" med gjennomsnittet
}
hist(averages, col = 'gray', breaks = 10)
```



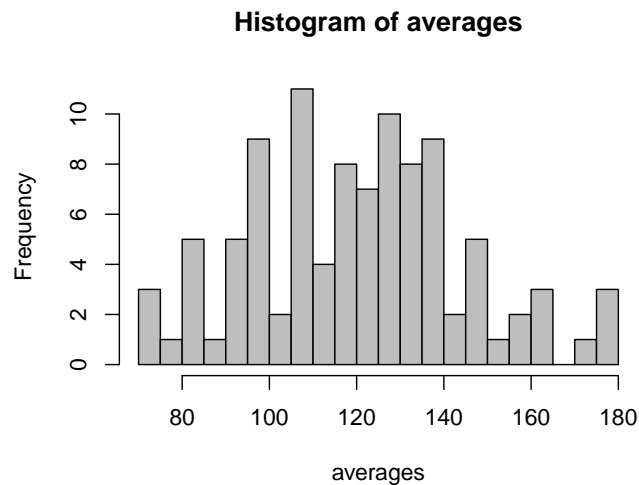
d)

Vi ser at midten av histogrammet ikke er så langt fra populasjonsgjennomsnittet på 120.5. Vi vet at gjennomsnittet av utvalget er en forventningsrett (unbiased) estimator av populasjonsgjennomsnittet. Det vil si at utvalgsgjennomsnittene våre skal fordele seg rundt populasjonsgjennomsnittet.

e)

Som vi diskuterte i d) ser vi nå at histogrammet fordeler seg rundt populasjonsgjennomsnittet på 120.5.

```
n = 100
averages = numeric(n) # lager en liste med 100 nullere (som vi skal fylle inn senere)
for (i in 1:n) { # repeterer 100 ganger
  sub = sample(x, 3) # trekker 3 tilfeldige individer
  avg = mean(sub) # regner ut gjennomsnitt
  averages[i] = avg # fyller inn plass "i" med gjennomsnittet
}
hist(averages, col = 'gray', breaks = 30)
```



Oppgave 5.28

- a) $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.24}{\sqrt{120}} = 0.113$.
- b) I 95 % av utvalgene vil vi ha et gjennomsnitt i intervallet $\mu \pm 2\sigma_{\bar{x}} = 6.78 \pm 2 \cdot 0.113 = [6.552, 7.006]$.
- c) $P(\bar{x} < 6.9) = P(Z < \frac{6.9-6.78}{0.113}) = P(Z < 1.06) = 0.8554$.

Oppgave 5.29

- a) Vi trenger et større utvalg enn 120, ettersom intervallet i Oppgave 5.28 b) er større enn det vi ønsker det skal være.
- b) Ettersom vi ønsker $\mu + 2\sigma_{\bar{x}} \leq \mu + 0.08$ trenger vi $\sigma_{\bar{x}} \leq \frac{0.08}{2} = 0.04$.
- c) Vi har at $\sigma_{\bar{x}} = \frac{1.24}{\sqrt{n}} \leq 0.04$. Hvis vi snur om på uttrykket får vi at $n \geq (\frac{1.24}{0.04})^2 = 961$, som vi si at vi trenger $n = 961$.

Oppgave 5.35

- a) At \bar{x} er forventningsrett (ikke er biased) betyr at \bar{x} ikke er systematisk for høy eller for lav. Hvis man gjentar utvalget vil man få \bar{x} 'er som er spredd rundt den sanne verdien μ .
- b) Jo større utvalget våres er, jo nærmere vil \bar{x} være den sanne verdien μ .

Oppgave 5.36

- a) Vi har bruker definisjonene på side 248 og 256 og får
- $$\mu = 4 \cdot 0.18 + 3 \cdot 0.31 + 2 \cdot 0.26 + 1 \cdot 0.13 + 0 \cdot 0.12 = 2.3.$$
- $$\sigma^2 = (4 - 2.3)^2 \cdot 0.18 + (3 - 2.3)^2 \cdot 0.31 + (2 - 2.3)^2 \cdot 0.26 + (1 - 2.3)^2 \cdot 0.13 + (0 - 2.3)^2 \cdot 0.12 = 1.55.$$
- $$\sigma = \sqrt{1.55} = 1.245.$$
- b) Vi har at $\mu_{\bar{x}} = \mu = 2.3$ og $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{25}} = 0.249$.
- c) $P(X \geq 3) = P(X = 4) + P(X = 3) = 0.18 + 0.31 = 0.49$.
- d) $P(\bar{x} \geq 3) = P(Z \geq \frac{3-2.3}{0.249}) = P(Z \geq 2.81) = 1 - P(Z < 2.81) = 0.0025$.

Oppgave 5.37

- a) $\mu_{\bar{x}} = \mu = 0.4$ og $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.9}{\sqrt{100}} = 0.09$.
- b) $P(\bar{x} > 0.5) = P(Z > \frac{0.5-0.4}{0.09}) = P(Z > 1.11) = 1 - P(Z \leq 1.11) = 0.1335$. Her har vi bruk sentralgrenseteoremet når vi antar at \bar{x} er normalfordelt.
- c) Ja, når n er så stor som 100 er det en rimelig antagelse at \bar{x} er normalfordelt.

Oppgave 5.38

- a) Hvis man bare selger 100 poliser vil sannsynligheten for å få noen krav være liten. Men på den andre siden vil et enkelt krav fort være større enn verdien av alle polisene (så selskapet går konkurs hvis det blir et krav). Hvis man selger mange tusen poliser sier sentralgrenseteoremet at gjennomsnittlig krav være cirka lik forventningsverdien, $\bar{x} \approx \mu$. Slik vil man være (neste helt) sikker på at verdien av alle polisene er større enn kravene.
- b) Vi har $\mu_{\bar{x}} = \mu = 500$ og $\sigma_{\bar{x}} = \frac{10000}{\sqrt{50000}} = 44.72$. Sannsynligheten for at gjennomsnittlig krav skal være større enn prisen for politen er derfor $P(\bar{x} > 600) = P(Z > \frac{600-500}{44.72}) = P(Z > 2.24) = 0.0125$. Altså vil dette bar inntreffe ca hvert hundrede år.

Oppgave 5.60

- a) Nei. Du kan ikke telle tid siden det er kontinuerlig.
- b) Ja, her er det naturlig å anta at $X \sim B(20, p)$, hvor p er sannsynligheten for et defekt par.
- c) Ja, her kan vi anta at $X \sim B(n, p)$ hvor n er størrelsen på utvalget og p er sannsynligheten for at en tilfeldig valgt student sier han spiser minst 5 om dagne.
- d) Ja, til dels. Vi kan anta at $X \sim B(n, p)$ hvor n er antall dager i skoleåret og p er sannsynligheten for at du skulker. Her må vi i så fall anta at alle dagene er uavhengig av hverandre (om du skulker i dag påvirker ikke om du skulker i morgen).

Oppgave 5.62

- a) Individene er trolig uavhengige av hverandre, og uten mer informasjon kan vi anta at alle har samme sannsynlighet for å fortsette med nedlastingen på $p = 0.33$. Vi kan derfor anta at $X \sim B(50, 0.33)$.

Oppgave 5.64

- a) Forventningsverdien til X vil være $\mu_X = 15 \cdot 0.75 = 11.25$. Forventningsverdien til \hat{p} vil være $\mu_{\hat{p}} = \frac{\mu_X}{n} = 0.75$.
- b) For $n = 150$ har vi $\mu_X = 150 \cdot 0.75 = 112.5$ og for $n = 1500$ har vi $\mu_X = 150 \cdot 0.75 = 1125$. På den annen side vil alltid $\mu_{\hat{p}} = 0.75$. Når antallet i utvalget øker (n øker) vil forventet antall øke, mens forventet proporsjon vil alltid være konstant.

Oppgave 5.70

Vi har $n = 1020$ og $p = 0.48$ som vil si at np og $n(1-p)$ er begge rundt 500. Vi kan derfor trygt bruke Normal-approximasjonen her. Da har vi at $\mu_{\hat{p}} = 0.48$ og $\sigma_{\hat{p}} = \sqrt{\frac{0.48 \cdot (1-0.48)}{1020}} = 0.0156$. Vi har ha sannsynligheten for at estimert \hat{p} ligger innenfor oppgitt feilmargin (på ± 0.04) er

$$\begin{aligned} P(0.44 < \hat{p} < 0.52) &= P\left(\frac{0.44 - 0.48}{0.0156} < Z < \frac{0.52 - 0.48}{0.0156}\right) \\ &= P(-2.56 < Z < 2.56) = 0.9948 - 0.0052 = 0.9896. \end{aligned}$$

Det er altså at det er veldig sannsynlig at estimatet ligger innen for en feilmargin på ± 0.04 .

Oppgave 5.72

Her gjentar vi bare utregningene fra Oppgave 5.70 for $n = 300$ og $n = 5000$, som vil si at vi må regne ut $\sigma_{\hat{p}}$ på nytt.

For $n = 300$ får vi da $P(0.44 < \hat{p} < 0.52) = 0.8354$, og for $n = 5000$ får vi $P(0.44 < \hat{p} < 0.52) \approx 1$. Altså ser vi at større utvalg har større sannsynlighet for å gi estimater nærme den sanne populasjonen.

Oppgave 5.76

Vi har $\sigma_{\hat{p}} = \sqrt{\frac{0.48 \cdot (1-0.48)}{n}} = 0.005$. Hvis vi snur om på dette uttrykket får vi at $n = \frac{0.48 \cdot (1-0.48)}{0.005^2} = 9984$.

Oppgave 5.77

- Man har $p = \frac{1}{4} = 0.25$ sannsynlighet for gjette riktig på et vilkårlig kort.
- Her har vi en binomisk fordeling $B(20, 0.25)$ som gir

$$\begin{aligned} P(X \geq 10) &= P(X = 10) + P(X = 11) + \dots + P(X = 20) \\ &= 0.0099 + 0.0030 + 0.0008 + 0.0002 + 0 + \dots + 0 \\ &= 0.0139. \end{aligned}$$

Vi har brukt Tabell C i boken for å finne tallene.

- Vi har forventning $\mu = np = 20 \cdot 0.25 = 5$, med standardavvik $\sigma = \sqrt{np(1-p)} = \sqrt{3.75} = 1.9365$.
- Nei, vi kan ikke lenger bruke binomisk fordeling her. Det står at forsøket utføres med å gå gjennom **ALLE** kortene og vi vet at det er 5 av hver type, så forsøkene er ikke uavhengig av hverandre. Hvis man i stedet hadde utført forsøket med å trekke tilfeldig fra bunken med tilbakelegging, kunne vi brukt binomisk fordeling.

Oppgave 5.83

- Siden vi har et intensitet på 0.5 per cm^2 og en overflate på 100 cm^2 har vi en Poisson-fordeling med forventning på $0.5 \cdot 100 = 50$.
- Vi har $\mu = 50$ og $\sigma = \sqrt{50}$. $P(X > 60) = P(Z > \frac{60-50}{\sqrt{50}}) = P(Z > 1.41) = 0.0793$.

Midtveiseksamen Høsten 2009

- $p_3 = 1 - 0.3 - 0.5 = 0.2$ som gir b).
- $-1 \cdot \frac{1}{6} + 0 + z \cdot \frac{1}{6} = 1.5$. Det betyr at $z = 1.5 \cdot 6 + 1 = 10$ som gir d).
- Median er $\frac{x+700}{2} = 650$ som betyr at $x = 600$ som gir d).
- $\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y = 3^2 + 1^2 + 2 \cdot 0.5 \cdot 3 \cdot 1 = 13$ som gir $\sigma_{X-Y} = 3.606$, som gir a).
- $P(X > 37.5) = P(Z > \frac{37.5-37}{0.3}) = P(Z > \frac{5}{3}) = 0.0478$, som gir c).
- $P(Z < z) = 0.75$ gir at $z = 0.647$. Vi får derfor temperaturen $0.647 \cdot 0.3 + 37.0 = 37.19$ som gir a).
- c).

10) Vi har d) fordi

$$\begin{aligned}P(M) &= P(M \cap S) + P(M \cap S^C) \\&= P(M | S)P(S) + P(M | S^C)P(S^C) \\&= 0.95 \cdot 0.007 + 0.035 \cdot (1 - 0.007) \\&= 0.041405.\end{aligned}$$