Kap. 14

Mål:

- Kunne rekne ut gjennomsnitt, varians og standardavvik
- Forstå kva desse størrelsane forteller om ei mengde tall.
- Vite kva ein stokastisk variabel er, og rekne ut
  - forventningsverdi
  - varians
  - standardavvik.
- sannsynlighetsmodeller som beskriver utfall frå endelige mengder.
- sannsynlighetsmodeller for kontinuerlege mengder.

1

Def. (Gjennomsnitt)

La $x_1, x_2, \dots, x_n$ vere n tall. Gjennomsnittet er

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum x_i$$

Variansen er:

$$S^2 = \frac{(\bar{x}-x_1)^2 + (\bar{x}-x_2)^2 + \cdots + (\bar{x}-x_n)^2}{n-1}$$

$$= \frac{1}{n-1}\sum (\bar{x}-x_i)^2$$

Standardavvik:

$$S = \sqrt{S^2}$$

Eksempel 1

$$X = \{5, 10, 30, 35\}$$

$$\bar{X} = \frac{5 + 10 + 30 + 35}{4} = \underline{20}$$ (gj. snitt)

$$S^2 = \frac{(20-5)^2 + (20-10)^2 + (20-30)^2 + (20-35)^2}{4-1}$$

$$= \frac{(+15)^2 + (+10)^2 + (-10)^2 + (-15)^2}{3}$$

$$= 216.66 = \underline{14.72^2}$$ (varians)

$$S = \underline{14.72}$$ (standardavvik)

Eksempel 2

$$Y = \{18, 19, 21, 22\}$$

$$\bar{Y} = \frac{18 + 19 + 21 + 22}{4} = \underline{20}$$

$$S^2 = \frac{(20-18)^2 + (20-19)^2 + (20-21)^2 + (20-22)^2}{4-1}$$

$$= 3.33 = \underline{1.83^2}$$

$$\underline{S = 1.83}$$

Standardavvik:

- liten verdi: tala er nærme gjennomsnittet
- stort verdi: tala er langt frå gjennomsnitt

Stokastisk variabel

- I sannsynlighet er eit <u>utfall</u> eit mulig resultat av eit tilfeldig forsøk.
- I sannsynlighet er <u>hendelse</u> er mengda av mulige utfall.

Eksempel

I eit terningkast er hendelsane $\{1, 2, 3, 4, 5, 6\}$

Eit utfall er t.d verdien 4.

Stokastisk variabel

Ein stokastisk variabel X er ein variabel gitt ved utfalla x av eit tilfeldig forsøk, sammen med sannsynligheta for at x skjer $P(X = x)$

Eksempel

Terningkast: $X = \{1, 2, 3, 4, 5, 6\}$

$P(X = 2) = 1/6$

$P(X = 6) = 1/6$

Døme (Roulette)

Vilkårlege tall: $\{0, 00, 1, 2, 3, \ldots, 36\}$

$x: \{2 \text{ grønne}, 18 \text{ røde}, 18 \text{ svarte}\}$

Anta at du satser 10 kroner på rød og 15 kroner på oddetal.

La X vere den stokastiske variabelen

$$X = \{\text{nettogevinst}\} = \{\underbrace{-25}_{\text{viss svart partall/grønn}}, \underbrace{-25+2\cdot 10}_{\text{rødt partal}}, \underbrace{-25+2\cdot 15}_{\text{svart oddetal}}, \underbrace{-25+2\cdot 10+2\cdot 15}_{\text{rødt oddetal}}\}$$

$$= \{-25, -5, 5, 25\}$$

$P(X = -25) = \frac{11}{38}$

$P(X = -5) = \frac{9}{38}$

$P(X = 5) = 9/38$

$P(X = 25) = \frac{9}{38}$

Kva kan eg forvente å vinne?

Def (Forventningsverdi, varians, standardavvik)

Forventningsverdien til den stokastiske variabelen X er

$$E(X) = \sum x \cdot P(X = x)$$

Variansen til X er:

$$\text{Var}(X) = \sum_x (x - E(X))^2 \cdot P(X = x)$$

!!

Standardavviket:

$$\text{std}(X) = \sqrt{\text{Var}(X)}.$$

Eksempel (fort)

$$E(X) = (-25)\cdot\frac{11}{38} + (-5)\cdot\frac{9}{38} + 5\cdot\frac{9}{38} + 25\cdot\frac{9}{38} = \underline{-1.32}$$

$$\text{Var}(X) = (-25-(-1.32))^2\cdot\frac{11}{38} + (-5-(-1.32))^2\cdot\frac{9}{38} + (5-(-1.32))^2\frac{9}{38} + (25-(-1.32))^2\cdot\frac{9}{38} = 338.9281 = \underline{18.41^2}$$

$$\text{std}(X) = \underline{18.41}$$

Tilfeldig uttrekk frå endelige mengder

- Uttrekk med tilbakelegging: Etter kvart forsøk legger ein trukket verdi tilbake i mengda.
  - trekke kort frå kort frå kortstokk, der vi legger kortet tilbake
  - trille terning
  - Roulette

- Uttrekk uten tilbakelegging: Etter kvart forsøk legger man ikkje verdi tilbake.
  - Trekker kuler frå ei skål, uten å legge tilbake
  - Trekke kort frå kortstokk, uten å legge kort tilbake
  - Lotto

Binomialfordelinga

(5)

La X vere ein stokastisk variabel som teller antall positive utfall på $\underline{n}$ forsøk med sannsynlighet $\underline{p}$.

$$X \sim bin(n, p)$$

Eksempel

Anta at vi kaster terning 1000 ganger. La $X = \{\text{verdi } 5\}$.

$$X \sim bin\left(1000, \tfrac{1}{6}\right)$$

Resultat

Viss $X \sim bin(n, p)$ er:

$$E(X) = n \cdot p$$

$$Var(X) = np(1-p)$$

Eksempel (fort)

$$X \sim bin\left(1000, \tfrac{1}{6}\right)$$

$$E(X) = 1000 \cdot \tfrac{1}{6} = \underline{166,67}$$

$$Var(X) = 1000 \cdot \tfrac{1}{6} \cdot \left(1 - \tfrac{1}{6}\right) = 138,88 = \underline{11,79^2}$$

Hypergeometrisk fordeling

La X vere ein stokastisk variabel som teller positive utfall frå ei endeleg mengde uten å legge tilbake. La

N = total antal element

M = total antal positive utfall

n = antall forsøk

$$X \sim \text{hypergeom}(N, M, n)$$

Forventningsverdi og varians

$$E(X) = n \cdot \frac{M}{N}$$

$$\text{Var}(X) = n \frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$$

Eksempel

Anta at ei skål har 20 kuler, der 5 kuler er raude.

La Y vere den stokastiske variabelen som teller antal raude kuler som vi trekker på 10 trekk.

$$Y \sim \text{hypergeom}(20, 5, 10)$$

$$E(Y) = 10 \cdot \frac{5}{20} = \underline{2.5}$$

$$\text{Var}(Y) = 10 \cdot \frac{5}{20}\left(1 - \frac{5}{20}\right)\frac{20-10}{20-1} \simeq \underline{1^2}$$

$$\text{sd}(Y) = \underline{\underline{1}}$$

Binomialkoeffisienter

La k og n vere heiltal $0 \leq k \leq n$. Da er binomial-koeffisienten $\binom{n}{k}$ gitt ved

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Repetisjon

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 1$$

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

Eksempel

$$\binom{8}{3} = 8 \text{ 'nCr' } 3 = \underline{56}$$

$$\binom{10}{7} = \underline{120}$$

Binomialkoeffisienten

La $0 \le k \le n$. Da

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = n \; \boxed{nCr} \; k$$

Pascals trekant

$\downarrow n$

1	$n=0$
1 1	$n=1$
1 2 1	$n=2$
1 3 3 1	$n=3$
1 4 6 4 1	$n=4$
1 5 10 10 5 1	$n=5$
1 6 15 20 15 6 1	$n=6$

$k=0$, $k=1$, ..., $k=4$ (under 15)

$$\binom{6}{4} = 15$$

Eksempel

La ei skål inneholde raude og svarte kuler.

Kor mange måter kan eg trekke 2 raude kuler viss eg trekker 5 kuler.

RRSSS SRRSS SSRRS SSSRR
RSRSS SRSRS SSRSR
RSSRS SRSSR
RSSSR

$$\binom{5}{2} = 10 \text{ mulige kombinasjoner}$$

- Binomialkoeffisienten av to tal n og k svarer til kor mange kombinasjoner ein kan velge k element frå ei mengde med n element.

Ekempel

La ei skål inneholde 20 kuler, 5 røde og 15 svarte.
Vi trekker 5 kuler uten og legger tilbake

$p(R) = \frac{1}{4}$

$p(S) = \frac{3}{4}$

Finn sannsynligheten for at nøyaktig 2 av kulene er røde.

$p(2 \text{ røde}, 3 \text{ svarte})$

$= p(RRSSS) + p(RSRSS) + \cdots + p(SSSRR)$

$= \frac{1}{4}\cdot\frac{1}{4}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{3}{4} + \frac{1}{4}\cdot\frac{3}{4}\cdot\frac{1}{4}\cdot\frac{3}{4}\cdot\frac{3}{4} + \cdots + \frac{3}{4}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{1}{4}\cdot\frac{1}{4}$

$= \left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3 + \left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3 + \cdots + \left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3$

$= \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3 = \underline{\underline{0.2637}}$

Binomialfordelinga

2

- n forsøk med lik sannsyn p.
- X teller antal positive utfall.

$$X \sim bin(n, p)$$

Vi veit

$$E(X) = np$$

$$V(X) = np(1-p)$$

Sannsynligheten for nøyaktig x positive utfall:

$$p(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Hypergeometrisk fordeling

- Anta at vi har ei mengde med N element, der M er positive utfall.
- Anta vi trekker n element uten å legge tilbake.

Eksempel
sannsyn for først 2 spar og så eitt hjerte.

$$\frac{13}{52}\cdot\frac{12}{51}\cdot\frac{13}{50}$$

$$X \sim \text{hypergeom}(N, M, n)$$

Vi har

$$E(X) = n\frac{M}{N}$$

$$Var(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$$

Sannsynligheten for at nøyaktig x av utfalla er positive:

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

14.5.5

105 studenter
42 pendler til campus
Trekker ut 8 studenter.

La
$$X = \{\text{antall pendlere}\}$$

a) $X \sim \text{hypergeom}(105, 42, 8)$

b) $E(X) = 8\cdot\frac{42}{105} = \underline{\underline{3.2}}$

c) $P(\text{nøyaktig } 4)$

$$= P(X=4) = \frac{\binom{42}{4}\binom{105-42}{8-4}}{\binom{105}{8}} = \frac{\binom{42}{4}\binom{63}{4}}{\binom{105}{8}} = \underline{\underline{0.239}}$$

d) $P(X \le 7) = P(X=0) + P(X=1) + \dots + P(X=7)$

$$= 1 - P(X=8) = 1 - \frac{\binom{42}{8}\binom{105-42}{8-8}}{\binom{105}{8}} = \underline{\underline{0.9996}}$$

Poisson fordeling

I Poissonfordelinga beskriver vi tilfelle der vi har eit endeleg antal positive utfall i ei kontinuerleg mengde.

- anta at vi har eit endeleg antal positive utfall som er uavhengige av kvarandre.
- La $\lambda > 0$ vere ein frekvenskonstant.
- La t vere antal enheter.

Da er λt gjennomsnittlig antall positive utfall i t enhet.

Da er $X = \{\text{antall positive utfall}\}$

$$\sim P_o(\lambda t)$$

Da er

$$E(X) = \lambda t, \quad Var(X) = \lambda t$$

Saunsynligheten for x utfall

$$P(X = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$

Eksempel (14.5.7)

Ein basestasjon mottar i snitt 4 sms pr. minutt.

La $X = \{\text{antall sms på 20 sekund}\}$.

Fordeling?

$$\lambda = 4 \text{ sms/min}$$

$$t = 20 \text{ sek} = \tfrac{1}{3} \text{ minutt}$$

Da er

$$X \sim P_o(4 \cdot \tfrac{1}{3}) = P_o(\tfrac{4}{3})$$

Forventningsverdi: $E(X) = \frac{4}{3}$

a) P(nøyaktig 2 sms)

$$= P(X = 2) = \frac{(4/3)^2}{2!} e^{-4/3} = \underline{\underline{0.234}}$$

b) P(minst ein sms)

$$= P(X \geq 1) = P(X=1) + P(X=2) + P(3) + \cdots\cdots$$

$$= 1 - P(X = 0) = 1 - \frac{(4/3)^0}{0!} e^{-4/3} = \underline{\underline{0.736}}$$

Q4 I hver oppgave skal du finne fordelingen og regne ut en sannsynlighet.

4.a) X = antall ganger vi får fem eller seks på 10 terningkast. Hvilken fordeling har X? Regn ut $P(X = 4)$.

4.b) Y = antall sykler som passerer et tellepunkt på ti minutt, når det i snitt passerer 50 i timen. Hvilken fordeling har Y? Regn ut $P(Y \geq 2)$.

4.c) I en skuff ligger det ledninger med ulike koblinger: 4 USB-C, 5 micro-USB og 2 USB-3. Z = antall USB-C vi får når fem ledninger trekkes tilfeldig. Hvilken fordeling har Z? Regn ut $P(Z = 3)$.

Eksamen (H18)

a) $X = \{$antall 5 eller 6 på 10 terningkast$\}$

$X \sim bin(10, \frac{1}{3})$

$P(X=4) = \binom{10}{4}\left(\frac{1}{3}\right)^4\left(1-\frac{1}{3}\right)^{10-4}$

$= \binom{10}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^6 = \underline{\underline{0.2276}}$

b) $Y = \{$antall sykler på 10 minutt$\}$.

- anta at sykler ikke kjem på vilkårlige tidspunkt.

$\lambda = 50$ sykler/time

$t = 10$ min $= 1/6$ time

$Y \sim Po(50 \cdot \frac{1}{6}) = Po(\frac{25}{3})$

⑤

P(meir eller lik 2 sykler)

$= 1 - P(Y=0) - P(Y=1)$

$= 1 - \frac{(25/3)^0}{0!}e^{-25/3} - \frac{(25/3)^1}{1!}e^{-25/3}$

$= \underline{\underline{0.9978}}$

c) Skuffe: $\begin{cases} 4 & \text{USB-C} \\ 5 & \text{mikro-USB} \\ 2 & \text{USB-3} \end{cases}$

Vi tok ut 5 vilkårlige koblinger.

$Z = \{$antall USB-C$\} \sim$ hypergeom$(11, 4, 5)$

$P(Z=3) = \frac{\binom{4}{3}\binom{11-4}{5-3}}{\binom{11}{5}} = \frac{\binom{4}{3}\binom{7}{2}}{\binom{11}{5}} = \underline{\underline{0.1818}}$