Sandsynlighedsregning 4. forelæsning

Bo Friis Nielsen

Anvendt Matematik og Computer Science

Danmarks Tekniske Universitet

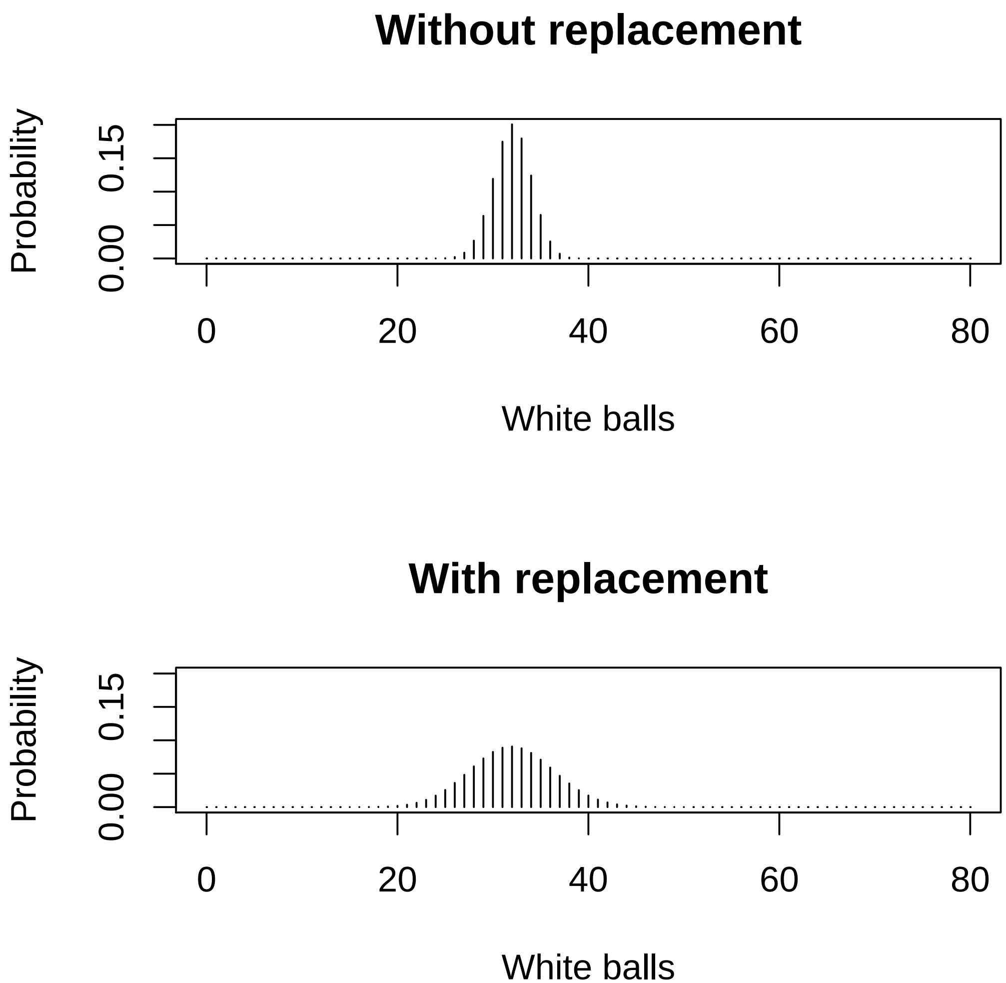
2800 Kgs. Lyngby – Danmark

Email: bfni@dtu.dk

Dagens emner: Afsnit 3.3 og 3.4

- Varians/standardafvigelse
$$\mathrm{Var}(X) = \mathsf{E}[(X - \mathsf{E}(X))^2] = \sum_{\text{alle } x} (x - \mathsf{E}(X))^2 \mathsf{P}(X = x)$$
- Normalfordelingsapproksimation/den Centrale Grænseværdisætning
$$\mathsf{P}\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \Phi(b) - \Phi(a)$$
- Markovs og Chebychevs uligheder for ekstreme udfald
$$\mathsf{P}(X \geq a) \leq \frac{\mathsf{E}(X)}{a} \qquad \mathsf{P}(|X - \mathsf{E}(X)| \geq k\mathsf{SD}(X)) \leq \frac{1}{k^2}$$
- Et udpluk af diskrete fordelinger
$$\mathsf{P}(T = i) = (1-p)^{i-1}p \qquad \mathsf{P}(Y_r = i) = \binom{i+r-1}{r-1} p^r (1-p)^i$$
- Indikatorfunktioner $I_A = \begin{cases} 1 & \text{hvis } A \text{ indtræffer} \\ 0 & \text{hvis } A^c \text{ indtræffer} \end{cases}$

Standardafvigelse/varians



Den hypergeometriske fordeling har mindre variation end binomialfordelingen.

Definition af varians

Definition af varians

$$\mathrm{Var}(X) = \mathsf{E}[(X - \mathsf{E}(X))^2] = \sum_{\text{alle } x} (x - \mathsf{E}(X))^2 \mathsf{P}(X = x)$$

Definition af varians

$$\mathrm{Var}(X) = \mathsf{E}[(X - \mathsf{E}(X))^2] = \sum_{\text{alle } x} (x - \mathsf{E}(X))^2 \mathsf{P}(X = x)$$

Beregningsformel for varians (vigtig) p.186

$$\mathsf{Var}(X) = \mathsf{E}(X^2) - (\mathsf{E}(X))^2$$

Definition af varians

$$\mathrm{Var}(X) = \mathsf{E}[(X - \mathsf{E}(X))^2] = \sum_{\text{alle } x} (x - \mathsf{E}(X))^2 \mathsf{P}(X = x)$$

Beregningsformel for varians (vigtig) p.186

$$\mathsf{Var}(X) = \mathsf{E}(X^2) - (\mathsf{E}(X))^2$$

"Shift and scale"

$$\mathsf{Var}(aX + b) = a^2 \mathsf{Var}(X)$$

Definition af varians

$$\mathrm{Var}(X) = \mathsf{E}[(X - \mathsf{E}(X))^2] = \sum_{\text{alle } x} (x - \mathsf{E}(X))^2 \mathsf{P}(X = x)$$

Beregningsformel for varians (vigtig) p.186

$$\mathsf{Var}(X) = \mathsf{E}(X^2) - (\mathsf{E}(X))^2$$

"Shift and scale"

$$\mathsf{Var}(aX + b) = a^2 \mathsf{Var}(X)$$

Sum af *uafhængige* variable

Hvis $X_1, \ldots, X_n$ er uafhængige, så er

$$\mathsf{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathsf{Var}(X_i)$$

Varians af et tal trukket fra en liste

Givet en liste af tal

$$(x_1, x_2, \ldots, x_n)$$

Lad X være en stokastisk variabel: Et tilfældigt element af listen.

$$\mathsf{E}(X) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

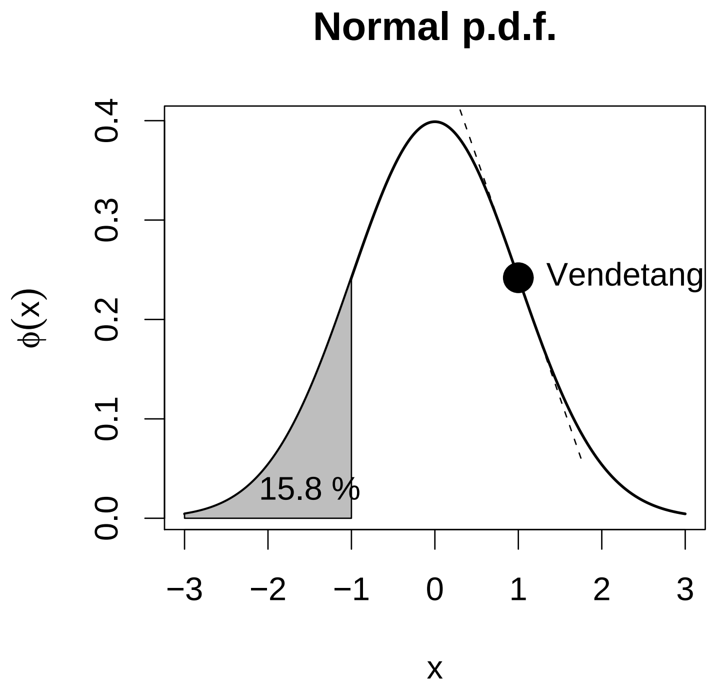
$$\mathsf{Var}(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$$

Standardafvigelse

$$\mathsf{SD}(X) = \sqrt{\mathsf{Var}(X)}$$

Standardafvigelse

$$\mathsf{SD}(X) = \sqrt{\mathsf{Var}(X)}$$



"Shift and scale"

$$\mathsf{SD}(aX + b) = |a|\mathsf{SD}(X)$$

Bestem variansen i binomialfordelingen (n, p)

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

Bestem variansen i binomialfordelingen (n,p)

Først, skriv $X \sim \mathrm{Bin}(n,p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

$$\mathsf{E}(X_i) = p \ ,$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

$$\mathsf{E}(X_i) = p \ , \ \mathsf{E}(X_i^2) = p \ ,$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

$$\mathsf{E}(X_i) = p \ , \ \mathsf{E}(X_i^2) = p \ , \ \mathsf{Var}(X_i)$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

$$\mathsf{E}(X_i) = p \ , \ \mathsf{E}(X_i^2) = p \ , \ \mathsf{Var}(X_i) = p$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n,p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

$$\mathsf{E}(X_i) = p\ ,\ \mathsf{E}(X_i^2) = p\ ,\ \mathsf{Var}(X_i) = p - p^2$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

Dernæst, variansen i en Bernoullifordeling (p):

$$\mathsf{E}(X_i) = p \ , \ \mathsf{E}(X_i^2) = p \ , \ \mathsf{Var}(X_i) = p - p^2 = p(1-p)$$

Bestem variansen i binomialfordelingen (n, p)

Først, skriv $X \sim \mathrm{Bin}(n, p)$ som

$$X = X_1 + \cdots + X_n \quad \text{hvor} \quad X_i \sim \mathrm{Bernoulli}(p)$$

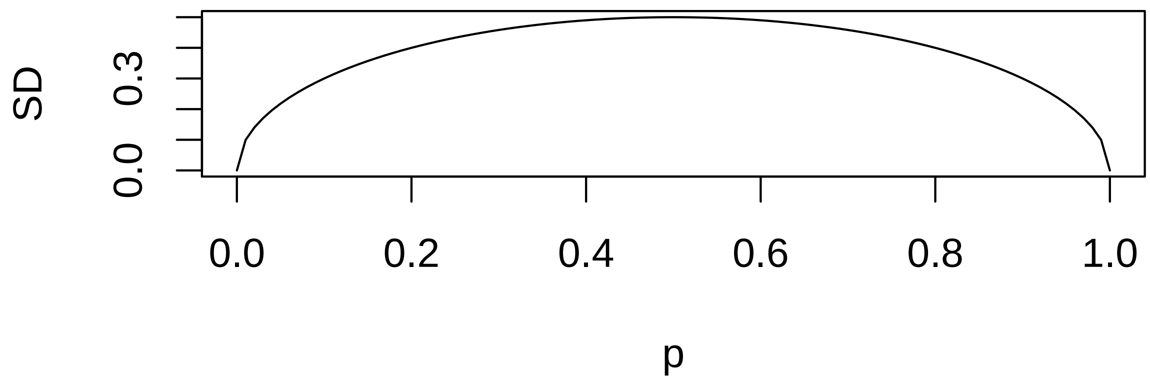
Dernæst, variansen i en Bernoullifordeling (p):

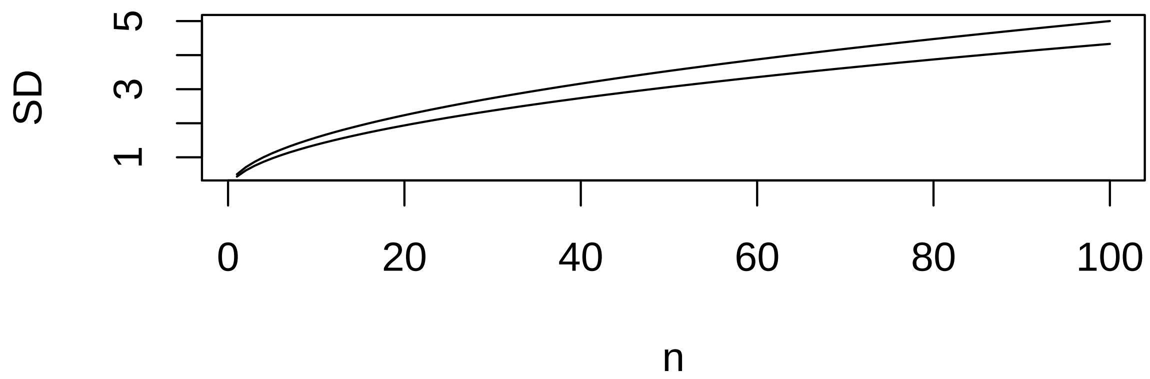
$$\mathsf{E}(X_i) = p \ , \ \mathsf{E}(X_i^2) = p \ , \ \mathsf{Var}(X_i) = p - p^2 = p(1-p)$$

Kombinér:

$$\mathsf{Var}(X) = np(1-p)$$

Spredningen i binomialfordelingen (n, p)





Lad X være antallet af dage i en måned valgt tilfældigt blandt årets tolv måneder i et ikke-skudår. Det anføres, at $\mathrm{E}(X) = \frac{365}{12}$.

Spørgsmål 1

Man finder $\mathrm{SD}(X)$ til

1 ☐ 0,71

2 ☐ 0,76

3 ☐ 0,81

4 ☐ 0,86

5 ☐ 0,91

6 ☐ Ved ikke

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n)$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad ,$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad , \qquad \mathsf{SD}(S_n)$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad , \qquad \mathsf{SD}(S_n) = \sqrt{n}\mathsf{SD}(X_i)$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad , \qquad \mathsf{SD}(S_n) = \sqrt{n}\mathsf{SD}(X_i)$$

$$\mathsf{E}(\bar{X}_n)$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad , \qquad \mathsf{SD}(S_n) = \sqrt{n}\mathsf{SD}(X_i)$$

$$\mathsf{E}(\bar{X}_n) = \mathsf{E}(X_n)$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad , \qquad \mathsf{SD}(S_n) = \sqrt{n}\mathsf{SD}(X_i)$$

$$\mathsf{E}(\bar{X}_n) = \mathsf{E}(X_n) \quad , \qquad \mathsf{SD}(\bar{X}_n)$$

Kvadratrodsloven

Lad X_i være I.I.D. (Independent and Identically Distributed).

Definér summen $S_n = X_1 + \cdots + X_n$

Definér gennemsnittet $\bar{X}_n = \frac{1}{n} S_n$.

$$\mathsf{E}(S_n) = n\mathsf{E}(X_i) \quad , \qquad \mathsf{SD}(S_n) = \sqrt{n}\mathsf{SD}(X_i)$$

$$\mathsf{E}(\bar{X}_n) = \mathsf{E}(X_n) \quad , \qquad \mathsf{SD}(\bar{X}_n) = \frac{1}{\sqrt{n}}\mathsf{SD}(X_n)$$

Store tals svage lov

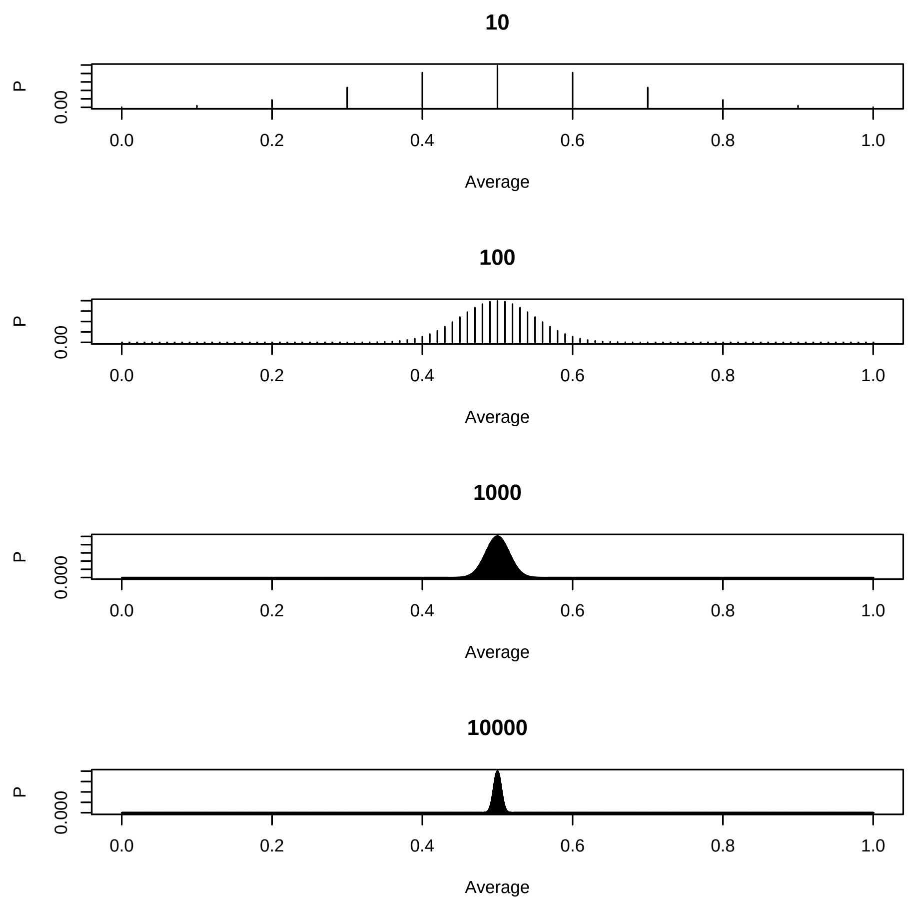
Lad X_i være I.I.D. med middelværdi μ og spredning σ.

$$\forall \epsilon > 0 :$$

$$n \to \infty \Rightarrow$$

$$\mathsf{P}(|\bar{X}_n - \mu| < \epsilon) \to 1$$

($\bar{X}_n$ konvergerer mod μ i sandsynlighed)



Den centrale grænseværdisætning p.196

Lad $X_1, \ldots, X_n$ være I.I.D. med middelværdi μ og varians σ^2.

Definér $S_n = X_1 + \cdots + X_n$.

For store n gælder approksimativt

$$\mathsf{P}\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \Phi(b) - \Phi(a)$$

Betydning

- CGS er et centralt resultat i sandsynlighedsregningen og statistikken (og findes i mange varianter)

Betydning

- CGS er et centralt resultat i sandsynlighedsregningen og statistikken (og findes i mange varianter)
- Budskabet er, at stokastiske variable, der er dannet gennem en sum af mange mindre bidrag, med god tilnærmelse kan beskrives ved en normalfordeling, som vi strengt taget først stifter bekendtskab med i næste uge

Betydning

- CGS er et centralt resultat i sandsynlighedsregningen og statistikken (og findes i mange varianter)
- Budskabet er, at stokastiske variable, der er dannet gennem en sum af mange mindre bidrag, med god tilnærmelse kan beskrives ved en normalfordeling, som vi strengt taget først stifter bekendtskab med i næste uge
- Vi har allerede set denne sætning i anvendelse for binomialfordelingen (og Poissonfordelingen)

Markovs ulighed p.174

Markovs ulighed p.174

- For ikke negative stokastiske variable, kan vi angive en øvre grænse for sandsynligheder ved brug af middelværdien.

$$\mathsf{P}(X \geq a) \leq \frac{\mathsf{E}(X)}{a}$$

Markovs ulighed p.174

- For ikke negative stokastiske variable, kan vi angive en øvre grænse for sandsynligheder ved brug af middelværdien.

$$\mathsf{P}(X \geq a) \leq \frac{\mathsf{E}(X)}{a}$$

Chebychevs ulighed p.191

Markovs ulighed p.174

- For ikke negative stokastiske variable, kan vi angive en øvre grænse for sandsynligheder ved brug af middelværdien.

$$\mathsf{P}(X \geq a) \leq \frac{\mathsf{E}(X)}{a}$$

Chebychevs ulighed p.191

- Hvis vi også kender variansen kan vi som regel skærpe denne grænse

$$\mathsf{P}(|X - \mathsf{E}(X)| \geq k\mathsf{SD}(X)) \leq \frac{1}{k^2}$$

Om en stokastisk variabel X, der beskriver et rumindhold, oplyses det, at den har middelværdi $\mathsf{E}(X) = 10$.

Spørgsmål 2

Den mindste øvre grænse for $\mathsf{P}(X \geq 20)$ findes til

1 ☐ $\frac{1}{4}$

2 ☐ $\frac{1}{2}$

3 ☐ $\frac{3}{4}$

4 ☐ $1 - \Phi\left(\frac{20-10}{\sqrt{20}}\right)$

5 ☐ En sådan grænse kan ikke bestemmes ud fra det oplyste

6 ☐ Ved ikke

Hvor Φ som sædvanlig betegner fordelingsfunktionen for en standardiseret normalfordelt variabel.

Et udvalg af diskrete fordelinger

- En række grundlæggende mekanismer
- Se pp.475 "Distribution summaries"

Spørgsmål 3

Hvis en række tal c_i udgør en sandsynlighedsfordeling, hvad må der så gælde om c_i?

1 ☐ Der er ingen specifikke krav

2 ☐ $\sum_i c_i = 1$

3 ☐ c_i kan bestemmes fra en formel som $c_i = \begin{pmatrix} n \\ i \end{pmatrix} p^i (1-p)^{n-i}$

4 ☐ $c_i \geq 0$

5 ☐ Ingen af de ovenstående er tilstrækkelige

6 ☐ Ved ikke

Den geometriske fordeling

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i)$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1 - p)$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T)$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T) = \sum_{i=0}^{\infty} \mathsf{P}(T > i)$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T) = \sum_{i=0}^{\infty} \mathsf{P}(T > i) = \frac{1}{p}$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T) = \sum_{i=0}^{\infty} \mathsf{P}(T > i) = \frac{1}{p}$$

Varians

$$\mathsf{Var}(T)$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T) = \sum_{i=0}^{\infty} \mathsf{P}(T > i) = \frac{1}{p}$$

Varians

$$\mathsf{Var}(T) = \frac{1-p}{p^2},$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T) = \sum_{i=0}^{\infty} \mathsf{P}(T > i) = \frac{1}{p}$$

Varians

$$\mathsf{Var}(T) = \frac{1-p}{p^2}, \qquad \mathsf{SD}(T)$$

Den geometriske fordeling

T : antal Bernoulli forsøg indtil første succes

$$\mathsf{P}(T = i) = (1-p)^{i-1}p$$

$$\mathsf{P}(T > i) = (1-p)^{i}$$

Middelværdi

$$\mathsf{E}(T) = \sum_{i=0}^{\infty} \mathsf{P}(T > i) = \frac{1}{p}$$

Varians

$$\mathsf{Var}(T) = \frac{1-p}{p^2}, \qquad \mathsf{SD}(T) = \frac{\sqrt{1-p}}{p}$$

T_r Antal forsøg indtil $r$te succes

Negativ binomialfordeling

- Y_r Antal fiaskoer til den r'te succes i en sekvens af Bernoulliforsøg
- Hvor mange kameraer skal vi kassere før vi får r, der passerer kvalitetskontrollen

$$T_r = Y_r + r$$

- (Har i øvrigt mange andre fortolkninger)

Spørgsmål 4

Hvad er middelværdi og varians for en $NB(r,p)$ fordeling?

1 ☐ $\mathsf{E}(Y_r) = \frac{r(1-p)}{p}, \mathsf{Var}(Y_r) = \frac{r(1-p)}{p^2}$

2 ☐ $\mathsf{E}(Y_r) = \frac{r(1-p)}{p}, \mathsf{Var}(Y_r) = \frac{(1-p)}{p^2}$

3 ☐ $\mathsf{E}(Y_r) = \frac{1}{p}, \mathsf{Var}(Y_r) = \frac{r(1-p)}{p^2}$

4 ☐ $\mathsf{E}(Y_r) = \frac{1}{p}, \mathsf{Var}(Y_r) = \frac{(1-p)}{p^2}$

5 ☐ $\mathsf{E}(Y_r) = \frac{r}{p}, \mathsf{Var}(Y_r) = \frac{r}{p^2}$

6 ☐ Ved ikke

Negativ binomialfordeling udledning

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Vi skal placere $r-1$ successer på $i+r-1$ forsøg før vores "endelige" succes.

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Vi skal placere $r-1$ successer på $i+r-1$ forsøg før vores "endelige" succes.

Derfor

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Vi skal placere $r-1$ successer på $i+r-1$ forsøg før vores "endelige" succes.

Derfor

$$\mathsf{P}(Y_r = i)$$

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Vi skal placere $r-1$ successer på $i+r-1$ forsøg før vores "endelige" succes.

Derfor

$$\mathsf{P}(Y_r = i) = \begin{pmatrix} i+r-1 \\ r-1 \end{pmatrix}$$

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Vi skal placere $r-1$ successer på $i+r-1$ forsøg før vores "endelige" succes.

Derfor

$$\mathsf{P}(Y_r = i) = \binom{i+r-1}{r-1} p^{r-1}(1-p)^i p$$

Negativ binomialfordeling udledning

Antal fiaskoer før r'te succes.

En sekvens af Bernoulli forsøg: fffsffsfs

Vi skal placere $r-1$ successer på $i+r-1$ forsøg før vores "endelige" succes.

Derfor

$$\mathsf{P}(Y_r = i) = \binom{i+r-1}{r-1} p^{r-1}(1-p)^i p$$

$$= \binom{i+r-1}{r-1} p^r(1-p)^i$$

Middelværdi og varians

$$Y_r \sim NB(r, p)$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.
W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.
W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p}$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.
W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p} \qquad \mathsf{Var}(W_i) = \frac{1-p}{p^2}$$

Vi har da

$$Y_r = \sum_{i=1}^{r} X_i = \sum_{i=1}^{r}(W_i - 1) = \sum_{i=1}^{r} W_i - r$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.
W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p} \qquad \mathsf{Var}(W_i) = \frac{1-p}{p^2}$$

Vi har da

$$Y_r = \sum_{i=1}^{r} X_i = \sum_{i=1}^{r}(W_i - 1) = \sum_{i=1}^{r} W_i - r$$

Derfor:

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.

W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p} \qquad \mathsf{Var}(W_i) = \frac{1-p}{p^2}$$

Vi har da

$$Y_r = \sum_{i=1}^{r} X_i = \sum_{i=1}^{r}(W_i - 1) = \sum_{i=1}^{r} W_i - r$$

Derfor:

$$\mathsf{E}(Y_r) = \frac{r}{p} - r$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.

W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p} \qquad \mathsf{Var}(W_i) = \frac{1-p}{p^2}$$

Vi har da

$$Y_r = \sum_{i=1}^{r} X_i = \sum_{i=1}^{r}(W_i - 1) = \sum_{i=1}^{r} W_i - r$$

Derfor:

$$\mathsf{E}(Y_r) = \frac{r}{p} - r = r\frac{1-p}{p}$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.
W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p} \qquad \mathsf{Var}(W_i) = \frac{1-p}{p^2}$$

Vi har da

$$Y_r = \sum_{i=1}^{r} X_i = \sum_{i=1}^{r} (W_i - 1) = \sum_{i=1}^{r} W_i - r$$

Derfor:

$$\mathsf{E}(Y_r) = \frac{r}{p} - r = r\frac{1-p}{p} \qquad \mathsf{Var}(Y_r)$$

Middelværdi og varians

$$Y_r \sim NB(r,p) \qquad Y_r = T_r - r \qquad T_r = W_1 + \cdots + W_r$$

hvor W_i er ventetiden fra den $i-1$te til den $i$te succes.

W_i er IID; geometrisk (p) fordelt: Lad X_i være antallet af fiaskoer mellem succes $i-1$ og i.

$$\mathsf{E}(W_i) = \frac{1}{p} \qquad \mathsf{Var}(W_i) = \frac{1-p}{p^2}$$

Vi har da

$$Y_r = \sum_{i=1}^{r} X_i = \sum_{i=1}^{r} (W_i - 1) = \sum_{i=1}^{r} W_i - r$$

Derfor:

$$\mathsf{E}(Y_r) = \frac{r}{p} - r = r\frac{1-p}{p} \qquad \mathsf{Var}(Y_r) = r\frac{1-p}{p^2}$$

Indikatorfunktionen (p.155), p.168

$$I_A = \begin{cases} 1 & \text{hvis } A \text{ indtræffer} \\ 0 & \text{hvis } A^c \text{ indtræffer} \end{cases}$$

Spørgsmål 5

Hvad er $\mathrm{E}(I_A)$

1 ☐ 0,5

2 ☐ 1

3 ☐ A

4 ☐ $\mathrm{P}(A)$

5 ☐ Kan man ikke sige generelt

6 ☐ Ved ikke

Hvornår kan man bruge indikatorfunktioner?

Hvornår kan man bruge indikatorfunktioner?

Man kan bruge indikatorfunktionen når vi kan skrive en stokastisk variabel X som

$$X = I_{A_1} + I_{A_2} + \cdots + I_{A_k}$$

Hvornår kan man bruge indikatorfunktioner?

Man kan bruge indikatorfunktionen når vi kan skrive en stokastisk variabel X som

$$X = I_{A_1} + I_{A_2} + \cdots + I_{A_k}$$

- dvs. X tæller, hvor mange af hændelsernene $A_1, \ldots, A_k$, der er indtruffet.

Hvornår kan man bruge indikatorfunktioner?

Man kan bruge indikatorfunktionen når vi kan skrive en stokastisk variabel X som

$$X = I_{A_1} + I_{A_2} + \cdots + I_{A_k}$$

- dvs. X tæller, hvor mange af hændelsernene $A_1, \ldots, A_k$, der er indtruffet.

$$\mathsf{E}(X)$$

Hvornår kan man bruge indikatorfunktioner?

Man kan bruge indikatorfunktionen når vi kan skrive en stokastisk variabel X som

$$X = I_{A_1} + I_{A_2} + \cdots + I_{A_k}$$

- dvs. X tæller, hvor mange af hændelsernene $A_1, \dots, A_k$, der er indtruffet.

$$\mathsf{E}(X) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \dots \mathsf{P}(A_k)$$

Hvornår kan man bruge indikatorfunktioner?

Man kan bruge indikatorfunktionen når vi kan skrive en stokastisk variabel X som

$$X = I_{A_1} + I_{A_2} + \cdots + I_{A_k}$$

- dvs. X tæller, hvor mange af hændelsernene $A_1, \ldots, A_k$, der er indtruffet.

$$\mathsf{E}(X) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \ldots \mathsf{P}(A_k)$$

bruges når $\mathsf{P}(A_i)$ kan bestemmes (let), men fordelingen af X ikke kan. (F.eks. opgave 3.2.14)

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ ,\quad \mathsf{P}(A_2) = \frac{1}{4}\ ,\quad \mathsf{P}(A_3) = \frac{1}{3}$$

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ , \quad \mathsf{P}(A_2) = \frac{1}{4}\ , \quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ ,\quad \mathsf{P}(A_2) = \frac{1}{4}\ ,\quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ ,\quad \mathsf{P}(A_2) = \frac{1}{4}\ ,\quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

- Variablen I_B antager værdien 1,

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ , \quad \mathsf{P}(A_2) = \frac{1}{4}\ , \quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

- Variablen I_B antager værdien 1, hvis B indtræffer,

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ , \quad \mathsf{P}(A_2) = \frac{1}{4}\ , \quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

- Variablen I_B antager værdien 1, hvis B indtræffer, ellers 0.

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ , \quad \mathsf{P}(A_2) = \frac{1}{4}\ , \quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

- Variablen I_B antager værdien 1, hvis B indtræffer, ellers 0.

$$N =$$

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ , \quad \mathsf{P}(A_2) = \frac{1}{4}\ , \quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

- Variablen I_B antager værdien 1, hvis B indtræffer, ellers 0.

$$N = I_{A_1}$$

Opgave 3.3.8

Lad A_1, A_2 og A_3 være hændelser med sandsynlighederne

$$\mathsf{P}(A_1) = \frac{1}{5}\ , \quad \mathsf{P}(A_2) = \frac{1}{4}\ , \quad \mathsf{P}(A_3) = \frac{1}{3}$$

Lad N betegne antallet af disse hændelser, der indtræffer.

Opskriv N udtrykt ved indikatorvariable

- Variablen I_B antager værdien 1, hvis B indtræffer, ellers 0.

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Find E(N)

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N)$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1} + I_{A_2} + I_{A_3})$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1} + I_{A_2} + I_{A_3})$$

Dermed

$$\mathsf{E}(N)$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1} + I_{A_2} + I_{A_3})$$

Dermed

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1}) + \mathsf{E}(I_{A_2}) + \mathsf{E}(I_{A_3}) \qquad \text{Hvorfor?}$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1} + I_{A_2} + I_{A_3})$$

Dermed

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1}) + \mathsf{E}(I_{A_2}) + \mathsf{E}(I_{A_3}) \qquad \text{Hvorfor?}$$

$$\mathsf{E}(N)$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1} + I_{A_2} + I_{A_3})$$

Dermed

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1}) + \mathsf{E}(I_{A_2}) + \mathsf{E}(I_{A_3}) \qquad \text{Hvorfor?}$$

$$\mathsf{E}(N) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3)$$

Find $\mathsf{E}(N)$

$$N = I_{A_1} + I_{A_2} + I_{A_3}$$

Vi tager forventning på begge sider

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1} + I_{A_2} + I_{A_3})$$

Dermed

$$\mathsf{E}(N) = \mathsf{E}(I_{A_1}) + \mathsf{E}(I_{A_2}) + \mathsf{E}(I_{A_3}) \qquad \text{Hvorfor?}$$

$$\mathsf{E}(N) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) = \frac{47}{60}$$

Beregn Var(N) når

- A_1, A_2 og A_3 er gensidigt udelukkende
- A_1, A_2 og A_3 er uafhængige
- $A_1 \subset A_2 \subset A_3$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$\mathsf{P}(N = 1)$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$$\mathsf{P}(N = 1) = \mathsf{P}(A_1 \cup A_2 \cup A_3)$$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$$\mathsf{P}(N = 1) = \mathsf{P}(A_1 \cup A_2 \cup A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3)$$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$$\mathsf{P}(N = 1) = \mathsf{P}(A_1 \cup A_2 \cup A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) = \mathsf{E}(N)$$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$$\mathsf{P}(N = 1) = \mathsf{P}(A_1 \cup A_2 \cup A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) = \mathsf{E}(N)$$

Eksperimentet er altså et Bernoullieksperiment og vi har

$$\mathsf{Var}(N)$$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$$\mathsf{P}(N = 1) = \mathsf{P}(A_1 \cup A_2 \cup A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) = \mathsf{E}(N)$$

Eksperimentet er altså et Bernoullieksperiment og vi har

$$\mathsf{Var}(N) = p(1-p)$$

Var(N) når A_i er gensidigt udelukkende

I dette tilfælde har vi

$$\mathsf{P}(N > 1) = 0$$

$$\mathsf{P}(N = 1) = \mathsf{P}(A_1 \cup A_2 \cup A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) = \mathsf{E}(N)$$

Eksperimentet er altså et Bernoullieksperiment og vi har

$$\mathsf{Var}(N) = p(1-p) = \frac{47}{60}\frac{13}{60} = 0,170 = 0,412^2$$

Var(N) når A_1, A_2 og A_3 er uafhængige

Var(N)

Var(N) når A_1, A_2 og A_3 er uafhængige

$$\mathsf{Var}(N) = \mathsf{Var}(I_{A_1} + I_{A_2} + I_{A_3})$$

Var(N) når A_1, A_2 og A_3 er uafhængige

$$\mathsf{Var}(N) = \mathsf{Var}(I_{A_1} + I_{A_2} + I_{A_3}) = \mathsf{Var}(I_{A_1}) + \mathsf{Var}(I_{A_3}) + \mathsf{Var}(I_{A_3})$$

Var(N) når A_1, A_2 og A_3 er uafhængige

$$\mathsf{Var}(N) = \mathsf{Var}(I_{A_1} + I_{A_2} + I_{A_3}) = \mathsf{Var}(I_{A_1}) + \mathsf{Var}(I_{A_3}) + \mathsf{Var}(I_{A_3})$$

For hver A_i har vi

$$\mathsf{Var}(I_{A_i}) = \mathsf{P}(A_i) \cdot \mathsf{P}(A_i^c)$$

Var(N) når A_1, A_2 og A_3 er uafhængige

$$\mathsf{Var}(N) = \mathsf{Var}(I_{A_1} + I_{A_2} + I_{A_3}) = \mathsf{Var}(I_{A_1}) + \mathsf{Var}(I_{A_3}) + \mathsf{Var}(I_{A_3})$$

For hver A_i har vi

$$\mathsf{Var}(I_{A_i}) = \mathsf{P}(A_i) \cdot \mathsf{P}(A_i^c)$$

Altså

$$\mathsf{Var}(I_{A_1}) = \frac{1}{5}\frac{4}{5} = \frac{4}{25} \qquad \mathsf{Var}(I_{A_2}) = \frac{3}{16} \qquad \mathsf{Var}(I_{A_3}) = \frac{2}{9}$$

Var(N) når A_1, A_2 og A_3 er uafhængige

$$\mathsf{Var}(N) = \mathsf{Var}(I_{A_1} + I_{A_2} + I_{A_3}) = \mathsf{Var}(I_{A_1}) + \mathsf{Var}(I_{A_3}) + \mathsf{Var}(I_{A_3})$$

For hver A_i har vi

$$\mathsf{Var}(I_{A_i}) = \mathsf{P}(A_i) \cdot \mathsf{P}(A_i^c)$$

Altså

$$\mathsf{Var}(I_{A_1}) = \frac{1}{5}\frac{4}{5} = \frac{4}{25} \qquad \mathsf{Var}(I_{A_2}) = \frac{3}{16} \qquad \mathsf{Var}(I_{A_3}) = \frac{2}{9}$$

Så

$$V(N) = \frac{2051}{3600} \approx 0,570 \approx 0.755^2$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N = 3)$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N = 3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N = 3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N = 3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N = 2)$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1)$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \setminus A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1)$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \backslash A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \backslash A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Dermed:

$$\mathsf{E}(N^2)$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \setminus A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \setminus A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Dermed:

$$\mathsf{E}(N^2) = \frac{9}{5} + \frac{4}{20} + \frac{1}{12}$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \backslash A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Dermed:

$$\mathsf{E}(N^2) = \frac{9}{5} + \frac{4}{20} + \frac{1}{12} = \frac{119}{60}$$

Beregn $\mathsf{Var}(N)$ når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \backslash A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Dermed:

$$\mathsf{E}(N^2) = \frac{9}{5} + \frac{4}{20} + \frac{1}{12} = \frac{119}{60}$$

$$\mathsf{Var}(N)$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \backslash A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Dermed:

$$\mathsf{E}(N^2) = \frac{9}{5} + \frac{4}{20} + \frac{1}{12} = \frac{119}{60}$$

$$\mathsf{Var}(N) = \mathsf{E}(N^2) - (\mathsf{E}(N))^2$$

Beregn Var(N) når $A_1 \subset A_2 \subset A_3$

$$\mathsf{P}(N=3) = \mathsf{P}(A_1) = \frac{1}{5}$$

Vi har ved brug af differensreglen p.22

$$\mathsf{P}(N=2) = \mathsf{P}(A_2 \backslash A_1) = \mathsf{P}(A_2) - \mathsf{P}(A_1) = \frac{1}{20}$$

$$\mathsf{P}(N=1) = \mathsf{P}(A_3 \backslash A_2) = \mathsf{P}(A_3) - \mathsf{P}(A_2) = \frac{1}{12}$$

Dermed:

$$\mathsf{E}(N^2) = \frac{9}{5} + \frac{4}{20} + \frac{1}{12} = \frac{119}{60}$$

$$\mathsf{Var}(N) = \mathsf{E}(N^2) - (\mathsf{E}(N))^2 = \frac{4931}{3600} = 1,370 = 1,170^2$$

Beregn Var(N) når

A_1, A_2 og A_3 er gensidigt udelukkende Var$(N) = 0.412^2$

A_1, A_2 og A_3 er uafhængige Var$(N) = 0,755^2$

$A_1 \subset A_2 \subset A_3$ Var$(N) = 1,17^2$

Kan vi forklare/forstå resultatet?

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X$$

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X = \sum_{i=1}^{X} 1$$

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X = \sum_{i=1}^{X} 1 = \sum_{i=1}^{\infty} I_{X \geq i}$$

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X = \sum_{i=1}^{X} 1 = \sum_{i=1}^{\infty} I_{X \geq i}$$

$$\mathsf{E}(X)$$

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X = \sum_{i=1}^{X} 1 = \sum_{i=1}^{\infty} I_{X \geq i}$$

$$\mathsf{E}(X) = \mathsf{E}\left(\sum_{i=1}^{\infty} I_{X \geq i}\right) =$$

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X = \sum_{i=1}^{X} 1 = \sum_{i=1}^{\infty} I_{X \geq i}$$

$$\mathsf{E}(X) = \mathsf{E}\left(\sum_{i=1}^{\infty} I_{X \geq i}\right) = \sum_{i=1}^{\infty} \mathsf{E}\left(I_{X \geq i}\right)$$

Tail sum formula

Hvis X kan antage værdier $0, 1, \ldots$

$$\mathsf{E}(X) = \sum_{i=1}^{\infty} \mathsf{P}(X \geq i)$$

(En tilsvarende formel gælder for kontinuerte variable; kap. 4)

Bevis:

$$X = \sum_{i=1}^{X} 1 = \sum_{i=1}^{\infty} I_{X \geq i}$$

$$\mathsf{E}(X) = \mathsf{E}\left(\sum_{i=1}^{\infty} I_{X \geq i}\right) = \sum_{i=1}^{\infty} \mathsf{E}\left(I_{X \geq i}\right) = \sum_{i=1}^{\infty} \mathsf{P}\left(X \geq i\right)$$

Et par nyttige sumformler

$$\sum_{i=0}^{N} a^i = \frac{1-a^{N+1}}{1-a}, \qquad |a| \neq 1$$

$$\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}, \qquad |a| < 1$$

$$\sum_{i=0}^{\infty} \frac{a^i}{i!} = e^a$$

Et par nyttige sumformler

$$\sum_{i=0}^{N} a^i = \frac{1-a^{N+1}}{1-a}, \qquad |a| \neq 1$$

$$\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}, \qquad |a| < 1$$

$$\sum_{i=0}^{\infty} \frac{a^i}{i!} = e^a$$

- For $a_i \geq 0$ og $b_i \geq 0$ har vi:

$$\left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{i=0}^{\infty} b_i\right) = \sum_{i=0}^{\infty}\left(\sum_{k=0}^{i} a_k b_{i-k}\right) = \sum_{i=0}^{\infty} c_i \qquad c_i = \sum_{k=0}^{i} a_k b_{i-k}$$

Sammenhæng mellem fordelinger

- $X \sim \text{Bin}(n_1, p)$ og $Y \sim \text{Bin}(n_2, p)$: $X + Y \sim \text{Bin}(n_1 + n_2, p)$
- $X \sim \text{Geom}(p)$: $X - 1 \sim \text{NB}(1, p)$.
- $X \sim \text{NB}(r_1, p)$ og $Y \sim \text{NB}(r_2, p)$: $X + Y \sim \text{NB}(r_1 + r_2, p)$
- $\text{HyperGeom}(n, N, G) \to \text{Bin}(n, p)$ når $N, G \to \infty$ mens $G/N \to p$.
- Summer kan generelt approksimeres med normalfordelingen

Nye begreber i afsnit 3.3 og 3.4

- Varians og standardafvigelse
- Markovs og Chebychevs uligheder
- Indikatorfunktion
- Den centrale grænseværdisætning (3.3)
- Geometrisk fordeling
- Negativ binomial fordeling

Afsnit 3.3 og 3.4

- Varians/standardafvigelse
$$\mathsf{Var}(X) = \mathsf{E}[(X - \mathsf{E}(X))^2] = \sum_{\text{alle } x} (x - \mathsf{E}(X))^2 \mathsf{P}(X = x)$$

- Normalfordelingsapproksimation/den Centrale Grænseværdisætning
$$\mathsf{P}\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right) \approx \Phi(b) - \Phi(a)$$

- Markovs og Chebychevs uligheder for ekstreme udfald
$$\mathsf{P}(X \ge a) \le \frac{\mathsf{E}(X)}{a} \qquad \mathsf{P}(|X - \mathsf{E}(X)| \ge k\mathsf{SD}(X)) \le \frac{1}{k^2}$$

- Et udpluk af diskrete fordelinger
$$\mathsf{P}(T = i) = (1-p)^{i-1}p \qquad \mathsf{P}(T_r = i) = \binom{x + r - 1}{r - 1} p^{r-1}(1-p)^i p$$

- Indikatorfunktioner $I_A = \begin{cases} 1 & \text{hvis } A \text{ indtræffer} \\ 0 & \text{hvis } A^c \text{ indtræffer} \end{cases}$