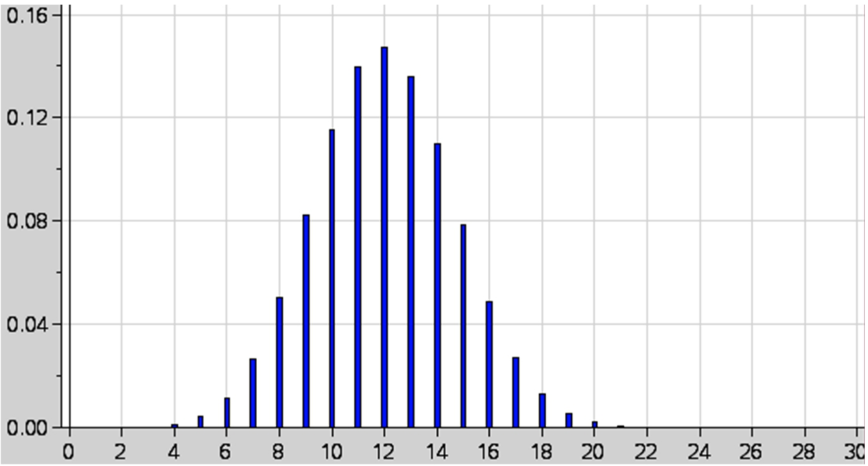
Sandsynlighed

for B-niveau i stx og hf



2020 Karsten Juul

1. Kombinatorik
1.1 Hvad er kombinatorik? ... 1
1.2 Multiplikations-princippet ... 1
1.3. Vælge r ting blandt n ting ... 1
1.4 Sammensat opgave ... 2

2. Stokastisk variabel
2.1 Eksempel på stokastisk variabel ... 3
2.2 Middelværdi ... 4
2.3 Spredning ... 6

3. Symmetrisk sandsynlighedsfelt
3.1 Hvad er et symmetrisk sandsynlighedsfelt ... 7
3.2 Regel for sandsynlighed i et symmetrisk sandsynlighedsfelt ... 7
3.3 Eksempel 1 på opgave om symmetrisk sandsynlighedsfelt ... 8
3.4 Eksempel 1 på opgave om symmetrisk sandsynlighedsfelt ... 8

4. Binomialfordeling
4.1 Hvad er en binomialfordeling? ... 9
4.2 Eksempel på binomialfordeling ... 10
4.3 Find sandsynligheden for ét udfald i en binomialfordeling ... 10
4.4 Find sandsynligheden for flere udfald i en binomialfordeling ... 10

5. Middelværdi og spredning for binomialfordeling
5.1 Formel for middelværdi for binomialfordeling ... 11
5.2 Mest sandsynlige udfald ... 11
5.3 Formel for spredning for binomialfordeling ... 11
5.4 Exceptionelle værdier og normale værdier ... 12

6. Formel for binomialfordeling
6.1 Fakultet ... 13
6.2 Formel for $K(n, r)$ ... 13
6.3 Formel for $P(X = r)$ hvor $X \sim b(n,p)$ ... 13
6.4 Eksempel på opgave med 6.3 ... 13
6.5 Eksempel 1 på opgave med 6.2 og 6.3 ... 14
6.6 Eksempel 2 på opgave med 6.2 og 6.3 ... 14

7. Konfidensinterval
7.1 Sådan finder man konfidensintervallet ... 15
7.2 Eksempel 1 på opgave med konfidensinterval ... 15
7.3 Eksempel 2 på opgave med konfidensinterval ... 16

8. Tosidet binomialtest
8.1 Hypotese ... 17
8.2 Kritisk område ... 17
8.3 Brug af det kritiske område ... 16

9. Ensidet binomialtest
9.1 Højresidet binomialtest 18
9.2 Venstresidet binomialtest 18

10. Normalfordelingsapproksimation
10.1 Tal der er normalfordelt 19
10.2 Frekvensfunktion for normalfordeling 19
10.3 Beregning af sandsynlighed i normalfordeling 19
10.4 Normalfordelingsapproksimation til binomialfordeling 19

11. Stolpediagram for binomialfordeling 20

1. Kombinatorik

1.1 Hvad er kombinatorik?

Kombinatorik er en del af matematikken.
Den består af metoder til at besvare spørgsmål af typen:
På hvor mange måder kan dette gøres?

Eksempler:

I to klasser er der henhodsvis 15 og 18 elever.

Der skal vælges én elev fra hver klasse.
På hvor mange måder kan dette gøres?

Der skal vælges 3 elever fra den første klasse.
På hvor mange måder kan dette gøres?

1.2 Multiplikations-princippet

I et spil skal man både vælge et af bogstaverne A, B, C og D,
og vælge et af tallene 2, 4 og 6.

Der er 4 måder at vælge et bogstav på.

Det første bogstav kan sættes sammen med hvert af de 3 tal.

Det andet bogstav kan også sættes sammen med hvert af de 3 tal.

Osv.

Det giver $4 \cdot 3 = 12$ muligheder for at vælge bogstav og tal.

	A2	1
1	**A4**	2
	A6	3
	B2	1
2	**B4**	2
	B6	3
	C2	1
3	**C4**	2
	C6	3
	D2	1
4	**D4**	2
	D6	3

Regel

Når et valg består af to delvalg der begge skal udføres, så gælder:

antal valgmuligheder

er

antal muligheder i første delvalg   gange   antal muligheder i andet delvalg .

1.3 Vælge r ting blandt n ting

Hvis vi skal vælge to bogstaver blandt de tre bogstaver **A** , **B** , **C** ,
så er der følgende muligheder:

AB
AC
BC

Antal måder er altså 3 .

Antal måder hvorpå 2 ting kan vælges blandt 3, kan skrives sådan:

$\boldsymbol{K(3, 2)}$

Der gælder altså at

$K(3, 2) = 3$.

Antal måder hvorpå 3 ting kan vælges blandt 7 er $K(7, 3)$.

1.3a Bestemme $K(7, 3)$ med formelsamling

Find ordet Pascals trekant i stikordsregisteret bag i formelsamlingen, og slå op på den relevante side.
I den øverste trekant finder vi symbolet $K(7, 3)$.
På den tilsvarende plads i den nederst trekant står tallet 35 .
Det betyder at

$$K(7, 3) = 35 \ .$$

1.3b Bestemme $K(7, 3)$ med Nspire

I et matematikfelt taster vi nCr(7, 3) :

$$\mathrm{nCr}(7,3) = 35$$

Dvs. $K(7, 3) = 35$.

1.4 Sammensat opgave

Opgaven

I et hold på 13 soldater er der 7 i grønne uniformer og 6 i brune uniformer.
Bestem hvor mange forskellige grupper bestående af 2 i grønne uniformer og 3 i brune uniformer, der kan dannes af de 13 på holdet.

Udregningen

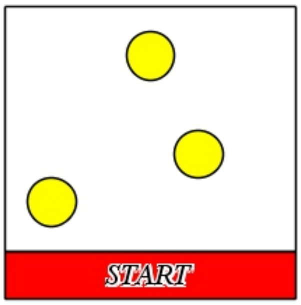
Valg 1: Der skal vælges 2 blandt de 7 grønne. Antal måder: $K(7, 2) = 21$.
Valg 2: Der skal vælges 3 blandt de 6 i brune. Antal måder: $K(6, 3) = 20$.

For at vælge gruppen skal både foretages valg 1 og valg 2, så antal måder gruppen kan vælges på er $21 \cdot 20 = \underline{\underline{420}}$. Se 1.2 .

Tallene $K(7, 2)$ og $K(6, 3)$ blev bestemt ved hjælp af Pascals trekant i formelsamlingen.

2. Stokastisk variabel

2.1 Eksempel på stokastisk variabel



Når man klikker på start-knappen, fremkommer et antal cirkler.

Antallet er tilfældigt.

Tabellen viser sandsynlighederne for de forskellige antal cirkler.

Antal cirkler	1	2	3	4
Sandsynlighed	0,2	0,4	0,1	0,3

Antal cirkler er en **stokastisk variabel** fordi hvert antal forekommer med en bestemt sandsynlighed.

En stokastisk variabel betegner vi ofte med X .

I tabellen ser vi at

sandsynligheden for at antal er 2 er 0,4 .

Dette kan vi skrive sådan:

$P(X = 2) = 0{,}4$.

2.1a Summen af sandsynlighederne skal altid være 1:

$0{,}2 + 0{,}4 + 0{,}1 + 0{,}3 = 1$.

Hvis sandsynlighederne angives i procent, gælder:

Summen af sandsynlighederne skal altid være **100 %** .

2.1b Eksempel på opave

Opgaven:

En stokastisk variabel X har følgende sandsynlighedestabel:

t :	5	10	15	20	25
$P(X = t)$:	0,3	0,1	0,1	p	0,2

Bestem den manglende sandsynlighed p .

Udregningen:

Summen af sandsynlighederne skal være 1 (se 2.1a):

$0{,}3 + 0{,}1 + 0{,}1 + p + 0{,}2 = 1$

$0{,}7 + p = 1$

$p = 0{,}3$

2.1c Eksempel på opave

Opgaven:

I et spil får man et vist antal stjerner. Den stokastiske variabel X angiver antal stjerner.

En del af sandsynlighedestabellen hørende til X ser sådan ud:

Antal stjerner:	0	1	2	3	4	. . .
Sandsynlighed:	5 %	7 %	6 %	9 %	14 %	. . .

a) Bestem sandsynligheden for at få mindst en stjerne.

b) Bestem $P(1 \leq X \leq 3)$, og fortolk resultatet.

Udregninger:

a) $100\,\% - 5\,\% = 95\,\%$ ifølge 2.1a.

Sandsynligheden for at få mindst en stjerne er 95 %.

b) $7\,\% + 6\,\% + 9\,\% = 22\,\%$
dvs. i 22 % af de gange man spiller, får man 1, 2 eller 3 stjerner.

2.2 Middelværdi

2.2a Sådan udregnes middelværdi

Sandsynlighedsfordeling for stokastisk variabel X:

x_i :	x_1	x_2	. . .	x_n
$P(X = x_i)$:	p_1	p_2	. . .	p_n

Middelværdien af den stokastiske variabel er

$$\mu = E(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \ . \ . \ . \ + x_n \cdot p_n$$

2.2b Eksempel på udregning af middelværdi

Hvis det er sandsynlighedsfordelingen fra 2.1, dvs.

Antal cirkler :	1	2	3	4
Sandsynlighed :	0,2	0,4	0,1	0,3

Så er middelværdien

$$\mu = E(X) = 1 \cdot 0{,}2 + 2 \cdot 0{,}4 + 3 \cdot 0{,}1 + 4 \cdot 0{,}3 = 2{,}5$$

Dvs: hvis vi mange gange klikker på start, så vil det **gennemsnitlige antal** cirkler være 2,5 .

2.2c Eksempel på opgave

Opgaven:

I et spil havner en kugle i et af tre felter.

I tabellen er vist gevinst og sandsynlighed for hvert felt.

	felt 1	felt 2	felt 3
Gevinst (kr.)	–5	2	3
Sandsynlighed	0,6	0,3	0,1

a) Bestem den gennemsnitlige gevinst pr. spil.

Spillets ejer vil ændre gevinsten i felt 3 så hun gennemsnitlig tjener 2 kr. pr. spil.

b) Bestem den gevinststørrelse x i felt 3 som vil bevirke at ejeren gennemsnitlig tjener 2 kr pr. spil.

Udregninger:

a) $\mu = (-5)\cdot 0{,}6 + 2\cdot 0{,}3 + 3\cdot 0{,}1 = -2{,}1$

Den gennemsnitlige gevinst pr. spil er $-2{,}1$ kr.

Dvs. man taber gennemsnitlig 2,1 kr. pr. spil,
og spillets ejer tjener gennemsnitlig 2,1 kr. pr. spil.

b) $(-5)\cdot 0{,}6 + 2\cdot 0{,}3 + x\cdot 0{,}1 = -2$

$-3 + 0{,}6 + 0{,}1x = -2$

$0{,}1x = 0{,}4$

$x = 4$

Hvis gevinststørrelsen i felt 3 ændres til 4 kr. ,
så vil ejeren gennemsnitlig tjene 2 kr. pr. spil.

2.2d Bestemme middelværdi med Nspire.

Tast den stokastiske variabels tabel i regnearket som vist.

	A antal	B sandsynlighed
=		
1	1	0.2
2	2	0.4
3	3	0.1
4	4	0.3
5		

Vælg i værktøjsmenuen:

Statistik / Statistiske beregninger / Statistik med én variabel /

Sæt Antal lister til 1 / Vælg som X1-liste søjlen **antal** (eller hvad du nu har kaldt den) /

Vælg som frekvensliste søjlen **Sandsynlighed** (eller hvad du nu har kaldt den) /

Klik på OK .

Middelværdien står ud for symbolet $\overline{x}$.

x̄	2.5

2.3 Spredning

Når eksperimentet hørende til en stokastisk variabel udføres flere gange, så fås en række tal. Disse tal kan ligge mere eller mindre spredt.
Man udregner et tal man kalder spredningen for den stokastiske variabel.
Dette tal er et udtryk for hvor langt tallene gennemsnitligt ligger fra middeltallet.

2.3a Bestemme spredning med Nspire.

Tast den stokastiske variabels tabel i regnearket som vist.

	A antal	B sandsynlighed
=		
1	1	0.2
2	2	0.4
3	3	0.1
4	4	0.3
5		

Vælg i værktøjsmenuen:

Statistik / Statistiske beregninger / Statistik med én variabel /

Sæt Antal lister til 1 / Vælg som X1-liste søjlen **antal** (eller hvad du nu har kaldt den) /

Vælg som frekvensliste søjlen **Sandsynlighed** (eller hvad du nu har kaldt den) /

Klik på OK .

Spredningen står ud for symbolet $\sigma x{:=}\sigma_n x$.

$\sigma x := \sigma_n x$	1.11803

3: Symmetrisk sandsynlighesfelt

3.1 Hvad er et symmetrisk sandsynlighedsfelt

I et spil på skærmen får man bogstaver i forskellige farver.
Sandsynlighedsfeltet ser sådan ud:

Udfald :	A	A	A	B	B	C	C	C	C	D
Sandsynlighed :	0,1	0,05	0,03	0,1	0,1	0,09	0,07	0,15	0,01	0,1

Lad H være hædelsen at vi får et grønt bogstav.
Sandsynligheden $P(H)$ for denne hændelse kan udregnes sådan:

$$P(H) = 0{,}05 + 0{,}1 + 0{,}07 = 0{,}22$$

I et andet spil ser sandsynlighedfeltet sådan ud:

Udfald :	A	A	A	B	B	C	C	C	C	D
Sandsynlighed :	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1

Lad H være hædelsen at vi får et grønt bogstav.
Sandsynligheden $P(H)$ for denne hændelse kan udregnes sådan:

$$P(H) = 0{,}1 + 0{,}1 + 0{,}1 = 0{,}3$$

I dette tilfælde kan sandsynligheden også udregnes sådan:

$$P(H) = \frac{3}{10} = 0{,}3$$

Dette skyldes at alle udfald har samme sandsynlighed.
Så kan vi blot dividere antallet af udfald i hændelsen med antallet af alle udfald.

De udfald vi skal finde sandsynligheden for, kaldes **gunstige udfald** ,
også selv om disse udfald er noget ubehageligt.

Et sandsynlighedsfelt kaldes **symmetrisk** hvis alle udfald har samme sandsynlighed.

3.2 Regel for sandsynlighed i et symmetrisk sandsynlighedsfelt

I et symmetrisk sandsynlighedsfelt udregnes sandsynligheden for en hændelse H sådan:

$$P(H) = \frac{antal\ gunstige}{antal\ mulige} .$$

3.3 Eksempel 1 på opgave om symmetrisk sandsynlighedsfelt

Opgaven:

En gruppe består 4 danskere og 4 svenskere.
Tilfældigt udtages 3.
Bestem sandsynligheden for at alle udtagne er danskere.

Udregningen:

Gruppen har $4+4 = 8$ medlemmer.
Antal måder hvorpå 3 personer kan vælges blandt 8 er $K(8, 3) = 56$.
Antal måder hvorpå 3 personer kan vælges blandt 4 (danskere) er $K(4, 3) = 4$.
Sandsynligheden for alle udtagne er danskere, er

$$\frac{antal\ gunstige}{antal\ mulige} = \frac{4}{56} = 0{,}071429$$

dvs. $7{,}1\,\%$.

3.4 Eksempel 2 på opgave om symmetrisk sandsynlighedsfelt

Opgaven:

I et skab står 30 ølglas.
12 af dem er en anelse højere end de andre 18 .
8 tilfældigt valgte af glassene stilles på et bord.
Bestem sandsynligheden for at 3 af dem er de høje og de andre 5 er de lave.

Udregningen:

Antal måder hvorpå 3 glas kan vælges blandt 12 er

$K(12, 3) = \text{nCr}(12,3) = 220$ udregnet af Nspire .

Antal måder hvorpå 5 glas kan vælges blandt 18 er

$K(18, 5) = \text{nCr}(18,5) = 8568$ udregnet af Nspire .

Antal måder hvorpå 8 glas kan vælges så 3 er høje og 5 er lave er

$$220 \cdot 8568 = 1884960$$

(da hver af de 220 muligheder kan sættes sammen med hver af de 8568 muligheder).

Antal måder hvorpå 8 kan vælges blandt 30 er

$K(30,8) = \text{nCr}(30,8) = 5852925$ udregnet af Nspire .

Sandsynligheden for at 3 af dem er de høje og de andre 5 er de lave, er

$$\frac{antal\ gunstige}{antal\ mulige} = \frac{K(12,3)\cdot K(18,5)}{K(30,8)} = 0{,}322054$$

dvs. $32{,}2\,\%$.

4: Binomialfordeling

4.1 Hvad er en binomialfordeling

Der står et bogstav på hver side af en skæv firesidet terning.

Når man kaster terningen mange gange, får man B i 20 % af kastene,
dvs. sandsynligheden for B er $p = 0{,}20$.

Hvis antal gange vi kaster er $n = 15$, så er antal gange vi får B en stokastisk variabel X med nedenstående sandsynlighedstabel (blå):

Sandsynlighed for B ved et kast : $p = 0{,}20$
Antal kast: $n = 15$
X = antal kast der giver B

t :	0	1	2	3	4	5	6	⋯	15
$P(X=t)$:	0,04	0,13	0,23	0,25	0,19	0,10	0,04	⋯	0,00

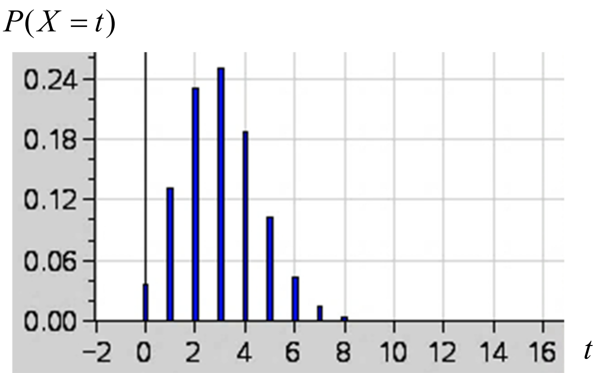
I tabellen ser vi f.eks. at:

Sandsynligheden for at det er to af kastene der giver B, er 0,23

Dvs.

Hvis vi mange gange kaster 15 gange, så vil der i 23 % af tilfældene være 2 kast der giver B.

Nedenfor er et stolpediagram der viser denne sandsynlighedstabel (sandsynlighedfordeling).
Vi ser at stolperne fra 9 til 15 er så må at man ikke kan se dem.



Sandsynlighedsfordeling ovenfor er en

binomialfordeling
med **antalsparameter** $n = 15$
og **sandsynlighedsparameter** $p = 0{,}20$

fordi

Vi udfører samme eksperiment (basiseksperimentet) 15 gange
og hver gang er sandsynligheden for at få B (basishændelsen)
lig 0,20 (basissandsynligheden).

Skrivemåde: At

X er binomialfordelt med $n = 15$ og $p = 0{,}20$

skrives kort sådan:

$X \sim b(15\,,\,0{,}20)$

4.2 Eksempel på binomialfordeling.

25 % af de der bruger en bestemt side på internettet er gymnasieelever.
Man har en liste med 40 tilfældigt valgte blandt brugerne af siden.
Lad X være antallet af gymnasieelever på listen.
Hver gang vi vælger 40 brugere, får X en en værdi.

Det viser sig at $P(X=10) = 0{,}14$, dvs. at

i 14 % af listerne med 40 personer gælder at

præcis 10 af personerne er gymnasieelever.

Begrundelse for at den stokastiske variabel X er binomialfordelt med antalsparameter $n=40$ og sandsynlighedsparameter $p=0.25$:

Basiseksperiment: En bruger undersøges for om vedkommende er gymnasieelev.

Basishændelse: Eleven er gymnasieelev.

Basissandsynlighed: Sandsynligheden for at en tilfældig bruger er gymnasieelev, dvs. 25 %.

Antal udførelser: Basiseksperimentet udføres for hver person på listen, dvs. 40 gange.

4.3 Find sandsynligheden for **ét udfald** i en binomialfordeling:

For at finde $P(X=4)$ for den stokastiske variabel fra 4.1 vælger vi i værktøjsmenuen:

Beregninger / Sandsynlighedsregning / Fordelinger / Binomial Pdf

og sætter

n til 15 , p til 0,20 og X til 4 .

Vi får $P(X=4) = 0{,}187604$.

4.4 Find sandsynligheden for **flere udfald** i en binomialfordeling:

For at finde $P(3 \leq X)$ for den stokastiske variabel fra 4.1 vælger vi i værktøjsmenuen:

Beregninger / Sandsynlighedsregning / Fordelinger / Binomial Cdf

og sætter

n til 15 , p til 0,20 , Nedre grænse til 3 og Øvre grænse til 15 .

Vi får $P(3 \leq X) = 0{,}601977$.

5: Middelværdi og spredning for binomialfordeling

5.1 Formel for middelværdi for binomialfordeling

Når

$X \sim b(n,p)$ dvs. X er binomialfordelt med antalsparameter n og sandsynlighedsparameter p

er

$E(X) = n \cdot p$ dvs. middelværdien er lig antalsparameter gange sandsynlighedsparameter

5.2 Mest sandsynlige udfald

I tabellen ser vi at sandsynligheden for 2 er 0,311 og at alle andre værdier af X har mindre sandsynlighed.

t:	0	1	2	3	4	5	6	7	8
$P(X=t)$:	0,100	0,267	0,311	0,208	0,087	0,023	0,004	0,000	0,000

Regel:

5.2a Hvis middelværdien er et helt tal:
Mest sandsynlige udfald er middelværdien.

5.2b Hvis middelværdien ikke er et helt tal:
Mest sandsynlige udfald er en af (eller begge) de heltallige naboer til middelværdien.

Eksempel på 5.2a

For $X \sim b(8, 0{,}25)$ er $E(X) = 8 \cdot 0{,}25 = 2$.

Middelværdien 2 er et helt tal, så 2 er den mest sandsynlige af de værdier X kan antage.

Eksempel på 5.2b

For $X \sim b(8,\ 0{,}2)$ er $E(X) = 8 \cdot 0{,}2 = 1{,}6$.

De heltallige naboer til middelværdien er altså 1 og 2 .

Disses sandsynligheder er

$P(X{=}1) = 0{,}336$ og $P(X{=}2) = 0{,}294$.

Altså er tallet 1 den mest sandsynlige af de værdier X kan antage.

Dette viser at det af tallene der er nærmest middelværdien IKKE behøver være den mest sandsynlige værdi.

5.3 Formel for spredning for binomialfordeling

Når

$X \sim b(n,p)$ dvs. X er binomialfordelt med antalsparameter n og sandsynlighedsparameter p

er

$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)}$ formlen for spredningen σ

5.3a Eksempel 1 på opgave med spredning

Opgaven delprøve 2

En stokastisk variabel X er binomialfordelt med antalsparameter $n = 40$ og sandsynlighedsparameter $p = 0{,}35$. Bestem spredningen σ .

Udregningen

$\sigma(X) = \sqrt{40 \cdot 0.35 \cdot (1-0.35)} = 3.01662 \approx 3.02$

5.3b Eksempel 2 på opgave med spredning

Opgaven delprøve 1

En stokastisk variabel X er binomialfordelt med antalsparameter $n = 100$ og sandsynlighedsparameter $p = 0{,}2$. Bestem spredningen $\sigma(X)$.

Udregningen

$$\sigma(X) = \sqrt{100 \cdot 0{,}2 \cdot (1-0{,}2)} = \sqrt{20 \cdot 0{,}8} = \sqrt{16} = 4$$

5.3c Eksempel 3 på opgave med spredning

Opgaven delprøve 1

En stokastisk variabel X er binomialfordelt med antalsparameter $n = 200$ og sandsynlighedsparameter $p = 0{,}2$. Undersøg om spredningen $\sigma(X)$ er mindre end 6 .

Udregningen

$$\sigma(X) = \sqrt{200 \cdot 0{,}2 \cdot (1-0{,}2)} = \sqrt{40 \cdot 0{,}8} = \sqrt{32} < \sqrt{36} = 6$$

Spredningen er mindre end 6.

5.4 Exceptionelle værdier og normale værdier

En værdi som en stokastisk variabel kan antage, er **exceptionel** hvis den afviger fra middelværdien med mere end 3 gange spredningen.

En værdi som en stokastisk variabel kan antage, er **normal** hvis den afviger fra middelværdien med 2 gange spredningen eller mindre.

5.4a Eksempel på opgave med exceptionelle og normale værdier

Opgaven delprøve 2

En stokastisk variabel X er binomialfordelt med antalsparameter $n = 48$ og sandsynlighedsparameter $p = \frac{1}{4}$.

Bestem de værdier af X der er normale, og de værdier der er exceptionelle

Udregningen

Middelværdien er $\mu = 48 \cdot \frac{1}{4} = 12$.

Spredningen er $\sigma = \sqrt{48 \cdot \frac{1}{4} \cdot \left(1 - \frac{1}{4}\right)} = \sqrt{12 \cdot \frac{3}{4}} = \sqrt{3 \cdot 3} = \sqrt{9} = 3$.

$$\mu - 2 \cdot \sigma = 12 - 2 \cdot 3 = 12 - 6 = \mathbf{6}$$
$$\mu + 2 \cdot \sigma = 12 + 2 \cdot 3 = 12 + 6 = \mathbf{18}$$

De **normale værdier** er tallene

6, 7, 8, · · · 17, 18 .

$$\mu - 3 \cdot \sigma = 12 - 3 \cdot 3 = 12 - 9 = \mathbf{3}$$
$$\mu + 3 \cdot \sigma = 12 + 3 \cdot 3 = 12 + 9 = \mathbf{21}$$

De **exceptionelle værdier** er tallene

0, 1, 2 samt 22, 23, 24, · · · 47, 48 .

6: Formel for binomialfordeling

6.1 Fakultet

Symbolet $7!$ betegner det tal man får når man

ganger de 7 første hele positive tal,

dvs.

$7! = 1{\cdot}2{\cdot}3{\cdot}4{\cdot}5{\cdot}6{\cdot}7 = 5040$.

I stedet for 7 kan man skrive ethvert positivt helt tal, f.eks.

$4! = 1{\cdot}2{\cdot}3{\cdot}4 = 24$.

Tallet $7!$ læses **7 fakultet** .

Nspire kan udregne tal som $7!$.

I formelsamlimgen står følgende formel:

Fakultet $\quad n! = n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot2\cdot1$

6.2 Formel for $K(n,r)$

Blandt 5 ting udtager vi 3 .

Antal måder det kan gøres på betegnes $K(5,3)$.

6.2a Tallet $K(5,3)$ kan vi finde i Pascals trekant i formelsamlingen. Vi ser at $K(5,3) = 10$.

6.2b Tallet $K(5,3)$ kan udregnes sådan:

$$K(5,3) = \frac{5!}{3!\cdot(5-3)!} = \frac{5!}{3!\cdot2!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\;\cdot\;2\cdot1} = \frac{5\cdot4}{2\cdot1} = 5\cdot2 = 10$$

6.2c Tallet $K(5,3)$ kan i Nspire udregnes sådan:

nCr(5,3) = 10

I formelsamlimgen står følgende formel:

Kombinationer $\quad K(n,r) = \dfrac{n!}{r!\cdot(n-r)!}$

6.3 Formel for $P(X=r)$ hvor $X \sim b(n,p)$

I formelsamlimgen står følgende formel:

$X \sim b(n,p)$

$P(X=r) = K(n,r)\cdot p^r\cdot(1-p)^{n-r}$

6.4 Eksempel på opgave med 6.3

Opgave delprøve 1

For en binomialfordelt stokastisk variabel X er sandsynligheden for $X=4$ givet ved

$$P(X=4) = K(18,4)\cdot0{,}3^4\cdot0{,}7^{14} .$$

Angiv sandsynlighedsparameteren og antalsparameteren for X .

Svar

Sandsynlighedsparameteren er $p = 0{,}3$ og antalsparameteren er $n = 18$.

6.5 Eksempel 1 på opgave med 6.2 og 6.3

Opgaven delprøve 1

Der udtages en stikprøve blandt eleverne på et gymnasium.
Den binomialfordelte stokastiske variabel X betegner antallet af elever i stikprøven som har matematik på B-niveau.
Det oplyses at sandsynligheden for at der er præcis 9 elever i stikprøven som har matematik på B-niveau, kan beregnes ved

$$P(X=9) \;=\; \frac{30!}{9!\cdot 21!}\cdot 0{,}27^{9}\cdot 0{,}73^{21}\,.$$

Forklar betydningen af tallene 30 og 0,27 i formlen.

Svar

Tallet 30 er antallet af elever i stikprøven (antalsparameteren n).
Tallet 0,27 (sandsynlighedsparameteren p) betyder at 27 % af skolens elever har matematik på B-niveau.

6.6 Eksempel 2 på opgave med 6.2 og 6.3

Opgaven delprøve 1

Antal røde klodser X i en pakke er binomialfordelt med antalsparameter 200 og sandsynlighedsparameter 0,1 .
Opstil et regneudtryk til beregning af sandsynligheden for at der er præcis 20 røde klodser i en pakke.

Svar

$$P(X=20) \;=\; \frac{200!}{20!\cdot 180!}\cdot 0{,}1^{20}\cdot 0{,}9^{180}$$

7: Konfidensinterval

7.1 Sådan finder man konfidensintervallet

Vi spørger 1000 vælgere hvilket parti de vil stemme på. 320 siger at de vil stemme på pip-partiet.

$$\frac{320}{1000} = 0{,}32 = 32\ \%\ .$$

Da det kun er en stikprøve, behøver de 32 % ikke være nøjagtig.
Vi angiver nøjagtigheden ved at angive 95 % konfidensintervallet.
I Nspire finder vi 95 % konfidensintervallet sådan:

<u>Vælg i værktøjsmenuen:</u>

Beregninger
Statistik
Konfidensintervaller
z-interval for en andel
<u>Sæt</u>
x til 320
n til 1000
C-niveau til 0,95

<u>Vi får så:</u>

C Lower : 0,291088
C Upper : 0,348912

<u>Dvs. 95 % konfidensintervallet er:</u>

[29,1% ; 34,9%]

Hvis der var valg nu ville pip-patiet bedømt ud fra stikprøven få mellem 29,1% og 34,9 % af stemmerne.

7.2 Eksempel 1 på opgave med konfidensinterval

Opgaven

I 2017 fik 44 % af eksaminanderne 10 eller 12.
I 2019 udtog man en stikprøve på 160 eksaminander.
I stikprøven var der 60 eksaminander der fik 10 eller 12.

a) Bestem ud fra stikprøven i 2019 et 95 %-konfidensinterval for andelen af eksaminander der fik 10 eller 12.

b) Undersøg om man på baggrund af konfidensintervallet kan konkludere at andelen af eksaminander med 10 eller 12 har ændret sig.

Svaret

a) Vi taster X-værdien 60, stikprøvestørrelsen 160 og niveauet 0,95 , hvorefter Nspire bestemmer 95 % konfidensintervallet [0,299986 ; 0,450014] .
95 %-konfidensinterval for andelen af eksaminander der fik 10 eller 12 er [30,0 % ; 45,0 %] .

b) 44 % ligger i konfidensintervallet, så man kan ikke konkludere at andelen af eksaminander med 10 eller 12 har ændret sig.

7.3 Eksempel 2 på opgave med konfidensinterval

Opgaven

Et stykke legetøj fås i en blå og en rød version.
I et område A er 30 % af de solgte blå.
En stikprøve fra et andet område B er 35 % blå.
Med udgangspunkt i 95 %-konfidensintervallet kan man konkludere at andelen af blå er anderledes i område B end i område A.

a) Hvor stor må stikprøven mindst have været?

Svaret

Vi taster X-værdien 148, stikprøvestørrelsen 424 og niveauet 0,95 , hvorefter Nspire bestemmer procent blå til 35% og 95 % konfidensintervallet til $[30\%;39\%]$.
Vi taster X-værdien 149, stikprøvestørrelsen 425 og niveauet 0,95 , hvorefter Nspire bestemmer procent blå til 35% og 95 % konfidensintervallet til $[31\%;40\%]$.
For at 30% ikke skal være med i 95 % konfidensintervallet, må stikprøvestørrelsen altså være 425 eller større.

8: Tosidet binomialtest

8.1 Hypotese

I en beholder ligger mange ens pakker. I hver pakke er én kage med et bogstav. Hypotese:

Der står A på 30 % af kagerne.

For at teste hypotesen udtager vi 60 gange en pakke og undersøger om det er en A-pakke.

Hvis hypotesen er rigtig:

For hver pakke er sandsynligheden 0,30 for at det er en A-pakke.

Antallet X af A-pakker blandt de 60 er binomialfordelt med $n = 60$ og $p = 0{,}30$.

8.2 Kritisk område

Vi forkaster hypotesen hvis antallet X af A-pakker er alt for stort eller alt for lille, for så er det nok ikke rigtigt at 30 % af pakkerne er A-pakker.
Nærmere bestemt forkaster vi hypotesen hvis antallet X er et af følgende tal:

0, 1, 2, …10   og   26, 27, 28,…60

Disse tal kaldes det kritiske område.

Vi har fundet frem til det kritiske område på følgende måde:

a) Først fastsætter vi et såkaldt signifikansniveau. Vi sætter det til 5 %.
Vi dividerer signifikansniveauet med 2 og får 2,5 %.

b) Vi finder det største tal k hvor $P(X \leq k)$ er 2,5 % eller mindre.
Vi prøver os frem ved at udregne nogle sandsynligheder på Nspire:

$P(X \leq 10) = 0{,}013878$ Dette er mindre end 2,5 % .

$P(X \leq 11) = 0{,}029475$ Dette er større end 2,5 % .

Vi ser at $k = 10$.

c) Vi finder så det mindste tal k hvor $P(k \leq X)$ er 2,5 % eller mindre.
Vi prøver os frem ved at udregne nogle sandsynligheder på Nspire:

$P(26 \leq X) = 0{,}019574$ Dette er mindre end 2,5 % .

$P(25 \leq X) = 0{,}036238$ Dette er større end 2,5 % .

Vi ser at $k = 26$.

d) Fra de to k-værdier får vi at det kritiske område er

0, 1, 2, …10   og   26, 27, 28,…60

e) De værdier af X hvor vi ikke forkaster, kaldes acceptområdet.
Dvs, acceptområdet er tallene

11, 12, 13,…, 24, 25

8.3 Brug af det kritiske område:

Hvis X er 8, dvs. antal A-pakker er 8, så forkaster vi hypotesen da 8 er i det kritiske område.

Begrundelse: Hvis hypotesen havde været rigtig, så var det usandsynligt at X skulle antage en værdi i det kritiske område. Hvis hypotesen havde været rigtig, så ville sandsynligheden for at få et tal i det kritiske område være under 5 %, altså det signifikansniveau vi har brugt.

Hvis det viser sig at X er 15, dvs. hvis antal A-pakker er 15, så forkaster vi ikke hypotesen da 15 ikke er i det kritiske område.

At vi ikke forkaster hypotesen, betyder IKKE at vi har vist at hypotesen er rigtig.
Det er kun hvis vi forkaster en hypotese, at vi har vist noget.

9. Ensidet binomialtest

Denne opgavetype kommer ikke til skriftlig eksamen.

9.1 Højresidet binomialtest

Sælgeren siger at 10 % af pakkerne har defekt indhold.
Kunden udtager en stikprøve på 30 pakker for at teste følgende hypotese:

10 % af pakkerne har defekt indhold.

Signifikansniveau er 5 %.

Kunden forkaster ikke hypotesen hvis der er få defekte.
Testens kritiske område skal derfor kun bestå af store tal.

Antallet X af defekte i stikprøven er binomialfordelt med $n = 30$ og $p = 0{,}10$.
Nspire udregner:

$P(X \geq 6) = 0{,}07319$ Dette er større end 5 % .

$P(X \geq 7) = 0{,}025827$ Dette er mindre end 5 % .

Altså er det kritiske område følgende tal:

7, 8, 9,…, 29, 30 .

9.2 Venstresidet binomialtest

Sælgeren siger at 90 % af pakkerne er uden fejl.
Kunden udtager en stikprøve på 30 pakker for at teste følgende hypotese:

90 % af pakkerne er uden fejl.

Signifikansniveau er 5 %.

Kunden forkaster ikke hypotesen hvis der er mange uden fejl.
Testens kritiske område skal derfor kun bestå af små tal.

Antallet X af defekte i stikprøven er binomialfordelt med $n = 30$ og $p = 0{,}90$.
Nspire udregner:

$P(X \leq 23) = 0.025827$ Dette er mindre end 5 % .

$P(X \leq 24) = 0.07319$ Dette er større end 5 % .

Altså er det kritiske område følgende tal:

0, 1, 2,…, 22, 23 .

10. Normalfordelingsapproksimation

Denne opgavetype kommer ikke til skriftlig eksamen.

10.1 Tal der er normalfordelt.

Vi måler højden af alle drenge i 3.g .

De målte tal afsætter vi som prikker på en tallinje.

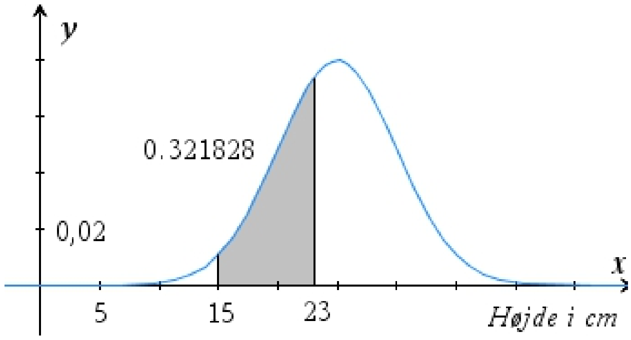
Tallene ligger ikke jævnt fordelt. De er fordelt sådan:

Der er et sted på tallinjen hvor tallene ligger tæt.
Jo længere man kommer fra dette sted, jo mindre tæt ligger tallene.

Når vi måler ting af samme slags, vil tallene ofte være fordelt på en bestemt måde som vi kalder **normalfordelt.** F.eks. er højderne af drengene i 3.g normalfordelt.

10.2 Frekvensfunktion for normalfordeling.

Vi har målt højden af nogle planter af samme type.
Højderne er normalfordelt med middelværdi 25 cm og spredning 5 cm.
Nedenfor er vist grafen for fordelingens frekvensfunktion.
Det ses at arealet under kurven i intervallet $15 \leq x \leq 23$ er 0,322.
Det betyder at 32,2 % af højderne er mellem 15 cm og 23 cm.



10.3 Beregning af sansynlighed i normalfordeling.

Hvis X er normalfordelt med middelværdi $\mu = 4{,}7$ og spredning $\sigma = 1{,}2$
er $P(4 \leq X \leq 5) = 0{,}318872 \approx 32{,}9$ %.

Dette er udregnet i Nspire ved i værktøjsmenuen at vælge
Beregninger / Sandsynlighedsregning / Fordelinger / Normal Cdf
og sætte
Nedre grænse til 4 , Øvre grænse til 5 , μ til 4,7 , σ til 1,2 .

10.4 Normalfordelingsapproksimation til binomialfordeling

Når

$$X \sim b(n, p) \quad \text{og} \quad Y \sim \mathrm{N}(\mu, \sigma)$$

hvor

$$\mu = n \cdot p \quad \text{og} \quad \sigma = \sqrt{n \cdot p \cdot (1-p)} \ ,$$

så gælder:

$$P(X \leq a) \approx P(Y \leq a + 0{,}5) \ .$$

Tilnærmelsen er god når

$$n \cdot p \cdot (1-p) > 9 \ .$$

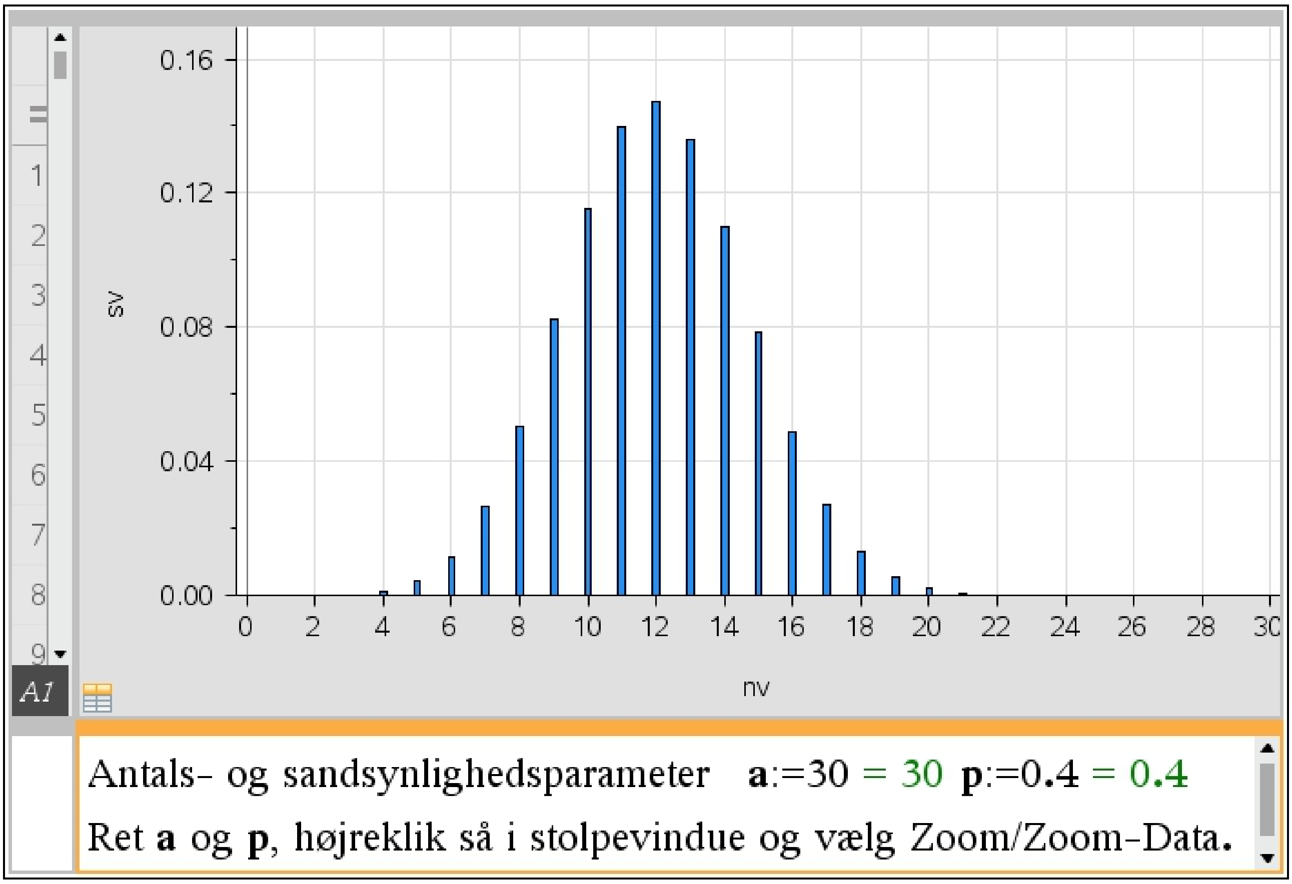
11. Stolpediagram for binomialfordeling

For at tegne et stolpediagram for en binomialfordeling skal du åbne dokumentet

mat1.dk/binom_stolpe.tns

og kopiere stolpe-vinduet over i dit eget dokument.

Sørg for at dokumrntet ligger på din computer.



A
antal måder 13
antalsparameter 9
B
basiseksperiment 9, 10
basishændelse 9
basissandsynlighed 9, 10
binomialfordeling 9
binomialfordeling i Nspire 10
binomialfordeling, formel 13, 14
E
ensidet binomialtest 18
exceptionelle værdier 12
F
fakultet 13
G
gennemsnitlig antal 4
gennemsnitlig gevinst 4, 5
gunstige udfald 7
H
hypotese 17
højresidet binomialtest 18
K
$K(n,r)$ 1
$K(n,r)$ i formelsamling 2
$K(n,r)$ i Nspire 2
$K(n,r)$, formel 13, 14
kombinatorik 1
konfidensinterval 15, 16
kritisk område 17, 18
M
mest sandsynlige udfald 11
middelværdi 4
middelværdi for binomialfordeling 11
middelværdi i Nspire 5
mulige udfald 7
multiplikationsprincippet 1
måder, antal 13
N
normale værdier 12
normalfordeling 19
normalfordeling i Nspire 19
normalfordelingsapproksimation 19
R
rækkefølge betyder ikke noget 13
S
sandsynlighedsparameter 9
signifikansniveau 17
spredning 6
spredning for binomialfordeling 11, 12
spredning i Nspire 6
stokastisk variabel 3
stolpediagram for binomialfordeling 20
sum af sandsynligheder 3
symmetrisk sandsynlighedsfelt 7, 8
T
tosidet binomialtest 17
V
venstresidet binomialtest 18
vælge r ting blandt n ting 1