
Sandsynlighedsregning

En note om sandsynlighedsregning. Via basal sandsynlighedsregning gøres læseren klar til forstå binomialfordelingen.

Henrik S. Hansen, Sct. Knud

Version 5.0

$$P(x) = \frac{\text{antal gunstige udfald}}{\text{antal mulige udfald}}$$

$$f(x) = P(X=x) = b(x; n, p) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$$

Opgaver til hæftet kan hentes her. [PDF](#)

Facit til opgaverne kan hentes her. [PDF](#)

$$P(X=4) = b(4; 10, 1/3) = \binom{10}{4} \cdot \left(\frac{1}{3}\right)^4 \cdot \left(\frac{2}{3}\right)^6 \approx 22.7\%$$

Indhold

Indledning.....	1
Sandsynlighedsregning	1
Udfald:	1
Udfaldsrum:	1
Sandsynlighedsfunktion:	1
Sandsynlighedsfelt:.....	2
Hændelse:.....	2
Komplementær hændelse:.....	2
Foreningsmængde:.....	2
Fællesmængde:	2
Kombinatorik	3
Permutation $P(n,r)$	4
Sætning: Permutation	4
Kombination $K(n,r)$	5
Sætning: Kombination	5
Kugletrækning / "kombinations"-trækning.....	6
Stokastisk variabel.....	7
Middelværdi	7
Definition: Middelværdien af en stokastisk variabel.....	8
Sætning: Regneregler for middelværdier af stokastiske variable	9
Binomial fordeling	10
Sætning: Binomialfordelingen	10
Kumulerede binomialsandsynligheder	13
Middelværdi, spredning og varians for binomialfordelt stokastisk variabel.....	15
Sætning: Middelværdi, varians og spredning for en stokastisk variabel der er binomialfordelt	15
Barske Bernhard og Lasse Lusk (binomialfordeling).....	16
Historie	18
De Mérés problem.....	18
Den delte indsats.....	18

Indledning

Vi kunne spørge os selv, hvor stor en sandsynlighed der er for at vinde i Lotto? Hvad er sandsynligheden for at jeg slår 44 seksere ud af 100slag?

Hvad betyder det, når vi siger, at sandsynligheden for at få plat ved kast med en ærlig mønt er $\frac{1}{2}$? Når vi skal til at kaste mønten, aner vi ikke, hvad mønten vil vise. Dette er tilfældet uanset om det er første gang eller nummer hundrede, vi kaster. Kaster vi mønten 10 gange, skulle man tro at halvdelen af gangene ville vise plat, men de fleste ved godt, at det ikke altid er tilfældet. Det er først når antallet af kast bliver meget stort, at frekvensen af plat kommer tættere og tættere på $\frac{1}{2}$.

Sandsynligheden udtaler sig med andre ord om de statistiske muligheder ved udførelsen af eksperimentet et stort antal gange (store tals lov).

Hvad fortæller disse procenter/sandsynligheder, og hvordan bestemmes de? Dette skal vi forsøge at svare på igennem denne note. ([video](#))

Sandsynlighedsregning

Sandsynlighedsregningen bygger på stokastiske eksperimenter (stokastisk = tilfældig), dvs. eksperimenter med tilfældigt udfald. Hertil får vi lige brug for at få introduceret nogle begreber. Jeg vil for at anskueliggøre begreberne tage udgangspunkt i et slag med en terning. ([video](#)) (gammel video)

Udfald: Resultatet af eks. ét terningkast. Jeg slår en 3er.

Udfaldsrum:

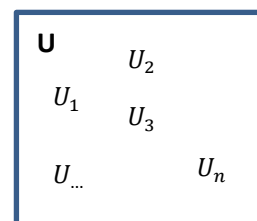
Et eksperiment kan have n forskellige udfald, som tilsammen udgør udfaldsrummet U .

En ægte (symmetrisk) terning vil have 6 mulige udfald.

$Udfaldsrummet = \{1, 2, 3, 4, 5, 6\}$

Mere generelt: Kalder vi udfaldene for $u_1 \dots u_n$, har vi

$U = \{u_1, u_2, \dots, u_n\}$



Sandsynlighedsfunktion:

Denne giver et tal mellem 0 og 1 til hvert af udfaldene i udfaldsrummet.

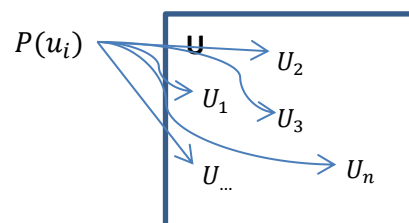
$P(u_i) \in [0, 1]$ hvor $(1 \leq i \leq n)$

$P(u_i)$ læses som "sandsynligheden for at u_i indtræffer". Eks. $P(1) = 1/6$ betyder

"sandsynligheden for at slå 1 er $1/6$ ".

Summen af alle sandsynlighederne i udfaldsrummet skal være 1.

$$P(U) = P(u_1) + P(u_2) + \dots + P(u_n) = \sum_{i=1}^n P(u_i) = 1$$



I eksemplet $P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 1$

Sandsynlighedsfelt:

Består af et udfaldsrum og en sandsynlighedsfunktion. I vores eksempel:

Udfald	1	2	3	4	5	6
P(u)	1/6	1/6	1/6	1/6	1/6	1/6

Feltet kaldes **symmetrisk**, hvis alle **n** udfald har samme sandsynlighed **p** = 1 / n.

Lav opgaver i hæftet. [PDF](#)

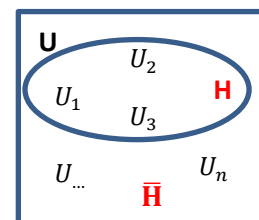
Hændelse:

En hændelse er en delmængde af udfaldsrummet. En hændelse betegnes H.

Eks. lige antal øjne med en terning. $H = \{2, 4, 6\}$.

Sandsynligheden for at en hændelse indtræffer, er summen af sandsynlighederne for de udfald, den indeholder. I eksemplet er sandsynligheden

$$P(H) = P(2) + P(4) + P(6) = \frac{1}{2}$$



Komplementær hændelse:

Den øvrige del af udfaldsrummet, ud over vores hændelse, kaldes

for den komplementære hændelse, og skrives som $\bar{H}$. I vores

eksempel vil alle ulige slag med terningen være vores komplementære hændelse,

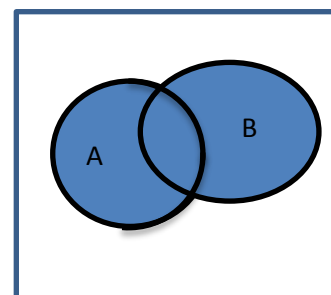
$\bar{H} = \{1, 3, 5\}$. Sandsynligheden for denne hændelse er $P(\bar{H}) = 1 - P(H)$. I vores tilfælde

$$P(\bar{H}) = 1 - \frac{1}{2} = \frac{1}{2}$$

Foreningsmængde:

Hvis man lægger to mængder sammen har man

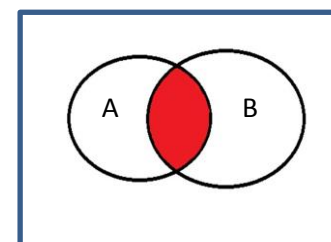
foreningsmængden. Den nye mængde indeholder altså alle elementerne fra de to oprindelige. Foreningsmængden af to mængder A og B betegnes $A \cup B$.



Fællesmængde:

Hvis man derimod kigger på hvad de to mængder har til fælles,

så kigger man på fælles mængden. Fællesmængden af A og B betegnes $A \cap B$.



Lav opgaver i hæftet. [PDF](#)

Kombinatorik

Kombinatorik er en matematisk disciplin, hvor man studerer, på hvor mange måder et sæt af elementer fra forskellige grupper kan sættes sammen ([video](#)). Hvis for eksempel 28 elever fra en klasse på Sct. Knuds gymnasium skulle stille sig i kø ved skolens kantine. På hvor mange forskellige måder kunne de gøre dette?

Lav et eksperiment.

Tag 4 genstande. Nu prøver i først med 1, så med 2, så med 3 og til sidst med 4. På hvor mange forskellige måder kan disse genstande ligge.

Kan I gætte hvor mange der ville være ved 5 genstande?

Man har indført en særlig skrivemåde for et sådant produkt af tal. Tallet $7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ betegnes $7!$ som læses ”7-fakultet”. $7!=5040$. I Nspire vælges blot ”udråbstegn”.

Dette går pludselig meget stærkt. Lad os lige kigge på klassen fra starten. Antal at måder de kan stille op på er altså

$$28! = 304888344611713860501504000000$$

Lav opgaver i hæftet. [PDF](#)

Enten-eller princippet (plus)

Hvis en hændelse kan indtræffe på m måder og en anden hændelse kan indtræffe på n måder, så er der der $m + n$ måder én af disse to hændelser kan indtræffe på.

Både-og princippet (gange)

Hvis en hændelse kan indtræffe på m måder og en anden hændelse kan indtræffe på n måder, så er der $m \cdot n$ måder disse to hændelser kan indtræffe sammen på.

Eks. I en klasse med 11 drenge og 17 piger, skal der vælges

- En klasserepræsentant,
Der skal vælges **enten** en dreng **eller** en pige, altså 1 ud af de i alt $11+17=28$ elever
- Et ”ugens dukse par”, hvor der **både** skal være en dreng **og** en pige
Hvis vi først vælger en dreng er der 11 muligheder. Hver af disse kan kobles sammen med en tilfældig af de 17 piger. For hver dreng er der altså 17 pige-kombinationer, så i alt er der $11 \cdot 17$ mulige par-kombinationer.

Permutation $P(n,r)$

Antal forskellige måder, hvorpå r elementer kan udvælges af n elementer i **bestemt rækkefølge** ([video](#))(gammel) er

$$P(n,r) = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-r+1)$$

Der skal **både** vælges en til 1. plads **og** en til 2. plads **og** osv., og som følge af *Både-og princippet*, skal antallet mulige kombinationer ganges.

Lad os kigge på $P(5,2)$. Jf. ovenstående så må $P(5,2) = 5 \cdot 4 = 20$. Så $P(5,2)$ ”på hvor mange måder kan 2 personer udvælges af 5 personer i en bestemt rækkefølge $P(5,2) = 20$ (eller i Nspire $nPr(5,2)$)

Men det bliver straks lidt mere besværligt hvis vi skal beregne eks. $P(67,56)$

Generelt

Sætning: Permutation

Antallet af r -mængder, der kan udtages fra en n -mængde, hvor rækkefølgen er vigtig, er givet ved:

$$P(n,r) = \frac{n!}{(n-r)!}$$

Bevis: ([video](#))

Vi har at $P(n,r) = n(n-1) \cdot (n-2) \cdot \dots \cdot (n-r+1)$

Ved at forlænge med $1 = \frac{(n-r)!}{(n-r)!}$ fås

$$\begin{aligned} P(n,r) &= n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-r+1) \cdot \frac{(n-r)!}{(n-r)!} \\ &= \frac{n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-r+1) \cdot (n-r) \cdot (n-r-1) \cdot \dots \cdot 2 \cdot 1}{(n-r)!} \\ &= \frac{n!}{(n-r)!} \end{aligned}$$

Hermed bevist

Eksempelvis:

På hvor mange måder kan jeg sammensætte tre ud af fem farver på.

$$P(5,3) = 5 \cdot 4 \cdot 3 = 60 \text{ eller } P(5,3) = \frac{5!}{(5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1} = 5 \cdot 4 \cdot 3 = 60$$

En pose indeholder bogstaverne A, B, C, D, E, F, G. Nu skal der trækkes fire bogstaver op. På hvor mange måder kan dette lade sig gøre?

Find selv på flere.

Kombination $K(n,r)$

Antal forskellige måder, hvorpå r elementer kan udvælges af n elementer **hvor rækkefølgen er lige meget** – dvs. i princippet **alle på én gang**. ([video](#))

Som eksempel: En gruppe på 5 personer beslutter at nedsætte et udvalg af 2 personer til sodavandsindkøb. Hvor mange måder kan udvalget sammensættes på?

Lav eksperiment: Tag 5 genstande.

På hvor mange forskellige måder kan der udvælges 5?

På hvor mange forskellige måder kan der udvælges 4?

På hvor mange forskellige måder kan der udvælges 3?

Lad os sammen kigge på $K(5,2)$.

Hvis vi som før kigger på den samlede kombinationsmængde så er den $5!$. Hvis rækkefølgen er vigtig, så har vi lige set, at vi kan bruge permutation, sådan at $P(5,2) = \frac{5!}{(5-2)!} = \frac{5!}{3!} = 5 \cdot 4$.

Da vi her er ligeglade med rækkefølgen, så må vi yderligere dividere med antallet af kombinationer af de to udvalgte, altså $2!$. Dermed fås antallet af mulige kombinationer (rækkefølgen er ligegyldig) til

$$K(5,2) = \frac{P(5,2)}{2!} = \frac{\frac{5!}{(5-2)!}}{2!} = \frac{5!}{2! \cdot (5-2)!} = \frac{5!}{2! \cdot 3!} = 10 \quad (\text{eller i Nspire } nCr(5,2))$$

Generelt kan vi samle dette til

Sætning: Kombination

Antallet af r -mængder, der kan udtages fra en n -mængde, hvor rækkefølgen er ligegyldig, er givet ved:

$$K(n,r) = \frac{n!}{(n-r)! \cdot r!}$$

Bevis: ([video](#))

Hvis vi vil udvælge r elementer af n i **bestemt rækkefølge**, kan vi benytte os af permutation med $P(n,r)$ muligheder.

Hvis rækkefølgen for den udvalgte r -mængde er **lige meget** så må vi kunne dividere $P(n,r)$ med antallet af kombinationsmuligheder for r -mængden.

$$\text{Altså } K(n,r) = \frac{P(n,r)}{r!} = \frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{(n-r)! \cdot r!}$$

Hermed bevist

Eksempelvis:

I en pose ligger bogstaverne A, B, C, D, E, F, G. På hvor mange måder kan vi trække fem bogstaver op?

$$\text{Svar: } K(7,5) = \frac{7!}{(7-5)! \cdot 5!} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{7 \cdot 6}{2 \cdot 1} = \frac{7 \cdot 3}{1} = 21$$

Lav opgaver i hæftet. [PDF](#)

Kugletrækning / "kombinations"-trækning

Vi har lige set hvordan vi kan finde antal kombinationer af en r-mængde fra en n-mængde. Nu skal vi prøve at kombinere dette med udtræk fra "blandede" n-mængder.

For at illustrere dette kan vi tage udgangspunkt i produktreglen og i et eksempel med kugletrækning.

I en pose har vi 8 røde og 6 blå kugler. Nu udtager vi tilfældigt eks. 5 kugler. Nu kunne opgaven lyde: Hvad er sandsynligheden for at det er 2 røde og 3 blå?

1. Antal mulige udfald med de 5 kugler må være $K(14,5) = 2002$
2. Antal mulige udfald for de 2 røde ud af de 8 røde. $K(8,2) = 28$
3. Antal mulige udfald for de 3 blå ud af de 6 blå $K(6,3) = 20$
4. Antallet af gunstige udfald bliver derfor (jvf. produktreglen)
 $K(8,2) \cdot K(6,3) = 28 \cdot 20 = 560$
5. Sandsynligheden bliver $\frac{560}{2002} \approx 0.28$ eller ca. 28 %

Lav opgaver i hæftet. [PDF](#)

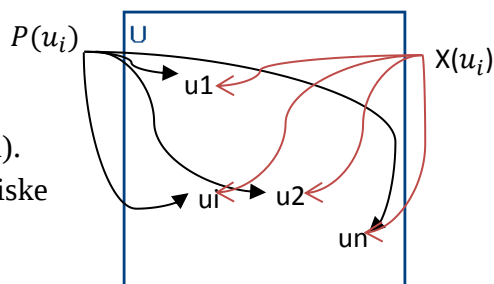
Stokastisk variabel

Som nævnt tidligere er stokastisk at sammenligne med tilfældig. En **stokastisk variabel** er altså en variabel der beskriver et *tilfældigt* forsøg, hvor udfaldet ikke er kendt. Et synonym for en stokastisk variabel vil altså være en *tilfældighedsvariabel*. ([video](#))

Til hver et hvert udfald knyttes en **stokastisk værdi**

$X(u_i) \in$ (Meget tit er X en "tæller", så $X(u)$ er et naturligt tal).

Den stokastiske variabel kobler et tal til vores udfald. Stokastiske variable betegnes ofte med store bogstaver som X , Y og Z .



Eksempel: To venner opfinder et spil, hvor de trækker kort et kort fra et almindeligt kortspil med 4 joker. Gevinsten er 30 kr. ved hjerter, 20 kr. ved klør, 10 kr. ved ruder, 0kr ved spar og 50 kr. for joker. Kortet ligges tilbage efter hvert træk. Det koster 20 kr. pr. spil. Giver spillet overskud på lang sigt?

En tabel med dette kaldes en sandsynlighedsfordeling.

t-gevinst	30 kr. (ved hjerter)	20 kr. (ved klør)	10 kr. (ved ruder)	0 kr. (ved spar)	50 kr. (ved joker)
$P(X=t)$	$\frac{13}{56} \approx 0.232$	$\frac{13}{56} \approx 0.232$	$\frac{13}{56} \approx 0.232$	$\frac{13}{56} \approx 0.232$	$\frac{4}{56} \approx 0.072$

Middelværdi

For at finde ud af om spillet giver overskud kunne vi prøve at beregne den gennemsnitlige gevinst pr. spil. Vi kalder det for middelværdien af den stokastiske variabel. ([video](#))(gammel)

$$\text{Vores gevinst i spillet } \mu = 30 \cdot \frac{13}{56} + 20 \cdot \frac{13}{56} + 10 \cdot \frac{13}{56} + 0 \cdot \frac{13}{56} + 50 \cdot \frac{4}{56} = 17.5$$

Som spiller vil man altså i gennemsnit tabe 2,50 kr. pr. spil, men som spiludbyder vinder man.

Definition: Middelværdien af en stokastisk variabel

For en stokastisk variabel X , der kan antage n værdier $t_1, t_2, t_3, \dots, t_n$

med sandsynlighederne $P(X = t_1), P(X = t_2), P(X = t_3), \dots, P(X = t_n)$

er middelværdien defineret som et vægtet gennemsnit af t -værdierne

$$\mu = E(X) = \sum_{i=1}^n t_i \cdot P(X = t_i) \quad \text{eller} \quad E(X) = \sum_{i=1}^n X(u_i) \cdot P(u_i) \quad \text{hvor } U = \{u_1, \dots, u_n\}$$

Prøv selv at lave et spil hvor der er et indskud, og hvor der er min 4 muligheder for gevinst.

Lav opgaver i hæftet. [PDF](#)



Sætning: Regneregler for middelværdier af stokastiske variable

Lad (U, P) være givet og X og Y være to stokastiske variable. Lad endvidere $a, b \in \mathbb{R}$ være konstanter. Der gælder så at:

1. $E(aX) = aE(X)$
2. $E(X + Y) = E(X) + E(Y)$
3. $E(X + a) = E(X) + a$
4. $E(a) = a$
5. $E(aX + b) = aE(X) + b$

Bevis: ([video](#))

Udgangspunktet i alle beviser er at $E(X) = \sum_{i=1}^n X(u_i) \cdot P(u_i)$

1.
$$\begin{aligned} E(a \cdot X) &= \sum_{i=1}^n (a \cdot X)(u_i) \cdot P(u_i) = \sum_{i=1}^n a \cdot X(u_i) \cdot P(u_i) \\ &= a \cdot \sum_{i=1}^n X(u_i) \cdot P(u_i) = a \cdot E(X) \end{aligned}$$
2.
$$\begin{aligned} E(X + Y) &= \sum_{i=1}^n (X + Y)(u_i) \cdot P(u_i) = \sum_{i=1}^n (X(u_i) + Y(u_i)) \cdot P(u_i) \\ &= \sum_{i=1}^n (X(u_i) \cdot P(u_i) + Y(u_i) \cdot P(u_i)) = \sum_{i=1}^n X(u_i) \cdot P(u_i) + \sum_{i=1}^n Y(u_i) \cdot P(u_i) \\ &= E(X) + E(Y) \end{aligned}$$
3.
$$\begin{aligned} E(X + a) &= \sum_{i=1}^n (X + a)(u_i) \cdot P(u_i) = \sum_{i=1}^n (X(u_i) + a) \cdot P(u_i) \\ &= \sum_{i=1}^n (X(u_i) \cdot P(u_i) + a \cdot P(u_i)) = \sum_{i=1}^n X(u_i) \cdot P(u_i) + \sum_{i=1}^n a \cdot P(u_i) \\ &= \sum_{i=1}^n X(u_i) \cdot P(u_i) + a \cdot \sum_{i=1}^n P(u_i) = E(X) + a \end{aligned}$$
4.
$$E(a) = \sum_{i=1}^n a(u_i) \cdot P(u_i) = \sum_{i=1}^n a \cdot P(u_i) = a \cdot \sum_{i=1}^n P(u_i) = a$$
5.
$$E(a \cdot X + b) = E(a \cdot X) + E(b) = a \cdot E(X) + b$$

Brugte punkt 1,3 og 4.

Hermed bevist

Binomial fordeling

I dette afsnit skal vi kigge på eksperimenter med to mulige udfald, som vi kalder *succes* og *fiasko*. Vi skal kigge på fordelingen af sandsynligheder af forskellige antal succeser. Derfor kaldes eksperimentet/forsøget for binomialfordelt. ([video](#))

En succes kunne være at slå en toer og alt andet er fiasko. Dette giver os $\frac{1}{6}$ sandsynlighed for succes, også kaldet for **primærsandsynlighed**. Sandsynligheden for fiasko er dermed $1 - P(\text{succes})$. I tilfældes med toerne er sandsynligheden for fiasko $1 - \frac{1}{6} = \frac{5}{6}$.

Det vi nu skal koncentrere sig om er gentagelser af ovenstående.

Vi slår 5 gange med en terning, dette kaldes for **antalsparameteren** og notere om det blev en succes (hvilket kunne være at slå en toer).

Denne gang blev det til 3 succeser illustreret i nedenstående

S	F	F	S	S
---	---	---	---	---

Sandsynlighederne er givet ved *succes* = 1/6 og *fiasko* = 5/6, som kan ses i tabellen

1/6	5/6	5/6	1/6	1/6
-----	-----	-----	-----	-----

Da vi har 5 uafhængige hændelser skal vi gange sandsynlighederne med hinanden.

$$\frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^2 = 0.003215$$

Denne sandsynlighed er uafhængig af kombinationen. (efterprøv selv).

Da denne sandsynlighed gælder pr. kombination, må den samlede sandsynlighed for alle kombinationer være givet ved: $K(5,3) \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^2 = 10 \cdot 0.003215 = 0.03215$

I praksis skrives dette som $P(X = 3)$ "sandsynligheden for at vores stokastiske variabel er tre, altså tre succeser."

I Nspire kan vi beregne ovenstående ved $\text{binompdf}(5, \frac{1}{6}, 3)$.

Mere generelt gælder der at:

Sætning: Binomialfordelingen

Når et eksperiment med primærsandsynlighed p for succes udføres n gange, er sandsynligheden for succes r gange givet ved:

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

*Er flere steder benævnt på følgende måde $b(n, p, r)$.

Bevis: ([video](#))

Lad S være succes og F være fiasko. Lad dertil hører $P(S) = p$ og $P(F) = 1 - P(S) = 1 - p$.

I eksperimentet er der n-udfald (samlet antal S og F) og r-succeser (S).

For hver eneste kombination vil der gælde at sandsynligheden er $p^r \cdot (1 - p)^{n-r}$, da p forekommer r gange i produktet og $(1 - p)$ så må forekomme n-r gange (resten). Eks. SFFS (ved $b(4,2,p)$)

Men da vi kan opstille $K(n, r)$ mulige kombinationer

Lille eksempel (4,2): SSFF, SFSF, SFFS, FSFS, FFSS, FSSF

og da sandsynligheden er den samme på kombination, så må den samlede sandsynlighed være

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

Hermed bevist

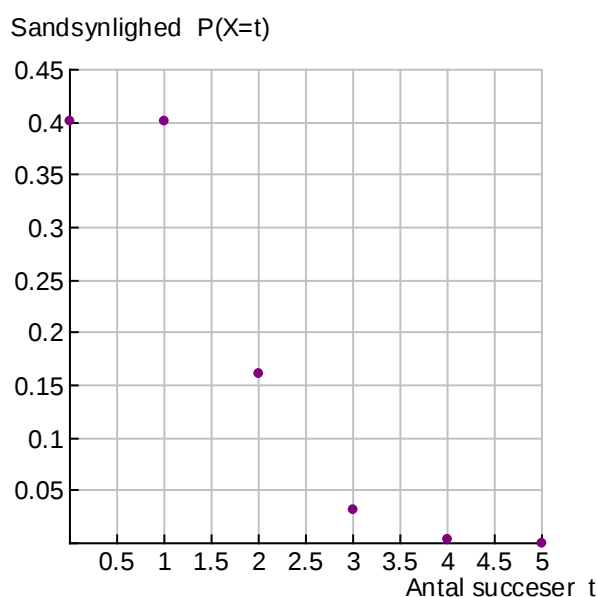
Skal vi beregne alle sandsynligheder for x-antal mulige succeser, kan vi få Nspire til at give os en liste med disse værdier i én beregning:

$\text{binompdf}(5, 1/6, \{0, 1, 2, 3, 4, 5\})$ giver dig en liste med svarene til $P(X = t)$ for alle t.

Tip: Skal du bruge en længere liste, kan du med fordel benytte $\text{seq}(i, i, 0, n)$, hvor n er vores antalsparameter. I eksemplet fra ovenstående $\text{binompdf}(5, \frac{1}{6}, \text{seq}(i, 0, 5))$

Vi kan på denne måde lave følgende skema: (til højre har vi plottet $(t, P(X = t))$)

Antal succeser t	$P(X=t)$
0	$K(5,0) \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^5 = 0.401878$
1	$K(5,1) \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{5}{6}\right)^4 = 0.401878$
2	$K(5,2) \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 = 0.160751$
3	$K(5,3) \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^2 = 0.03215$
4	$K(5,4) \cdot \left(\frac{1}{6}\right)^4 \cdot \left(\frac{5}{6}\right)^1 = 0.003215$
5	$K(5,5) \cdot \left(\frac{1}{6}\right)^5 \cdot \left(\frac{5}{6}\right)^0 = 0.000129$



Vi siger at antallet af succeser X i dette eksperiment er binomialfordelt med primærsandsynlighed $\frac{1}{6}$ og antalsparameter 5. Grafen er en diskret sandsynlighedsfordeling.

Eksempelvis:

I et terningspil skal der slås 200 gange med en terning og det gælder om at slå seksere.

Tegn sandsynlighedsfordelingen for forsøget, og bestem sandsynligheden for at slå 43 seksere ud af 200 forsøg.

Jeg opretter to lister og plotter disse

Liste med antal succeser

`suceser:=seq(i,i,0,200)`

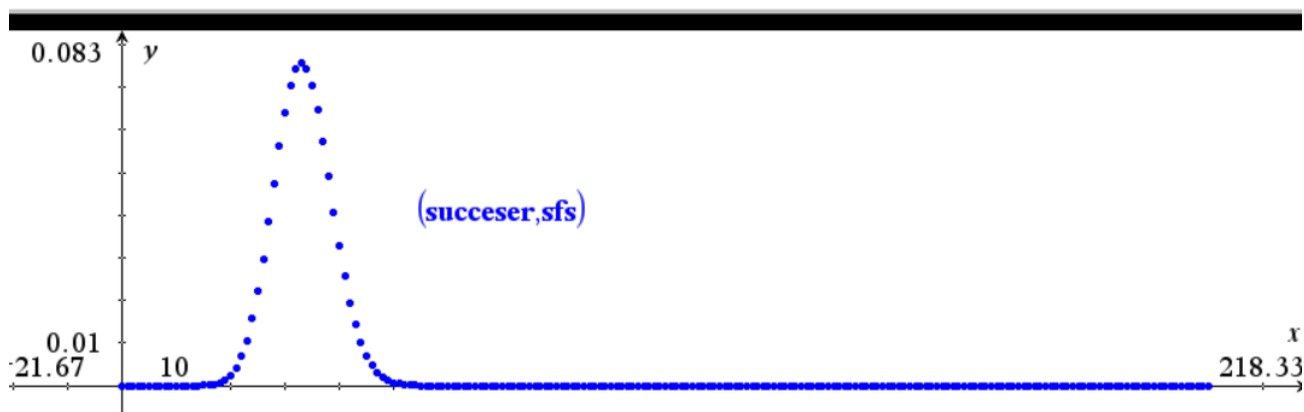
► { 0.,1.,2.,3.,4.,5.,6.,7.,8.,9.,10.,11.,12.,13.,14.,15.,16.,17.,18.,19.,20.,21.,22.,23.,24.,25.,26.,27.

Liste med sandsynligheden for netop x-antal succeser.

`sfs:=binomPdf(200, $\frac{1}{6}$ ,suceser)`

► { 1.458E-16,5.832E-15,1.161E-13,1.532E-12,1.509E-11,1.183E-10,7.69E-10,4.262E-9,2.057E-8,8.7

For 43 sekser er sandsynligheden `s43:=binomPdf(200, $\frac{1}{6}$ ,43)` ► 0.0143



Der er tale om succes eller fiasko så ”eksperimentet er binomialfordelt. Antalsparameteren er 200, den primære sandsynlighed er $\frac{1}{6}$, så jeg kan blot bestemme $\text{binopdf}\left(200, \frac{1}{6}, 43\right) = 0,014$ hvilket svarer til ca. 1,4%. På grafen ses en lodret linje gennem 43, og den skærer et punkt. Andenkoordinaten til dette punkt er 0,014.

I et kortspil gælder det om at trække spar. Hvad er sandsynligheden for at trække 15 spar ud af 50 forsøg? Tegn sandsynlighedsfordelingen.

Lav opgaver i hæftet. [PDF](#)

Kumulerede binomialsandsynligheder

Hvis vi lægger alle sandsynlighederne sammen så skulle vi gerne få 1 (100 %). ([video](#))

I Nspire kan vi bestemme den kumulerede sandsynlighed op til et bestemt antal ved

$\text{binomcdf}\left(5, \frac{1}{6}, 2\right) = 0.964506$, altså er der 96,45% sandsynlighed for at få 0,1 eller 2 succeser ud af 5 forsøg med en primær sandsynlighed på $\frac{1}{6}$.

Hvis vi skal bruge hele listen, så kan vi skrive $\text{binomcdf}\left(5, \frac{1}{6}, \text{seq}(i, 0, 5)\right)$

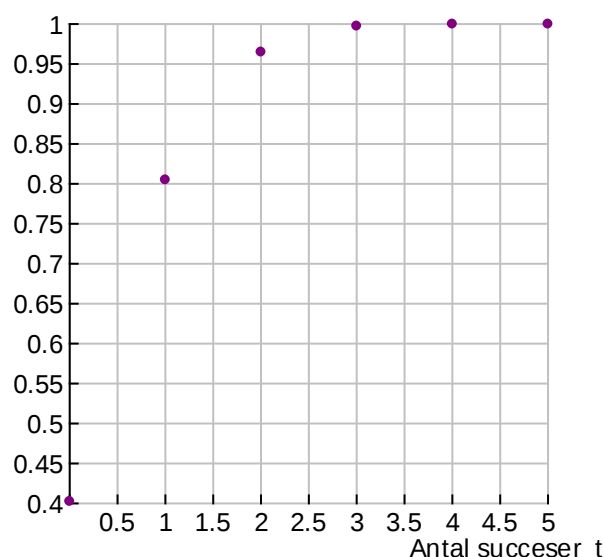
Antal succeser t	$P(X=t)$	(Kumulerede SS) $P(X \leq t)$
0	$K(5,0) \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^5 = 0.401878$	0.401878
1	$K(5,1) \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{5}{6}\right)^4 = 0.401878$	0.803755
2	$K(5,2) \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 = 0.160751$	.964506
3	$K(5,3) \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^2 = 0.03215$	.996656
4	$K(5,4) \cdot \left(\frac{1}{6}\right)^4 \cdot \left(\frac{5}{6}\right)^1 = 0.003215$	.999871
5	$K(5,5) \cdot \left(\frac{1}{6}\right)^5 \cdot \left(\frac{5}{6}\right)^0 = 0.000129$	1

Ud fra de kumulerede sandsynligheder kan vi som under deskriptiv statistik bestemme sandsynligheder som

$$P(X \leq 3) = 0.996656$$

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.964506 = 0.035494 \quad \text{osv...}$$

Kumulerede sandsynlighed

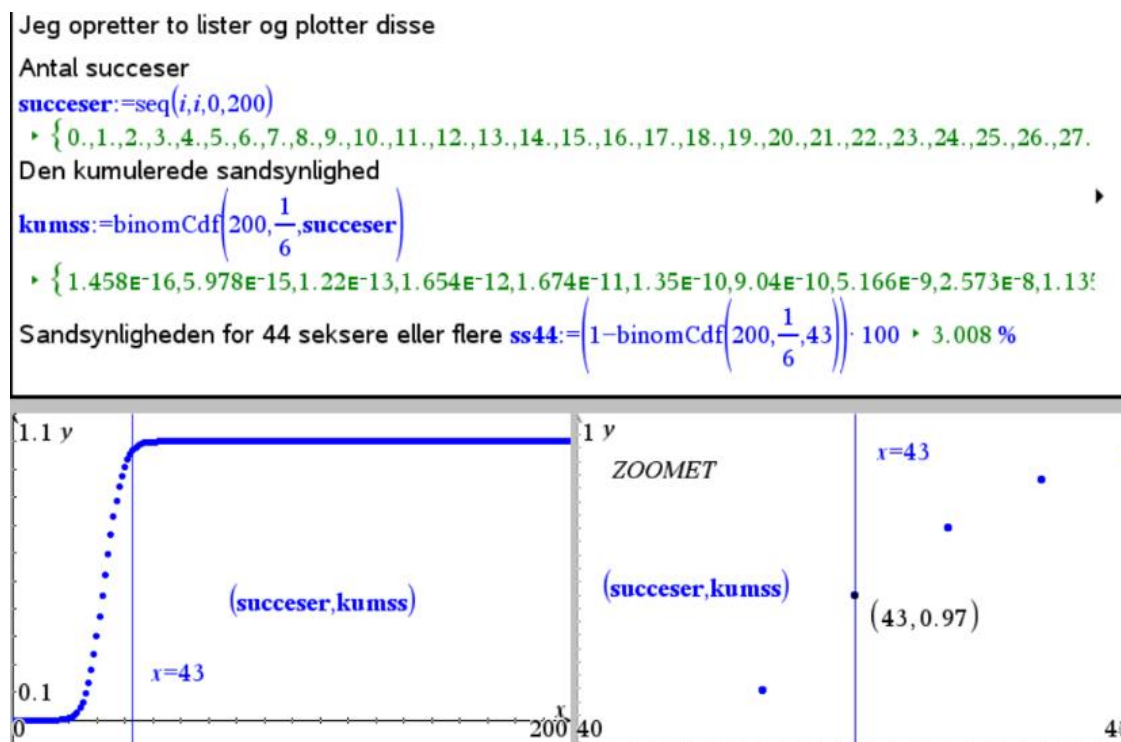


Denne fordeling af sandsynligheder kaldes en diskret fordeling, og punkterne forbindes IKKE. Årsagen er helt simpelt, at der kun eksisterer hele antal succeser.

Eksempelvis:

I et terningspil gælder det om at slå så mange seksere som muligt ud af 200 slag

Tegn den kumulerede sandsynlighedsfordeling for $b(200, \frac{1}{6}, r)$, hvor $0 \leq r \leq 200$, og bestem sandsynligheden for at slå mindst 44 seksere.



I et kortspil går det ud på at trække spar. Hvad er sandsynligheden for at trække mindst 15 spar ud af 50 forsøg? Tegn den kumulerede sandsynlighedsfordeling.

Lav opgaver i hæftet. [PDF](#)

Middelværdi, spredning og varians for binomialfordelt stokastisk variabel

Jeg vil om lidt vise, at middelværdien for stokastisk variabel, som er binomialfordelt, er givet ved $\mu = E(X) = n \cdot p$

Vi kan i første omgang nøjes med at bestemme den, som nævnt i tidligere afsnit om middelværdi

$$\mu = 0 \cdot 0.401878 + 1 \cdot 0.401878 + 2 \cdot 0.160751 + 3 \cdot 0.03215 + 4 \cdot 0.003215 + 5 \cdot 0.000129 = 0.83225 \approx \frac{5}{6}$$

Lav opgaver i hæftet. [PDF](#)

Sætning: Middelværdi, varians og spredning for en stokastisk variabel der er binomialfordelt

Lad (U, P) være givet og X være binomialfordelt med antalsparameter n og primærsandsynlighed p , dvs. $X \sim b(n, p)$

Så gælder der at:

1. Middelværdien for X er givet ved:

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

2. Variansen for X er givet ved:

$$\text{var}(X) = n \cdot p \cdot (1 - p)$$

3. Spredningen for X er givet ved:

$$\sigma(X) = \sqrt{\text{var}(X)}$$

Bevis: *Her vil vi kun bevise middelværdien. ([video](#))

For at beregne middelværdien indføres n stokastiske variable $X_i, i = 1, \dots, n$. Hvor X_i antager værdien 0 hvis der er tale om fiasko og værdien 1 ved succes. Dertil hører nu følgende sandsynligheder:

$$P(X_i = 0) = 1 - p \text{ og } P(X_i = 1) = p.$$

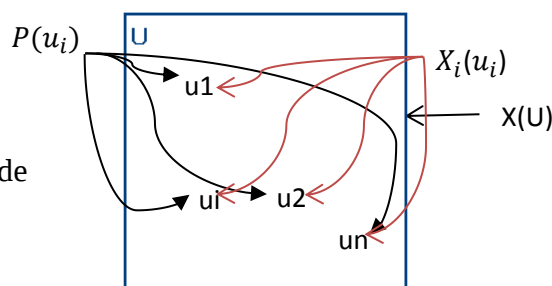
Middelværdien for hver enkelt X_i kan nu beregnes som:

$$E(X_i) = \sum_{i=1}^n X_i(u_i) \cdot P(u_i) = 0 \cdot (1 - p) + 1 \cdot p = p$$

Da X er summen af alle succeser må der gælde, at $X = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$

$$E(X) = E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n) = p + \dots + p = np$$

Hermed bevist



Lav opgaver i hæftet. [PDF](#)

Barske Bernhard og Lasse Lusk (binomialfordeling)

En lille historie, som kan sætte et hele lidt i perspektiv, og hvordan I måske kan anvende jeres viden senere ☺. Den er skrevet af AM.

”””””

Barske Bernhard og Lasse Lusk mødes på beverding og spiller terning. Den, der slår færrest seksere, skal give øl. Lasse vinder og Barske Bernhard bliver sur: ”Du snyder! Terningen er ikke ægte.” Det benægter Lasse Lusk selvfølgelig: ”Det er en ren tilfældighed – har du aldrig lært noget om sandsynligheder?” Det har Barske Bernhard faktisk, men han har glemt det igen, og det gør ham ikke mildere stemt. Tvisten udarter og Barske Bernhards kølle svinges mod stakkels Lasses hoved, da Statistik Math, som har siddet ved et af nabobordene og overhørt diskussionen, blander sig: ”Hej, slap nu lidt af – hvad er problemet?” ”Han snyder” fnyser Bernhard ”Vi har spillet 200 spil, og han har slået seksere de 44 af gangene!” Så dukker der lidt af den glemte sandsynlighedslærdom frem på lystavlen. ”Sandsynligheden for at slå en sekser med en ægte terning er en sjettedel, er vi ikke enige om det?”, siger Bernhard. De to andre nikker. ”Jamen, så skulle han jo kun slå en sjettedel af 200” – regne – regne ”altså under 35 seksere, så 44 er alt for mange!”

”Det har du måske ret i, men lad os lige undersøge det nærmere.”, siger Statistik Math. ”Vi kaster en terning, og det eneste, vi interesserer os for, er, om det er en sekser eller ikke.” Der nikkes. ”Dette eksperiment gentages 200 gange, og udfaldet hver gang er uafhængigt af, hvad vi fik sidste gang. Altså er antallet af 6’ere, som vi kan kalde X, binomialfordelt med antalsparameter 200 og sandsynlighedsparameter 1/6, hvis terningen er i orden, ikke?” Barske Bernhard, som undervejs i Math’s udredning har fremkaldt nogle skjulte erindringer fra sine matematiktimer, nikker bekræftende, mens Lasse, som ikke tør indrømme, at han ikke aner, hvad de snakker om, tøvende accepterer.

”OK, lad os så se”, siger Math og haler sin lommecomputer med NSPIRE frem. ”Hvor mange seksere, ville du acceptere, at Lasse slår, Bernhard, hvis du skal godtage, at terningen er ægte?” ”Mellem 30 og 40!”

”Nu skal du bare se noget smart. Da du gik i skole skulle du selv regne ud eller slå op i en tabel. NSPIRE klarer det nemt!”

”binomcdf(200,1/6,40) angiver, sandsynligheden for højst at få 40 seksere. Derfra skal vi så trække sandsynligheden for mindre end 30 - dvs. binomcdf(200,1/6,29) - når vi skal finde mellem 30 og 40.” Math taster og viser Nspire-resultaterne for Bernhard og Lasse.

$$\text{binomcdf}\left(200, \frac{1}{6}, 40\right) - \text{binomcdf}\left(200, \frac{1}{6}, 29\right) = .673969$$

”Lasse har altså kun 2/3 chance for at tilgodese dine krav Bernhard, selv om terningen er ægte. I en tredjedel af tilfældene vil du helt uretfærdigt komme til at anklage ham for at snyde. Du risikerer at begår den fejl at du forkaster en sand hypotese (hypotese: at det er en ægte terning) - **fejl af 1. art.**

Synes du det er fair?” ”Nå nej, det er måske lige groft nok” medgiver Bernhard ”men hvad så? Hvis jeg giver ham længere snor, så risikerer jeg jo selv at blive snydt - hvis jeg godtager en større margin, og terningen så faktisk er falsk.” ”Der har du fat i en vigtig pointe. Hvis du accepterer en forkert hypotese (hypotese: at det er en ægte terning), begår du også en fejl – det kaldes **fejl af 2.art.** I statistik er det almindelig praksis så at sige at lade tvivlen komme den anklagede til gode, så man

vil ikke lade sandsynligheden for fejl af 1. art være ret stor. Standarden er 5 %, men man ser også 2 % og 10 %. Denne grænse kaldes signifikansniveauet.”

”Jamen jeg vil da ikke være urimelig”, siger Bernhard ”Jeg holder mig til standarden på 5 %, men 44 ud af 200 er altså mange. Er det ikke for mange?”

”Nu skal du høre. De 5 % deles op med 2,5 % til hver side.” ”Det forstår jeg ikke!” ”Jeg prøver lige at lave en grafisk oversigt.” ”Hvis du ser på den, kan du godt se, at det både er få seksere og mange seksere, der er lidet sandsynlige, ikke?” ”Jo, det er klart nok, og så er det derfor, man deler op i 2,5 % *usandsynligt få* og 2,5 % *usandsynligt mange*?” ”Nemlig. Det kaldes en tosidet test – både for små og for store værdier er kritiske. Vi kan så prøve at finde ud af, hvilket antal der giver en sandsynlighed på omkring 2,5 % for dem, der er mindre end det antal, og hvilket antal der giver en sandsynlighed på ca. 2,5 % for dem, der er større. Disse 2 tal kaldes de kritiske værdier, fordi de sætter grænsen for, antallet af seksere du vil acceptere, uden at beskyldte Lasse for at snyde. Vi benytter igen Nspire til at give os de kumulerede sandsynligheder – husk lige at 2,5 % er 0.025.”

$$\text{binomcdf}\left(200, \frac{1}{6}, 22\right) = .016277 \quad \text{binomcdf}\left(200, \frac{1}{6}, 23\right) = .026928$$

”Den ene (lave) kritiske værdi er altså 22, så alt under 23 betegnes som *FOR USANDSYNLIGT*. Det kritiske niveau nedadtil, dvs. risikoen for at begå fejl af 1. art ved små værdier er 1.6 %. I den anden ende ser det sådan ud:”

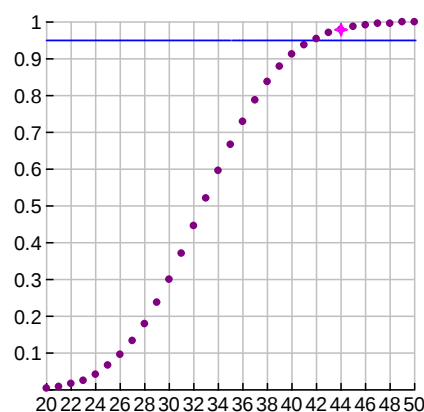
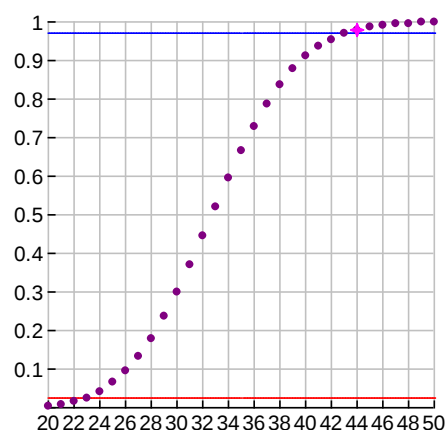
$$P(X \geq 44) = 1 - \text{binomcdf}\left(200, \frac{1}{6}, 43\right) = .03008$$

$$P(X \geq 45) = 1 - \text{binomcdf}\left(200, \frac{1}{6}, 44\right) = .019894$$

”Den anden (høje) kritiske værdi er altså 45, så alt over 44 betegnes som *FOR USANDSYNLIGT*. Det kritiske niveau opadtil, dvs. risikoen for at begå fejl af 1. art ved store værdier er 2,0 %. Det kritiske niveau er altså 1,6 % + 2,0 %

= 3,6 % < 5 %, som var dit signifikansniveau.” ”Så må jeg jo acceptere, at Lasse slog 44 6’ere med en ærlig terning”, konstaterer Bernhard ærgerligt. ”Tja ...Hov vent nu lidt. Lasse kunne da aldrig drømme om at snyde, så han fik få 6’ere. Vi kan se bort fra det lave antal.” Lasse tømmer sit glas og lusker ud ad døren. ”Du har fat i det rigtige. Man skal vurdere, den alternative hypotese ud fra

kendskabet til virkeligheden. Du argumenterer her for, at testen skal være ensidet – kun høje værdier skal være kritiske. Så kan du anvende hele signifikansniveauet i den høje ende.” ”Hallo, så ligger Lasses resultat på 44 jo i den kritiske mængde!!!!!! Jeg tænkte det nok!” Bernhard vender sig rasende om – men Lasse er forsvundet. Han skriver den kritiske værdi ned på en ølbrik for det tilfældes skyld, at han skulle støde på Lasse Lusk igen, og går så op og henter to Guinness til Math og sig selv.



Historie

(afsnittet er taget fra Mat B2, Systime)

Man regner normalt italieneren *Girolamo Cardano* (1501-1576) som grundlæggeren af sandsynlighedsregningen. I hans afhandling *Liber de ludo alen* fra 1564 omtales sandsynligheder for kortspil og terningkast.

De Mérés problem

En begyndende matematisk teori for sandsynlighed skyldes den franske matematiker og filosof 1654 *Blaise Pascal* (1623-1662) og hobbymatematikeren *Pierre de Fermat* (1601-1665). Desuden optrådte den lidenskabelige spiller *Chevalier de Méré*. Han så på følgende to hændelsen

A: at få mindst én sekser ved fire kast med en terning

B: at få mindst én dobbeltsekser ved 24 kast med to terninger

og mente, at den første var mest sandsynlig – det sagde hans erfaring ham.

Vi kan nu med vores viden beregne sandsynlighederne for dette, og se om han har ret.

A: Sandsynligheden for at få mindst én sekser må være $P(\text{mindst én sekser}) = 1 - P(\text{ingen sekser}) = 1 - \left(\frac{5}{6}\right)^4 \approx 0.5178$

B: Sandsynligheden for at få mindst én dobbeltsekser må være

$P(\text{mindst én dobbeltsekser}) = 1 - P(\text{ingen dobbeltsekser}) = 1 - \left(\frac{35}{36}\right)^{24} \approx 0.4913$

Chevalier de Méré havde altså ret.

Den delte indsats

Et problem, der i en del år optog spillere, var følgende:

”Hvordan skal man dele udbyttet i et spil, der inden sin afslutning afbrydes pludseligt”

Pascal og *Chevalier de Méré* diskuterede netop dette problem. Vi kan som illustration se på følgende terningspil mellem A og B, der skiftes til at kaste med en terning. De indskyder hver 150kr. inden spillets begyndelse, dvs. der er 300kr. i puljen.

Reglerne er følgende: Hver af de to spillere vælger et tal mellem 1-6 og de skal være forskellige. Nu går det ud på at slå sit slag tre gange i alt.

Lad os antage at spiller A har valgt 2 og spiller B har valgt 5, samt at deres spil kunne se således ud

Spiller	A	B	A	B	A	B	A	B	A	B	A	B
Øjne	1	6	2*	3	2*	2	6	5*	4	1	4	1

Med * er markeret de heldige slag for hver af spillerne. Nu må spillet imidlertid afbrydes, og spørgsmålet er, hvordan de 300kr., der er indskudt, skal fordeles mellem A og B.

Argumentet fra A er følgende: A har slået sit tal to gange, og det bedste, der kan ske, er at han i næste slag slår 2, hvorefter han har vundet de 300kr. Det værste der kan ske er at han ikke slår 2 og at B slår 5, og i så fald står de lige. Derfor er det naturligt at A først gør krav på de 150kr., han er sikker på, samt halvdelen af de 150kr. han er usikker på, dvs. 225kr.

Argumentet fra B er følgende: B's chance for at slå to 5'ere er halvt så stor som A's chance for at slå en 2'er. Derfor skal puljen fordeles så A får $\frac{2}{3}$ og B $\frac{1}{3}$, dvs. A skal have 200kr. og B 100kr.

Pascal kunne ikke løse problemet og præsenterede det derfor i et brev til *Fermat* i 1654. Det udviklede sig til korrespondance mellem de to matematikere som med denne brevveksling skabte grundlag for den klassiske sandsynlighedsregning. Det medførte, at det i det nævnte eksempel var A, der fik ret.