

1. Fundamentale begreper



1.1 Sannsynlighetsmodell

1.1.1 Eksperiment

Enhver statistisk analyse starter med et **eksperiment**.

Definisjon 1.1.1 — eksperiment

Et eksperiment er en *fysisk* begrenset serie av praktiske steg som ender i ett eller flere **utfall**.

■ Eksempel 1.1 — eksempel på eksperiment

Et eksperiment kan f.eks. være:

1. å ta et legemiddel og observere utfallet
2. å slå en mynt og observere utfallet
3. å slå en mynt hvor begge sidene viser krone og observere utfallet
4. å slå en mynt om igjen og om igjen inntil krone vises første gang og observere antall kast som måtte brukes
5. å kaste en terning og observere utfallet
6. å kaste 2 terninger 1 gang og observere utfallet
7. å spille en fotballkamp og observere sluttstillingen
8. å måle temperaturen ute en gitt dag
9. å kaste en gjenstand opp i luften og observere om den faller ned igjen
10. å telle hvor mange kunder som kommer innom butikken en gitt dag
11. å observere hvor mye neste kunde kjøper for i en gitt butikk
12. å observere hvilket klokkeslett neste akuttanrop oppstår (nødtelefon)

K Et eksperiment representerer en reell situasjon, og er således ingen *matematisk* definisjon.

K Et eksperiment *må* ha et utfall, selv om det i prinsippet kan ta svært lang tid. ¹

K Eksperimentet er koblingen mellom den matematiske delen og den *intuitive* og *praktiske* delen av en statistisk analyse, og er ofte utelatt eller har fått liten oppmerksomhet i den typiske fremstillingen av statistikk i de fleste lærebøker.

K Ved å *alltid* beskrive hvilket eksperiment man gjennomfører ved en statistisk analyse så får man en praktisk og intuitiv forståelse av situasjonen. Foran hvert eksempel eller hver oppgave i denne boken, vil derfor det underliggende eksperimentet alltid bli beskrevet. ²

¹Man kan ha eksperimenter som varer uendelig lenge, som f.eks. hvis man venter på at noe skal skje. Når blir f.eks. jorda truffet av en meteoritt av størrelse med den som utryddet dinosaurerne? Dette er en hendelse som kanskje aldri vil inntreffe, men det finnes en *sannsynlighet* for at hendelsen inntreffer. At hendelsen ikke inntreffer, er dermed et av utfallene av eksperimentet.

²Såfremt det ikke er en del av selve oppgaven å beskrive eksperimentet.

1.1.2 Utfallsrom

I et eksperiment har vi alltid et sett med *utfall*. Mengden av alle utfallene, kalles *utfallsrommet* til eksperimentet.

■ Eksempel 1.2 — krone og mynt

Anta vi har et eksperiment hvor vi slår en mynt og observerer om det blir krone eller mynt. Hvilke *utfall* har eksperimentet?

Eksperimentet har *to* utfall:³

”krone” (1.3)

”mynt” (1.4)

Utfallsrommet er *mengden* av utfallene:

$\{ \text{”krone”, ”mynt”} \}$ (1.5)

Utfallsrommet er dermed den første *matematiske* størrelsen vi trenger for å gjennomføre en statistisk analyse.

■



Figur 1.1: Eksperiment med to utfall - “krone” og “mynt”.

³Legg merke til at her har vi *navngitt* utfallene ”krone” og ”mynt”. Vi kunne ha navngitt utfallene forskjellig, f.eks. ”0” og ”1”, hvor ”0” står for krone og ”1” står for mynt. Vi kunne også ha kalt utfallene ”heads” og ”tails” dersom vi var engelstalende. Tilsvarende hadde utfallsrommene blitt:

$$\{ 1, 2 \} \quad (1.1)$$

eller

$$\{ \text{”heads”, ”tails”} \}. \quad (1.2)$$

Utfallsrommet er derfor helt avhengig av hvordan vi *navngir* utfallene, noe som gjør teorien vanskeligere, siden fremstillingen blir forskjellig alt etter hvordan utfallsrommet defineres. I kapittel 3 skal vi forklare hvorfor *ett* av disse alternativene er bedre enn de andre.

Et eksperiment og dets utfallsrom er så viktige begreper at de får sine egne symboler,

$$E \text{ og } \Omega_E \quad (1.6)$$

hhv.

Definisjon 1.1.2 — eksperiment og utfallsrom

Resultatet av et gitt eksperiment E har et visst antall *mulige utfall*. Denne mengden av mulige utfall kalles *utfallsrommet* til E og skrives:

$$\Omega_E = \left\{ \text{mengden av } \underline{\text{alle}} \text{ mulige utfall ved eksperiment } E \right\} \quad (1.7)$$

-  Den greske bokstaven kalles “omega”.
-  Dersom det er underforstått hvilket eksperiment det er snakk om, brukes bare Ω istedet for Ω_E .
-  Siden utfallsrommet til et eksperiment er definert som en *mengde*, er det lurt å friske opp kunnskapene rundt mengdelære. Appendix ?? gir en kort introduksjon til mengdelære.

Vi beskriver nå flere eksperimenter med tilhørende utfallsrom:

■ Eksempel 1.3 — frisk og ikke frisk

Eksperiment E : En pasient får et legemiddel, og deretter observeres det om pasienten blir frisk eller ikke etter forventet virketid.

Utfallsrom Ω_E : $\Omega_E = \{ \text{”frisk”}, \text{”ikke frisk”} \}$

Vi har i dette eksempelet navngitt utfallene ”frisk” og ”ikke frisk”

■

■ Eksempel 1.4 — mynt

Eksperiment E : Vi slår en mynt hvor begge sidene viser krone og observerer utfallet.

Utfallsrom Ω_E : $\Omega_E = \{ \text{"krone"} \}$

Legg merke til at fordi begge sidene viser krone på denne mynten, så har vi kun *ett* utfall: "krone".



Figur 1.2: En mynt med to like sider har bare *ett* utfall.

■

I eksempel 1.4 har vi et eksperiment med kun *ett* utfall: "krone". Vi sier da at eksperimentet er *forutbestemt* eller som vi skal kalle det: **deterministisk**.

Definisjon 1.1.3 — deterministisk eksperiment

Et eksperiment E er *deterministisk* dersom utfallsrommet Ω_E har bare *ett* element. Matematisk skriver vi at

$$|\Omega_E| = 1 \tag{1.8}$$

dvs. mengden har *kardinalitet* 1.

- K** Husk definisjonen av kardinalitet: **antall elementer i en mengde**.
- K** Et deterministisk eksperiment har *ingen usikkerhet* knyttet til seg, siden det finnes bare ett utfall av eksperimentet.
- K** Dersom man gjennomfører et deterministisk eksperiment, benyttes *ikke* statistikk for å analysere eksperimentet. Istedet tyr man gjerne til tradisjonell matematikk.

Det motsatte av et deterministisk eksperiment er et **stokastisk** eksperiment. Et stokastisk eksperiment har *mer enn* ett utfall som kan inntreffe.⁴

■ Eksempel 1.5 — stokastisk eksperiment

Å slå krone og mynt med en *balansert* mynt, er et klassisk stokastisk eksperiment. Eksperimentet har to utfall med like stor sannsynlighet for å inntreffe.⁵ ■

Definisjon 1.1.4 — **stokastisk** eksperiment

Et eksperiment E er *stokastisk* dersom antall utfall som kan inntreffe er **minst 2**.

- K Forskjellen mellom en deterministisk situasjon og en stokastisk situasjon er ofte misforstått og misbrukt. Denne definisjonen gir en svært god *intuitiv* forståelse av disse begrepene.
- K Statistikk er godt egnet for analysere et stokastisk eksperiment for å få svar på vesentlige spørsmål rundt eksperimentet.
- K Vi sier at et stokastisk eksperiment har tilknyttet en **usikkerhet** rundt utfallet, siden vi ikke vet hvilket utfall som inntreffer.

■ Eksempel 1.6 — deterministisk eksperiment

Anta vi gjennomfører eksperimentet E i eksempel 1.4 hvor vi slår en mynt hvor begge sidene viser krone, men nå definerer vi utfallsrommet

$$\Omega_E = \{ \text{”krone”, ”mynt”} \}. \quad (1.9)$$

Spørsmål: Er eksperimentet *stokastisk*?

Svar: Man kan fort tro at eksperimentet er stokastisk siden vi har to utfall, men siden det ene utfallet (mynt) ikke kan inntreffe, så har eksperimentet kun ett utfall som faktisk kan skje. (Man kan gjerne si at det ikke er hensiktsmessig å definere to element i utfallsrommet siden kun ett kan inntreffe).

Eksperimentet er derfor *deterministisk*. ■

⁴Formuleringen ”...som kan inntreffe” er svært viktig i denne sammenhengen. Hva menes med ”kan inntreffe”? ”Kan inntreffe” betyr at det finnes en viss sjanse for at utfallene er mulige. Vi sier at *sannsynligheten* for at utfallene inntreffer er større enn null.

⁵En *balansert* mynt, betyr at det er 50-50 sjanse for å krone eller mynt. Vi skal heretter anta at alle mynter er balanserte, men mindre det står noe annet.

■ Eksempel 1.7 — antall kast for å få krone

Eksperiment E: Vi slår en mynt om igjen og om igjen inntil krone vises første gang og teller antall kast som vi måtte bruke.

Utfallsrom Ω_E : $\Omega_E = \mathbb{N}^+ = \{ 1, 2, 3, 4, \dots \}$

Det er naturlig å kalle utfallene "1", "2", "3", etc. siden vi i eksperimentet *teller* antall kast vi trenger før vi får første krone. Legg merke til at vi ikke kan ha 0 som utfall, siden vi må ha minst et kast for å få krone.

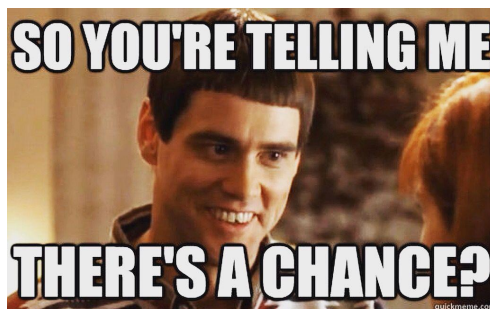
Siden vi i prinsippet kan bruke et *vilkårlig* antall kast før vi får første krone, må vi inkludere *alle* de positive heltallene.

K I eksempel 1.7 må vi ta med *alle* de mulige utfallene, uansett hvor usannsynlige de er. Hva er f.eks. sannsynligheten for at utfallet til eksperimentet er 100?

At utfallet blir 100, betyr at vi fikk mynt 99 ganger på rad før vi fikk kronen. Siden vi måtte kaste 100 ganger, hver gang med sannsynlighet 0.5, blir sannsynligheten for å få utfallet

[illegible]

Det er mao. en *svært* liten sannynlighet å få utfallet 100, men den er *ikke* 0. Vi må altså ta med alle utfallene, uansett hvor usannsynlig de er, så lenge sannsynligheten ikke er null. Mer om sannsynligheter kommer i neste avsnitt.



Figur 1.3: Et kjent sitat fra filmen ”*Dumb & Dumber*”, hvor ”Dumber” blir glad for å høre at han har en milliontedels sjanse for å bli sammen med damen han er forelsket i.

■ Eksempel 1.8 — tidspunkt for første nødrops

Eksperiment E : På ett gitt dogn fra kl. 00:00 til 24:00 registrerer vi hvilket klokkeslett første nødrops oppstår.

Utfallsrom Ω_E : $\Omega_E = [0, 24) \cup \{ \text{"aldri"} \}$

Dersom det forekommer et nødrops ila. dognet, må det skje mellom 0 og 24. Hvis det skjer akkurat kl. 00:00, som vi representerer ved 0, så må vi ha med dette utfallet. Vi kan ikke ha med akkurat 24:00, siden dette er en del av neste dag. Utfallsrommet bør derfor inneholde alle tallene mellom 0 og 24, hvor 0 er inkludert mens 24 er ekskludert. Dette intervallet betegnes ved $[0, 24)$.

I tillegg må vi ta høyde for at det kanskje aldri forekommer et nødrops. Vi har valgt å kalle dette utfallet "aldri". Vi inkluderer derfor mengden $\{ \text{"aldri"} \}$ til utfallsrommet, ved hjelp av unionsoperatoren $\cup$.⁶



Figur 1.4: Nødrops.

■

⁶For mer om operasjoner mellom mengder, se appendiks ??.

1.1.3 Sannsynlighet

Utfallssannsynlighet

Vi har hittil definert hva vi mener med et stokastisk eksperiment E og dets utfallsrom Ω_E . Siden E er et *stokastisk* eksperiment, vet vi at det finnes minst 2 utfall som kan inntreffe.

La $u \in \Omega_E$ være ett av disse utfallene.

Spørsmål: Hva er **sannsynligheten** p_u for at utfallet u inntreffer? ⁷

Dette spørsmålet ligger i hjertet av sannsynlighetsteori og forklares nærmere i dette avsnittet. Vi starter med et innledende eksempel.



Figur 1.5: Sannsynlighet.

■ Eksempel 1.9 — terningkast

Eksperiment E : Anta vi kaster en balansert terning og observerer antall øyne vi får.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, 5, 6 \}$

Utfallene representeres her ved tallene 1 til 6, som betegner antall øyne vi får.

Spørsmål: Hva er **sannsynligheten** p_5 for at utfallet 5 inntreffer?

Svar: Sannsynligheten for at utfall 5 inntreffer er $\frac{1}{6}$, dvs $p_5 = \frac{1}{6}$.

Siden terningen er balansert, er hvert utfall like *sannsynlige*. Siden vi har 6 sider i en terning, vil hvert utfall ha sannsynligheten $\frac{1}{6}$ for å inntreffe.

■

⁷Vi skal heretter bruke betegnelsen p_u for sannsynligheten for at utfall u inntreffer. Bokstaven p er ikke tilfeldig valgt - den stammer fra det engelske ordet for sannsynlighet: "probability".

I eksempel 1.9 hvor vi kastet en balansert terning, fant vi at sannsynligheten for å få 5 øyne er $\frac{1}{6}$. Faktisk fant vi at *alle* utfallene har samme sannsynlighet $\frac{1}{6}$ for å inntreffe.

Fra eksempelet trekker vi ut 2 fundamentale egenskaper ved begrepet *sannsynlighet*:

1. Ethvert utfall $u \in \Omega_E$ ved et eksperiment E tilordnes en sannsynlighet p_u .
2. Sannsynligheten p_u er et tall mellom 0 og 1, dvs. $p_u \in [0, 1]$.

Disse to egenskapene definerer *matematisk* hva som menes med *sannsynligheten for at et utfall inntreffer*.

Notasjon 1.1. Til hvert utfall $u \in \Omega_E$ i et stokastisk eksperiment E , kan vi tilordne sannsynligheten $p_u \in [0, 1]$ for at utfallet inntreffer. Vi kaller p_u *utfallssannsynligheten*.

K Legg merke til at vi ikke forteller hvordan vi *finder* p_u , kun at den eksisterer for hvert utfall $u \in \Omega_E$ og at den er et tall mellom 0 og 1.

K Dersom det er klart fra konteksten at vi snakker om en utfallssannsynlighet, sier vi ofte bare "sannsynlighet".

■ Eksempel 1.10 — terningkast

For eksperimentet i eksempel 1.9 har vi følgende *utfallssannsynligheter*:

$$p_1 = \frac{1}{6}, \quad p_2 = \frac{1}{6}, \quad p_3 = \frac{1}{6}, \quad p_4 = \frac{1}{6}, \quad p_5 = \frac{1}{6}, \quad p_6 = \frac{1}{6} \quad (1.11)$$



Figur 1.6: Terningkast.

■ Eksempel 1.11 — antall kast for å få krone

Vi skal finne utfallssannsynlighetene til eksperimentet E i eksempel 1.7, hvor vi teller antall kast vi trenger for å få krone. Vi husker at utfallsmengden var definert ved:

$$\Omega_E = \{ 1, 2, 3, 4, 5, \dots \} \quad (1.12)$$

Utfallsannsynlighetene er gitt ved:

$$p_1 = \frac{1}{2} \quad (1.13)$$

$$p_2 = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2^2} \quad (1.14)$$

$$p_3 = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2^3} \quad (1.15)$$

$$p_4 = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2^4} \quad (1.16)$$

$$\begin{aligned} & \vdots \\ p_n &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = \frac{1}{2^n} \\ & \vdots \end{aligned} \quad (1.17)$$

Sannsynligheten for å få utfall n er derfor $p_n = \frac{1}{2^n}$ for alle $n \in \Omega_E$.



Figur 1.7: Myntkast.

For et vilkårlig eksperiment E vil utfallssannsynlighetene generelt være *ukjente*, selv om de finnes.

■ Eksempel 1.12 — vaksinerings

Anta at et legefirma har nettopp utviklet det de tror er en ny vaksine for en sykdom.

Eksperiment E : En person får vaksinen og utsettes for smitte av sykdommen. Det registreres så om hvorvidt personen blir smittet eller ikke.

Utfallsrom Ω_E : $\Omega_E = \{\text{”smittet”}, \text{”ikke smittet”}\}$

Spørsmål: Hva er utfallssannsynlighetene for dette eksperimentet?

Svar: Utfallssannsynlighetene er *ukjente*.

Vi kjenner med andre ord ikke sannsynlighetene

$$p_{\text{smittet}} = ? \quad (1.18)$$

$$p_{\text{ikke smittet}} = ? \quad (1.19)$$

Vi har ingen måte å finne de *eksakte* sannsynlighetene på, selv om vi antar at de finnes. Vi må derfor reformulere spørsmålet:

Spørsmål: Hvordan kan vi finne **tilnærmede** verdier for utfallssannsynlighetene for dette eksperimentet?

Svar: Vi kan gjenta eksperimentet på $N = 5000$ forskjellige personer og telle hvor mange som ble smittet N_{smittet} .

Anta f.eks. at $N_{\text{smittet}} = 71$ personer (tilsvarende $N_{\text{ikke smittet}} = 4929$).

Vi kan da beregne de *eksperimentelle* sannsynlighetene:

$$p_{\text{smittet}} \approx \frac{N_{\text{smittet}}}{N} = \frac{71}{5000} = 0.0142 \quad (\text{andel som ble smittet})$$

$$p_{\text{ikke smittet}} \approx \frac{N_{\text{ikke smittet}}}{N} = \frac{4929}{5000} = 0.9858 \quad (\text{andel som ikke ble smittet})$$

Eksperimentell sannsynlighet

I eksempel 1.12 spurte vi hva vi gjør dersom vi ikke kjenner utfallssannsynlighetene til et eksperiment.⁸

En måte var å finne **tilnærmede** sannsynligheter ved å repetere eksperimentet svært mange ganger. Disse tilnærmede sannsynlighetene kaller vi **eksperimentelle** sannsynligheter.

Men hva er eksperimentell sannsynlighet og hvordan vet vi at disse faktisk tilnærmer utfallssannsynlighetene? Isåfall, er det slik at jo flere eksperimenter vi gjør desto bedre blir tilnærmingene?

Definisjon 1.1.5 — eksperimentell sannsynlighet for et utfall

Anta et eksperiment E med utfallsrom Ω_E repeteres totalt N ganger (antall forsøk). La N_u være antall ganger utfall $u \in \Omega_E$ forekom.

Den *eksperimentelle sannsynligheten* for utfall u er da definert som

$$f_u = \frac{N_u}{N} \quad \left(\frac{\text{antall ganger } u \text{ inntreffer}}{\text{antall forsøk}} \right) \quad (1.20)$$

K Eksperimentell sannsynlighet kallas også **empirisk sannsynlighet** eller **relativfrekvens**.⁹ Størrelsen N_u kalles også **hyppigheten** eller **frekvensen** av utfall u .

K Frekvensene N_u og de tilhørende eksperimentelle sannsynlighetene f_u er størrelsene som brukes i såkalte **histogram**.

K Vi kommer til å bruke den kollokviale kortformen

$$f_u = \frac{N_u}{N} = \frac{\text{antall ganger } u \text{ inntreffer}}{\text{antall forsøk}} = \frac{\# \text{ gunstige}}{\# \text{ mulige}} \quad (1.21)$$



Figur 1.8: Telling.

⁸I anvendt statistikk kjenner vi som regel ikke utfallssannsynlighetene, men må enten bruke statistisk inferens (del 2) eller gjøre à priori antagelser (dvs. beste gjett fra informasjonen som foreligger).

⁹Betegnelsen f_u gjenspeiler at eksperimentell sannsynlighet egentlig er en relativfrekvens.

■ Eksempel 1.13 — ubalansert terning

Ekspirement E : En ubalansert terning kastes og antall øyne registreres.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, 5, 6 \}$

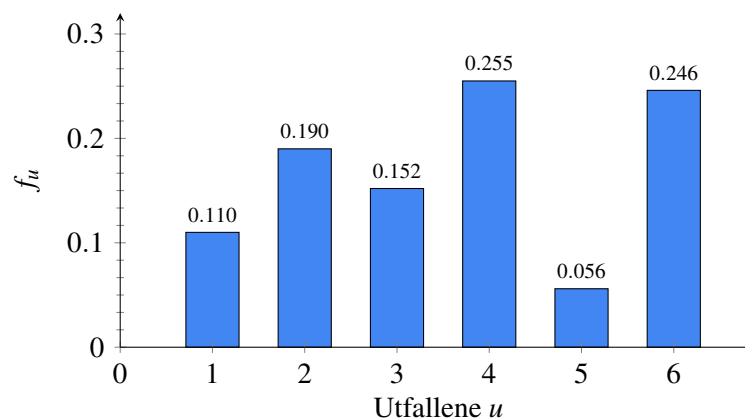
Siden terningen er *ubalansert* kjenner vi ikke utfallssannsynlighetene. Vi ønsker derfor å finne tilnærmede sannsynligheter ved hjelp av eksperimentelle sannsynligheter.

Tabell 1.2 viser frekvensene og de eksperimentelle sannsynlighetene etter å ha kastet terningen $N = 1\,000$ ganger.

Utfall $u \in \Omega_E$	1	2	3	4	5	6
Frekvenser N_u	110	190	152	255	56	249
Eksp. sannsynligheter f_u	0.11	0.19	0.152	0.255	0.056	0.249

Tabell 1.1: Frekvenser og eksperimentelle sannsynligheter etter $N = 1\,000$ kast.

I figur 1.10 har vi plottet inn de eksperimentelle sannsynlighetene i et *histogram*.¹⁰



Figur 1.9: Histogram av de **eksperimentelle sannsynlighetene**.

Fra de eksperimentelle sannsynlighetene får vi noenlunde gode estimater for utfallssannsynlighetene. Men hvis vi skal ha mer nøyaktige tilnærminger, må vi øke antall kast.

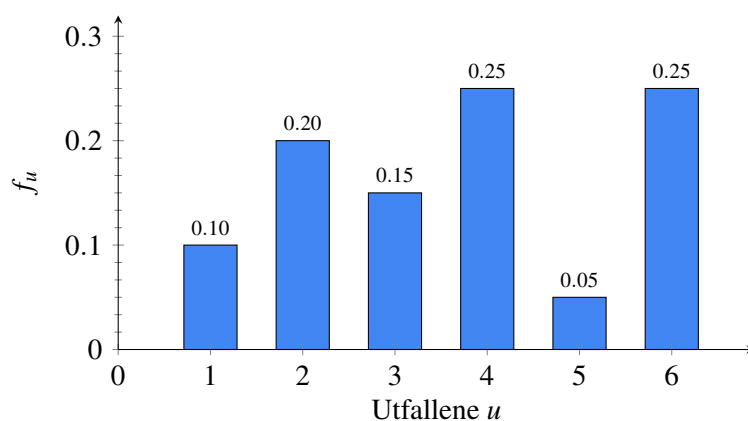
¹⁰I et histogram markeres utfallene langs x -aksen og de eksperimentelle sannsynlighetene langs y -aksen.

Tabell 1.2 viser frekvensene og de eksperimentelle sannsynlighetene etter å ha kastet terningen $N = 1\,000\,000$ ganger.¹¹

Utfall $u \in \Omega_E$	1	2	3	4	5	6
Frekvenser N_u	100 023	200 010	149 985	250 021	49 924	250 037
Eksper. sannsynligheter f_u	0.10	0.20	0.15	0.25	0.05	0.25

Tabell 1.2: Frekvenser og eksperimentelle sannsynligheter etter $N = 1\,000\,000$ kast.

I figur 1.10 har vi plottet inn de eksperimentelle sannsynlighetene i et nytt histogram.



Figur 1.10: Histogram av de eksperimentelle sannsynlighetene.

På bakgrunn av de eksperimentelle sannsynlighetene er det rimelig å konkludere med følgende utfallssannsynligheter:

$$p_1 = 0.1 \quad (\text{en 1er forekommer 10\% av gangene}) \quad (1.22)$$

$$p_2 = 0.2 \quad (\text{en 2er forekommer 20\% av gangene}) \quad (1.23)$$

$$p_3 = 0.15 \quad (\text{en 3er forekommer 15\% av gangene}) \quad (1.24)$$

$$p_4 = 0.25 \quad (\text{en 4er forekommer 25\% av gangene}) \quad (1.25)$$

$$p_5 = 0.05 \quad (\text{en 5er forekommer 5\% av gangene}) \quad (1.26)$$

$$p_6 = 0.25 \quad (\text{en 6er forekommer 25\% av gangene}) \quad (1.27)$$

■

¹¹I tabellen har de eksperimentelle sannsynlighetene blitt avrundet til to desimaler.

Dersom vi gjentar et stokastisk eksperiment *uendelig* mange ganger, vil den eksperimentelle sannsynligheten f_u bli nøyaktig lik utfallssannsynligheten p_u (**store talls lov**).¹²

Setning 1.1.1 — store talls lov

Anta vi har et eksperiment E med utfallsrom Ω_E . La p_u være sannsynligheten for at utfall u inntreffer.

La N være antall ganger vi gjentar eksperimentet. Da kan vi med sikkerhet si at

$$p_u = \lim_{N \rightarrow \infty} f_u = \lim_{N \rightarrow \infty} \frac{N_u}{N} \quad (\text{i det lange løp}) \quad \left(\frac{\# \text{ gunstige}}{\# \text{ mulige}} \right) \quad (1.28)$$

hvor $f_u = \frac{N_u}{N}$ er den eksperimentelle sannsynligheten for utfall u .

K Store talls lov er *svært* fundamental og gir en intuitiv tolkning av hva sannsynlighet er:

Utfallssannsynligheten p_u kan tolkes som *andelen av antall ganger u inntreffer i det lange løp*.¹³

Jo flere ganger vi repeterer eksperimentet, desto bedre tilnærmer den eksperimentelle sannsynligheten f_u den virkelige sannsynligheten p_u .

K Det er ikke alltid mulig å bruke store talls lov rent praktisk, da det kan både koste tid og penger å repetere eksperimentet.

K Store talls lov forteller oss hvordan vi finner selve *verdien* for en utfallssannsynlighet og gir oss en mulighet til å *definere* hva utfallssannsynlighet er.

Definisjon 1.1.6 — utfallssannsynlighet

Anta vi har et eksperiment E med utfallsrom Ω_E og la $u \in \Omega_E$ være et vilkårlig utfall. *Utfallssannsynligheten* til u er definert ved

$$p_u = \lim_{N \rightarrow \infty} f_u = \lim_{N \rightarrow \infty} \frac{N_u}{N} \quad (\text{i det lange løp}) \quad \left(\frac{\# \text{ gunstige}}{\# \text{ mulige}} \right) \quad (1.29)$$

hvor f_u er den eksperimentelle sannsynligheten for utfall u .

¹²I prinsippet må eksperimentet gjentas under *helt like omgivelser*, hvor et eksperiment ikke påvirkes av annet.

¹³Begrepet ”i det lange løp” betyr at vi gjentar eksperimentet uendelig mange ganger.

Begivenhetssannsynlighet

I tillegg til utfallssannsynlighetene ved et eksperiment E , er man også interessert i sannsynligheten for at resultatet av eksperimentet havner i en *delmengde* $S \subset \Omega_E$ av utfallsmengden.

■ Eksempel 1.14 — sannsynlighet for å havne i en delmengde av Ω_E

Eksperiment E : En *ubalansert* terning kastes og antall øyne registreres.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, 5, 6 \}$

Vi spør:

Hva er sannsynligheten for at eksperimentet gir ett av utfallene i mengdene:

$$S_1 = \{ 2, 4, 6 \} \quad (\text{kun partall}) \quad (1.30)$$

$$S_2 = \{ 1, 3, 5 \} \quad (\text{kun oddetall}) \quad (1.31)$$

$$S_3 = \{ 1, 2, 3 \} \quad (\text{mindre enn 4}) \quad (1.32)$$

$$S_4 = \{ 4, 5, 6 \} \quad (\text{større enn 3}) \quad (1.33)$$

Vi kaller slike delmengder av utfallsrommet *begivenheter*.¹⁴

■

Notasjon 1.2. Vi sier at en begivenhet *forekommer* eller *inntreffer* dersom eksperimentet resulterer i et utfall i begivenheten.

For å *utvide* definisjonen for sannsynlighet til å omfatte *begivenheter*, må vi først definere hva vi mener med den *eksperimentelle* sannsynligheten for en begivenhet.

Definisjon 1.1.7 — eksperimentell sannsynlig for en begivenhet

Anta et eksperiment E med utfallsrom Ω_E repeteres N ganger (antall forsøk).

La N_S være antall ganger begivenhet $S \subseteq \Omega_E$ inntreffer.

Den *eksperimentelle sannsynligheten* for at begivenheten S inntreffer er gitt ved

$$f_S = \frac{N_S}{N} \quad \left(\frac{\# \text{ gunstige}}{\# \text{ mulige}} \right) \quad (1.34)$$

¹⁴Noen bruker også “*hendelse*” istedet for “begivenhet”. På engelsk brukes ofte ordet “*event*”.

■ Eksempel 1.15 — eksperimentell sannsynlig for en begivenhet

Anta vi ønsker å beregne eksperimentelle sannsynligheter for begivenhet $S_1 = \{2, 4, 6\}$ i eksempel 1.14. Siden terningen er ubalansert, kjenner vi ikke utfallssannsynlighetene.

$N = 100$: Anta vi fikk følgende treff etter $N = 100$ kast:

	N_1	N_2	N_3	N_4	N_5	N_6
Antall treff	26	5	23	7	25	14

For å finne den eksperimentelle sannsynligheten for begivenheten S_1 må vi telle antall ganger et forsøk ga enten 2, 4 eller 6 øyne. Vi må derfor summere N_2 , N_4 og N_6 :

$$f_{S_1} = \frac{N_S}{N} = \frac{N_2 + N_4 + N_6}{N} = \frac{5 + 7 + 14}{100} = \frac{26}{100} = 0.26 \quad (1.35)$$

$N = 1\,000$: Anta vi fikk følgende treff etter $N = 1\,000$ kast:

	N_1	N_2	N_3	N_4	N_5	N_6
Antall treff	259	59	262	72	223	125

Vi summerer N_2 , N_4 og N_6 :

$$f_{S_1} = \frac{N_S}{N} = \frac{N_2 + N_4 + N_6}{N} = \frac{59 + 72 + 125}{1000} = \frac{256}{1000} = 0.256$$

$N = 10\,000$: Anta vi fikk følgende treff etter $N = 10\,000$ kast:

	N_1	N_2	N_3	N_4	N_5	N_6
Antall treff	2525	601	2520	710	2458	1186

Vi summerer N_2 , N_4 og N_6 :

$$f_{S_1} = \frac{N_S}{N} = \frac{N_2 + N_4 + N_6}{N} = \frac{601 + 710 + 1186}{10000} = \frac{2497}{10000} = 0.2497$$

Vi ser at f_{S_1} nærmer seg 0.25 når antall forsøk øker. Det vil si at terningen er balansert slik at ca. 25% av kastene gir et partall.

■

Store talls lov er også korrekt for begivenheter, så vi har følgende svært viktige definisjon: ¹⁵

Definisjon 1.1.8 — begivenhetssannsynlighet

Anta vi gjennomfører et eksperiment E med utfallsrom Ω_E . La $S \subseteq \Omega_E$ være en mulig *begivenhet*.

La N være antall ganger eksperiment E repeteres (antall forsøk), og la N_S være antall ganger begivenheten S inntreffer.

Begivenhetssannsynligheten, dvs. sannsynligheten for at eksperimentet gir ett av utfallene i begivenheten S er definert ved

$$P(S) = \lim_{N \rightarrow \infty} f_S = \lim_{N \rightarrow \infty} \frac{N_S}{N} \quad (\text{i det lange løp}) \quad \left(\frac{\# \text{ gunstige}}{\# \text{ mulige}} \right) \quad (1.36)$$

hvor f_S er den eksperimentelle sannsynligheten for begivenheten S .

- K** Legg merke til at vi bruker ”stor” P for begivenhetssannsynlighet og liten p for utfallssannsynlighet. Det er flere grunner for dette:
- Vi ønsker å poengtere forskjellen når vi snakker om en utfallssannsynlighet og når vi snakker om en begivenhetssannsynlighet.
 - Utfallssannsynlighetene kan oppfattes som *byggeklossene* for begivenhetssannsynlighetene, og derfor er det hensiktsmessig med en litt enklere notasjon for disse.

- K** Begivenhetssannsynlighet er en *utvidelse* av utfallssannsynlighet siden vi alltid kan definere begivenheten som består kun av ett utfall. La f.eks.

$$S_u = \{u\} \quad (\text{begivenheten for utfall } u) \quad (1.37)$$

hvor $u \in \Omega_E$ er et utfall. Siden denne begivenheten representerer utfallet u , må sannsynligheten for at begivenheten inntreffer være lik utfallssannsynligheten, dvs.:

$$P(S_u) = P(\{u\}) = p_u \quad (1.38)$$

for alle utfall $u \in \Omega_E$.

- K** Siden et utfall også er en begivenhet (fra kommentaren ovenfor), vil vi heretter bare si *sannsynlighet* som fellesbegrep for både utfallssannsynlighet og begivenhetssannsynlighet, med mindre det er et poeng å skille mellom disse to.

¹⁵Definisjon 1.1.8 bruker ordlyden ”mulig begivenhet” fordi det viser seg at når utfallsrommet f.eks. er alle reelle tall større enn 0 (f.eks. temperatur), finnes det delmengder hvor sannsynligheten ikke er vel-definert. Vi antar for enkelthets skyld at alle begivenheter er mulige begivenheter.

■ Eksempel 1.16 — begivenhetssannsynlighet via definisjonen

Eksperiment E : En balansert terning kastes og antall øyne registreres.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, 5, 6 \}$

Regn ut sannsynligheten for begivenhetene ved hjelp av definisjonen¹⁶

$$S_1 = \{ 2 \} \quad (\text{utfallet 2}) \quad (1.39)$$

$$S_2 = \{ 1, 3, 5 \} \quad (\text{oddetall}) \quad (1.40)$$

$$(1.41)$$

Siden terningen er *balansert* kan vi anta at utfallssannsynlighetene er $p_u = \frac{1}{6}$ for $u \in \Omega_E$.

$$P(S_1) = P(\{2\}) = p_2 = \lim_{N \rightarrow \infty} \frac{N_2}{N} = \frac{1}{6} \quad (1.42)$$

$$P(S_2) = P(\{1, 3, 5\}) = \lim_{N \rightarrow \infty} \frac{N_{S_1}}{N} \quad (1.43)$$

$$= \lim_{N \rightarrow \infty} \frac{N_1 + N_3 + N_5}{N} = \lim_{N \rightarrow \infty} \left(\frac{N_1}{N} + \frac{N_3}{N} + \frac{N_5}{N} \right) \quad (1.44)$$

$$= \lim_{N \rightarrow \infty} \frac{N_1}{N} + \lim_{N \rightarrow \infty} \frac{N_3}{N} + \lim_{N \rightarrow \infty} \frac{N_5}{N} \quad (1.45)$$

$$= p_1 + p_3 + p_5 \quad (1.46)$$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \quad (1.47)$$

$$= \frac{3}{6} \quad (1.48)$$

$$= \frac{1}{2} \quad (1.49)$$

■

Som vi ser fra eksempel 1.16 er det ganske vanskelig å *regne* ut sannsynligheter direkte fra definisjonen. I avsnitt 1.2 skal vi utlede vi fundamentale **regneregler** som vil forenkle beregninger av sannsynligheter drastisk.

Men først oppsummerer vi teorien vi har utviklet i en *sannsynlighetsmodell*.

¹⁶Vi mener egentlig *begivenhetssannsynlighetene*, men vi skriver bare *sannsynlighet* når det er klart utifra konteksten.

1.1.4 Sannsynlighetsmodell

Vi oppsummerer begrepene om et eksperiment, utfallsrom og sannsynlighet i et samlebegrep: en *sannsynlighetsmodell*.

Definisjon 1.1.9 — sannsynlighetsmodell

En *sannsynlighetsmodell* er en tripplett $(E, \Omega_E, P(S))$ hvor

$$E = \text{eksperimentet} \quad (1.50)$$

$$\Omega_E = \text{utfallsrommet} \quad (1.51)$$

$$P(S) = \text{sannsynlighet} \quad (1.52)$$

hvor

$$P(S) = \lim_{N \rightarrow \infty} f_S = \lim_{N \rightarrow \infty} \frac{N_S}{N} \quad (\text{i det lange løp}) \left(\frac{\# \text{gunstige}}{\# \text{mulige}} \right) \quad (1.53)$$

hvor $S \subseteq \Omega_E$ er en mulig begivenhet.

Vi skiller mellom to typer sannsynlighetsmodeller: **diskrete** og **kontinuerlige** sannsynlighetsmodeller.

Definisjon 1.1.10 — diskret sannsynlighetsmodell

En sannsynlighetsmodell $(E, \Omega_E, P(S))$ er *diskret* dersom utfallsrommet Ω_E er **tellbart**, dvs.:

$$\Omega_E = \{ u_1, u_2, u_3, u_4, \dots \} \quad (1.54)$$

Definisjon 1.1.11 — kontinuerlig sannsynlighetsmodell

En sannsynlighetsmodell $(E, \Omega_E, P(S))$ er *kontinuerlig* dersom utfallsrommet Ω_E **ikke** er **tellbart**.

K Utfallsrommet er *tellbart* dersom man kan **nummerere** utfallene, dvs. utfall nr. 1, utfall nr. 2, utfall nr. 3 osv.

K En diskret modell er *endelig*, dersom det er et endelig antall utfall.

K Vi skiller mellom disse to typene fordi det vil vise seg at man beregner sannsynligheter ganske forskjellig gitt den ene eller andre typen.

■ Eksempel 1.17 — diskrete vs kontinuerlige sannsynlighetsmodeller

Avgjør om sannsynlighetsmodellene er diskret eller kontinuerlig.

Modell:

Eksperiment E : En balansert terning kastes og antall øyne registreres.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, 5, 6 \}$

Modellen er **diskret** siden utfallsrommet er tellbart (det er 6 utfall).

Modell:

Eksperiment E : En balansert mynt slås inntil krone vises. Antall kast registreres.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, \dots \}$

Modellen er **diskret** siden utfallsrommet er tellbart (man kan *nummerere* utfallene selv om det er uendelig mange av dem).

Modell:

Eksperiment E : Varigheten til et prosjekt.

Utfallsrom Ω_E : $\Omega_E = [0, \infty)$

Modellen er **kontinuerlig** siden utfallsrommet *ikke* er tellbart (utfallene kan ikke nummereres).

Modell:

Eksperiment E : Prosentvis valgdeltakelse i et gitt demokrati.

Utfallsrom Ω_E : $\Omega_E = [0, 1]$

Modellen er **kontinuerlig** siden utfallsrommet *ikke* er tellbart (utfallene kan ikke nummereres).

Modell:

Eksperiment E : Sjekker om prisen til en aksje stiger eller synker ila neste time.

Utfallsrom Ω_E : $\Omega_E = \{ \text{'stiger'}, \text{'synker'} \}$

Modellen er **diskret** siden utfallsrommet er tellbart (det er 2 utfall).

■

Egenskaper ved diskrete sannsynlighetsmodeller

Den viktigste egenskapen ved en diskret sannsynlighetsmodell er at utfallssannsynlighetene p_u er *byggesteinene* i modellen.

Følgende setning sier at enhver begivenhetssannsynlighet er lik *summen* av utfallssannsynlighetene for alle utfallene som inngår i begivenhetene.¹⁷

Setning 1.1.2 — begivenhetssannsynlighet

La $(E, \Omega_E, P(S))$ være en **diskret** sannsynlighetsmodell hvor

$$\Omega_E = \{u_1, u_2, u_3, \dots\} \quad (1.55)$$

Da gjelder

$$P(S) = \sum_{u \in S} p_u \quad (1.56)$$

for alle mulige begivenheter $S \in \Omega_E$.

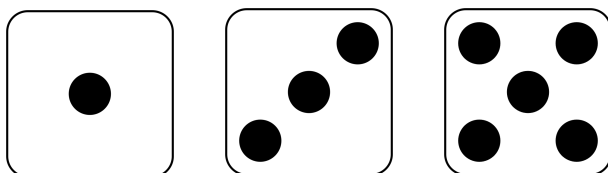


Denne setningen er ekstremt viktig for diskrete sannsynlighetsmodeller, siden den gir oss en enkel måte å beregne sannsynlighet på uten å gå via definisjonen (som vi så ifra eksempel 1.16 ikke var enkelt).

■ Eksempel 1.18 — sannsynlighet som sum av utfallssannsynligheter

Vi kan nå bruke setning 1.1.2 til å regne ut sannsynligheten for begivenheten $S = \{1, 3, 5\}$ fra eksempel 1.16:

$$P(S) = P(\{1, 3, 5\}) = \sum_{u \in S} p_u = p_1 + p_3 + p_5 = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} \quad (1.57)$$



Figur 1.11: Sannsynligheten for å få en 1er, 3er eller 5er.

■

¹⁷Symbolet $\sum_{u \in S}$ i definisjonen betyr ”summen over alle utfallene i begivenheten S ”.

■ Eksempel 1.19 — tre ubalanserte mynter slås en gang hver

Eksperiment E : Tre ubalanserte mynter kastes 1 gang hver, hvor resultatet for hver av dem registreres. Mynt nr. 1 er farget **rød**, mynt nr.2 er **blå** og mynt nr.3 er **gyllen**.

$$\text{Utfallsrom } \Omega_E: \quad \Omega_E = \left\{ \begin{array}{cccc} (k, k, k), & (k, k, m), & (k, m, k), & (k, m, m), \\ (m, k, k), & (m, k, m), & (m, m, k), & (m, m, m) \end{array} \right\}$$

Et utfall er her definert som en 3-tupple (triplett) $(-, -, -)$ hvor hver mynt tar enten verdien k for krone eller m for mynt.

Utfallet (k, k, m) betyr dermed at første mynt ble krone, andre mynt ble krone og tredje mynt ble mynt. Totalt består utfallsrommet av 8 utfall, og dermed har vi en *endelig* sannsynlighetsmodell.

Anta nå at vi vet at de respektive sannsynlighetene for mynt/krone for hver mynt er oppgitt i tabellen:

	mynt 1	mynt 2	mynt 3
Sannsynlighet for krone	$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{4}$
Sannsynlighet for mynt	$\frac{2}{3}$	$\frac{4}{5}$	$\frac{3}{4}$

Tabell 1.3: Sannsynlighetene for krone/mynt for hver av myntene.

Anta videre at **utfallssannsynlighetene** (liten p) fås ved å *multiplisere* sammen de tilhørende sannsynlighetene i tabell 1.3. For eksempel

$$P(k, k, m) = \frac{1}{3} \cdot \frac{1}{5} \cdot \frac{3}{4} = \frac{1}{20} \quad (1.58)$$

De andre utfallssannsynlighetene regnes ut på tilsvarende måte.

Spørsmål: Hva er sannsynligheten for at nøyaktig en av myntene viser krone?

Svar: Vi definerer *begivenheten* hvor nøykatig en krone forekommer som

$$S = \{ (k, m, m), (m, k, m), (m, m, k) \} \quad (1.59)$$

Siden vi har en *diskret* sannsynlighetsmodell, har vi fra setning 1.1.2

$$\begin{aligned} P(S) &= P(k, m, m) + P(m, k, m) + P(m, m, k) = \frac{1}{3} \cdot \frac{4}{5} \cdot \frac{3}{4} + \frac{2}{3} \cdot \frac{1}{5} \cdot \frac{3}{4} + \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{1}{4} \\ &= \frac{1}{5} + \frac{2}{20} + \frac{2}{15} = \frac{13}{30} \approx 0.433 \end{aligned}$$

■

Egenskaper ved kontinuerlige sannsynlighetsmodeller

En kontinuerlig sannsynlighetsmodell oppfører seg ganske annerledes enn en diskret. La oss se på et konkret eksempel for å få en bedre forståelse av hvordan en kontinuerlig sannsynlighetsmodell fungerer.

■ Eksempel 1.20 — temperatur på en gitt dag

Eksperiment E : Du registrerer temperaturen (målt i Kelvin K) ute en gitt dag på et gitt tidspunkt.

Utfallsrom Ω_E : $\Omega_E = [0, \infty)$

Et utfall er mao. et tall større enn det absolutte nullpunkt 0K. Utfallsrommet er dermed *ikke* tellbart, siden vi *ikke* kan nummerere utfallene. Vi har mao. en kontinuerlig sannsynlighetsmodell.¹⁸

Det neste som kommer nå kommer til å **overraske** og forvirre.

Spørsmål: Hva er sannsynligheten for at temperaturen 273.15 Kelvin (0 grader) inntreffer, dvs. hva er utfallssannsynligheten $p_{273.15}$?

Svar: Sannsynligheten for at temperaturen blir akkurat 273.15K er **null**, dvs.

$$p_{273.15} = 0 \quad (1.60)$$

Faktisk er *alle* utfallssannsynlighetene i en kontinuerlig sannsynlighetsmodell **null**.

■

Den overraskende konklusjonen fra eksempel 1.20 kan generaliseres til *alle* kontinuerlige sannsynlighetssmodeller:

Setning 1.1.3 — utfallssannsynlighet for kontinuerlig sannsynlighetsmodell

La $(E, \Omega_E, P(S))$ være en kontinuerlig sannsynlighetsmodell. Da er alle utfallssannsynlighetene **null**, dvs.

$$p_u = 0 \quad (1.61)$$

for alle utfall $u \in \Omega_E$.



Årsaken til at utfallssannsynlighetene alltid er null i et kontinuerlig utfallssrom, er litt løst sagt at det er for (altfor) mange utfall, så sannsynligheten for at nøykatig *ett* av dem inntreffer blir null.

¹⁸La oss prøve å nummerere utfallene: La utfall nr. 1 være 0. Hva skal utfall nr. 2 være? Det finnes ingen måte å nummerere de relle tallene på. Hvis man velger et tall si 0.0000001 som utfall nr. 2, så har man faktisk "hoppet" over alle tall mellom 0 og 0.0000001, som (faktisk) er like mange tall som hele tallinjen! At tallinjen $\mathbb{R}$ ikke kan nummereres er et dypt og fundamentalt resultat i grunnleggende matematikk, og er årsaken til at diskrete og kontinuerlige sannsynlighetsrom må behandles forskjellig.

Betyr setning 1.1.3 at vi ikke kan regne på sannsynligheter i kontinuerlige modeller? Svaret er **nei**, men det betyr at **utfallssannsynlighetene ikke** er **byggesteinene** for kontinuerlige modeller.

Byggesteinene i en kontinuerlig sannsynlighetsmodell er **intervaller**.

■ Eksempel 1.21 — temperatur på en gitt dag

Anta vi reformulerer spørsmålet i eksempel 1.20 til å omfatte et **intervall**:

Spørsmål: Hva er sannsynligheten for at temperaturen ligger mellom 273.15 Kelvin (0 grader) og 283.15 Kelvin (10 grader)?

Svar: Vi danner intervallet

$$I = [273.15, 283.15] \quad (1.62)$$

Intervallet I er en *begivenhet* og kan dermed estimeres ved hjelp av den eksperimentelle sannsynligheten.

Anta f.eks. at vi har målt temperaturen hver dag i et år, dvs. anta $N = 365$. Av de 365 dagene var temperaturen mellom 0 og 10 grader totalt $N_I = 137$ ganger. Den eksperimentelle sannsynligheten blir dermed

$$P(I) \approx f_I = \frac{N_I}{N} = \frac{137}{365} \approx 0.38 \quad (1.63)$$



Figur 1.12: Hva er sannsynligheten for at temperaturen havner mellom 0 og 10 grader?

■

- (K) Fra eksempel 1.21 innser vi at selv om utfallssannsynlighetene er alle null, kan et **intervall** ha sannsynlighet større enn null. Legg merke til at sannsynligheten blir **mindre** jo **kortere** intervallet blir.
- (K) Utfallssannsynligheter, eller “punktsannsynligheter”, for en kontinuerlig sannsynlighetsmodell er alltid null, mens “punktsannsynligheter” for en diskret sannsynlighetsmodell er ikke null.
- (K) Vi går nærmere inn på kontinuerlige sannsynlighetsmodeller i kapittel 5.

Uniform sannsynlighetsmodell

Vi skal nå se på en helt spesiell klasse av **endelige** sannsynlighetsmodeller hvor utfallssannsynlighetene er **like store**.

Slike sannsynlighetsmodeller kalles *uniforme* og forekommer svært ofte ved **spill og tipping**.

Definisjon 1.1.12 — uniform sannsynlighetsmodell

En endelig sannsynlighetsmodell $(E, \Omega_E, P(S))$ hvor Ω_E har totalt n utfall, dvs.

$$|\Omega_E| = n \quad (\Omega_E \text{ har } n \text{ utfall}) \quad (1.64)$$

kalles *uniform* dersom alle utfallene er like sannsynlige, dvs.:

$$p_u = \frac{1}{n} \quad (1.65)$$

for alle utfallene $u \in \Omega_E$.

Den viktigste egenskapen ved en uniform sannsynlighetsmodell er at det er nok å kjenne hvor *mange* utfall en begivenhet har for å regne ut sannsynligheten for at begivenheten inntreffer.

Setning 1.1.4 — fra sannsynlighetslære til kombinatorikk

Anta $(E, \Omega_E, P(S))$ er en *uniform* sannsynlighetsmodell med n utfall, og la S være en vilkårlig begivenhet med totalt $k \leq n$ utfall, dvs.

$$|S| = k \quad (S \text{ har } k \text{ utfall}) \quad (1.66)$$

Da er sannsynligheten for at begivenheten inntreffer gitt ved

$$P(S) = \frac{|S|}{|\Omega_E|} = \frac{k}{n} \quad \left(\frac{\text{\# gunstige utfall}}{\text{\# mulige utfall}} \right) \quad (1.67)$$

K For en *uniform* sannsynlighetsmodell kan en dermed beregne sannsynligheten for enhver begivenhet ved **å telle antall utfall begivenheten består av**. Denne egenskapen til uniforme sannsynlighetsmodeller medfører at beregning av sannsynligheten for en begivenhet blir et *telleproblem*. Dette fagfeltet kalles **kombinatorikk**, se vedlegg ??.

K Ligning 1.67 kan enkelt vises fra setningen om at begivenhetssannsynligheten er gitt som en *sum* av utfallssannsynlighetene (setning 1.1.2):

$$P(S) = \sum_{u \in S} p_u = p_{u_1} + p_{u_2} + \cdots + p_{u_k} = \overbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}^{k \text{ ganger}} = \frac{k}{n} \quad (1.68)$$

■ **Eksempel 1.22 — balansert terning gir uniform sannsynlighetsmodell**

Ekspirement E : En balansert terning kastes og antall øyne registreres.

Utfallsrom Ω_E : $\Omega_E = \{ 1, 2, 3, 4, 5, 6 \}$, $|\Omega_E| = 6$

Siden terningen er balansert kan vi anta at alle utfallssannsynlighetene er like, dvs. $p_u = \frac{1}{6}$ for utfall $u \in \Omega_E$. Vi har altså et **uniformt** utfallsrom.

Spørsmål: Hva er sannsynligheten for at vi får et oddetall **eller** mer enn 3 øyne?

Svar: Vi kaller begivenheten S . Hele oppgaven handler om å finne utfallene i S . La oss ta hvert utfall ett etter ett:

- $u = 1$: Siden 1 er et oddetall må $1 \in S$.
- $u = 2$: Siden 2 hverken er et oddetall *eller* er større enn 3, er $2 \notin S$.
- $u = 3$: Siden 3 er et oddetall må $3 \in S$.
- $u = 4$: Siden 4 er større enn 3 må $4 \in S$.
- $u = 5$: Siden 5 er et oddetall må $5 \in S$.
- $u = 6$: Siden 6 er større enn 3 må $6 \in S$.

Vi har dermed funnet at begivenheten S består av 5 utfall:

$$S = \underbrace{\{ 1, 3, 4, 5, 6 \}}_{5 \text{ utfall}} , \quad |S| = 5 \quad (1.69)$$

Siden S har $k = 5$ utfall kan vi konkludere fra setning 1.1.4 at

$$P(S) = \frac{|S|}{|\Omega_E|} = \frac{5}{6} \quad (1.70)$$

ved hjelp av telling.

■

K Fra eksempel 1.22 kan man få inntrykk av at telling er ganske enkelt, men som vi skal se i neste eksempel, så kan telling ofte bli ganske komplisert.



Figur 1.13: Balanserte terninger gir uniforme sannsynlighetsmodeller.

■ Eksempel 1.23 — trekking av pasienter

Anta det er totalt 9 pasienter på en avdeling på et sykehus er. Vi har fått i oppgave å trekke 3 tilfeldige pasienter fra de 9. Alle pasientene skal ha like stor sjanse for å bli trukket. Anta videre at

4 er **menn**, m_1, m_2, m_3, m_4

5 er **kvinner**, k_1, k_2, k_3, k_4, k_5

Eksperiment E : Vi trekker 3 tilfeldige pasienter fra de 9 pasientene.

Utfallsrom Ω_E : $\Omega_E = \{ (m_1, m_2, m_3), (m_1, k_1, m_2), \dots, (k_1, k_2, k_3) \}$

Legg altså merke til at et utfall er en *triplett*

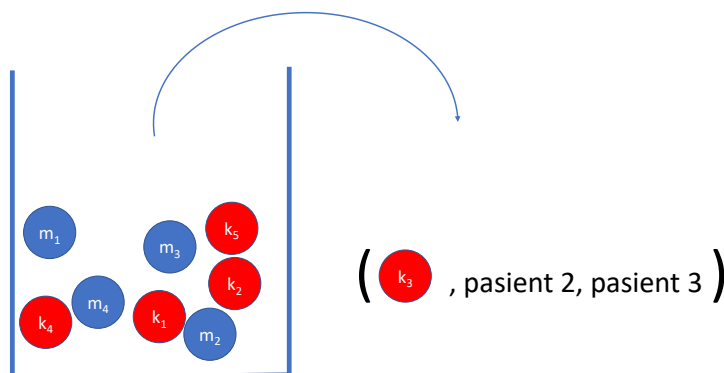
$$(\text{pasient nr. 1, pasient nr. 2, pasient nr. 3}) \quad (1.71)$$

Siden trekkingen er gjort *tilfeldig*, kan vi anta at hver pasient har like stor sannsynlighet for å bli trukket. Vi har altså et **uniformt** utfallsrom.

Spørsmål: Hvor mange utfall består utfallsrommet Ω_E av?

Svar: Vi kan modellere situasjonen som en urnemodell, hvor vi trekker 3 pasienter fra totalt 9 *uten* tilbakelegging:

$$|\Omega_E| = \binom{9}{3} = 84 \quad (\text{antall utfall}) \quad (1.72)$$



Figur 1.14: Urnemodellen hvor vi trekker 3 pasienter fra totalt 9 uten tilbakelegging.

Spørsmål: Hva sannsynligheten for å trekke 3 menn?
Kall begivenheten S_1 .

Svar: Vi kan modellere begivenheten som en urnemodell, hvor vi trekker 3 menn fra totalt 4 og 0 kvinner fra totalt 5, begge *uten* tilbakelegging. Siden er det er *to kategorier* pasienter så blir det et produkt av to binomialkoeffisienter:

$$|S_1| = \binom{4}{3} \binom{5}{0} = 4 \cdot 1 = 4 \quad (\text{antall gunstige utfall}) \quad (1.73)$$

Sannsynligheten er dermed gitt ved setning 1.67:

$$P(S_1) = \frac{|S_1|}{|\Omega_E|} = \frac{4}{84} = \frac{1}{21} = 0.048 \quad \left(\frac{\text{\# gunstige utfall}}{\text{\# mulige utfall}} \right) \quad (1.74)$$

Spørsmål: Hva sannsynligheten for å trekke 3 kvinner?
Kall begivenheten S_2 .

Svar: Vi kan modellere begivenheten som en urnemodell, hvor vi trekker 3 kvinner fra totalt 5 og 0 menn fra totalt 4, begge *uten* tilbakelegging:

$$|S_2| = \binom{5}{3} \binom{4}{0} = 10 \cdot 1 = 10 \quad (\text{antall gunstige utfall}) \quad (1.75)$$

Sannsynligheten er dermed gitt ved setning 1.67:

$$P(S_2) = \frac{|S_2|}{|\Omega_E|} = \frac{10}{84} = \frac{5}{42} = 0.119 \quad \left(\frac{\text{\# gunstige utfall}}{\text{\# mulige utfall}} \right) \quad (1.76)$$

Spørsmål: Hva sannsynligheten for å trekke 2 kvinner og 1 mann?
Kall begivenheten S_3 .

Svar: Vi kan modellere begivenheten som en urnemodell, hvor vi trekker 2 menn fra totalt 5 og 1 mann fra totalt 4, begge *uten* tilbakelegging:

$$|S_3| = \binom{5}{2} \binom{4}{1} = 10 \cdot 4 = 40 \quad (\text{antall gunstige utfall}) \quad (1.77)$$

Sannsynligheten er dermed gitt ved setning 1.67:

$$P(S_3) = \frac{|S_3|}{|\Omega_E|} = \frac{40}{84} = \frac{10}{21} = 0.476 \quad \left(\frac{\text{\# gunstige utfall}}{\text{\# mulige utfall}} \right) \quad (1.78)$$

1.2 Fundamentale regneregler

I dette avsnittet går vi gjennom 8 fundamentale regneregler for beregning av sannsynligheter for blant annet **sammensatte** begivenheter: ¹⁹

Setning 1.2.1 — fundamentale regneregler

La $(E, \Omega_E, P(S))$ være en sannsynlighetsmodell og anta at A og B er to vilkårlige begivenheter. Da gjelder følgende regneregler:

1. Utfallsrommet Ω_E

$$P(\Omega_E) = 1 \quad (1.79)$$

2. Komplementærsetningen

$$P(\bar{A}) = 1 - P(A) \quad (1.80)$$

3. Den tomme begivenheten $\emptyset$

$$P(\emptyset) = 0 \quad (1.81)$$

4. Den generelle addisjonssetningen

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.82)$$

5. Den spesielle addisjonssetningen

$$P(A \cup B) = P(A) + P(B) \quad \text{hvis } A \text{ og } B \text{ er disjunkte, dvs } A \cap B = \emptyset \quad (1.83)$$

6. Differansesetningen

$$P(A - B) = P(A) - P(B) \quad \text{hvis } A \subseteq B \quad (1.84)$$

7. Tvillingssetningene

$$P(A \cap B) = 1 - P(\bar{A} \cup \bar{B}) \quad \text{og} \quad P(A \cup B) = 1 - P(\bar{A} \cap \bar{B}) \quad (1.85)$$

8. Total sannsynlighet

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) \quad (1.86)$$

¹⁹Sammensatte begivenheter betyr at vi slår sammen to eller flere begivenheter vha. mengdeoperasjoner som komplementet $\bar{A}$, unionen $A \cup B$, snittet $A \cap B$ og differansen $A - B$. Les mer om dette i vedlegg ?? om mengdelære.

Gjennom hele dette avsnittet om regnereglene i sannsynlighetslære, skal vi bruke oss av følgende eksperiment for å belyse regelen. Vi kaller eksempelet **terningeksempelet**.

■ **Eksempel 1.24 — terningeksempelet**

Eksperiment E : En blå og rød terning kastes og vi leser av antall øyne hver av terningene viser. Anta at begge terningene er balanserte.

$$\text{Utfallsrom } \Omega_E: \quad \Omega_E = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\}$$

Dette eksperimentet har 36 utfall, hvor hvert utall er like sannsynlig siden begge terningene er balanserte. Vi har derfor et *uniformt* utfallsrom hvor utfallssannsynlighetene er gitt ved

$$p_u = \frac{1}{36} \quad (1.87)$$

for alle utfallene $u \in \Omega_E$. Siden vi har et uniformt utfallsrom så bruker vi lign.(1.67) i setning 1.1.4 og regne ut sannsynligheten for enhver begivenhet S :

$$P(S) = \frac{|S|}{|\Omega_E|} \quad \left(\frac{\text{\# gunstige utfall}}{\text{\# mulige utfall}} \right) \quad (1.88)$$



Figur 1.15: Blå og rød terning.

■

Regel 1: Minst ett utfall inntreffer

Denne regelen sier at sannsynligheten for at minst et utfall inntreffer er lik 1. Dette er ensbetydende med at et eksperiment *alltid* har et utfall.²⁰

Setning 1.2.2 — minst ett utfall inntreffer

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell. Da gjelder

$$P(\Omega_E) = 1 \quad (1.89)$$

Bevis. Anta vi gjennomfører eksperimentet E totalt N ganger. Siden utfallsrommet inneholder *alle* utfall, må antall ganger begivenheten $S = \Omega_E$ inntreffer være lik N , dvs. $N_{\Omega_E} = N$.

Den eksperimentelle sannsynligheten for begivenheten $S = \Omega_E$ er dermed gitt ved

$$f_{\Omega_E} = \frac{N_{\Omega_E}}{N} = \frac{N}{N} = 1 \quad (1.90)$$

Fra definisjonen av sannsynlighet i lign.(1.53) har vi dermed

$$P(\Omega_E) = \lim_{N \rightarrow \infty} f_{\Omega_E} = \lim_{N \rightarrow \infty} 1 = 1 \quad (1.91)$$

■

K Dersom vi har en **diskret** sannsynlighetsmodell, sier vi gjerne at **summen utfallssannsynlighetene er lik 1**. Dette kommer som en direkte følge av setning 1.1.2 lign.(1.56) med $S = \Omega_E = \{u_1, u_2, u_3, u_4, \dots\}$

$$P(\Omega_E) = \sum_{u \in \Omega_E} p_u = p_{u_1} + p_{u_2} + p_{u_3} + p_{u_4} + \dots = 1 \quad (1.92)$$

■ **Eksempel 1.25** I **terningeksempelet** 1.24, la $S = \Omega_E$. Siden vi har 36 utfall totalt, gir lign.(1.88)

$$P(\Omega_E) = \frac{|S|}{|\Omega_E|} = \frac{36}{36} = 1 \quad (1.93)$$

Vi kan også bruke lign.(1.92):

$$P(\Omega_E) = p_1 + p_2 + \dots + p_{36} = \frac{1}{36} + \frac{1}{36} + \dots + \frac{1}{36} = \frac{36}{36} = 1 \quad (1.94)$$

■

²⁰Hvis eksperimentet ikke ender i et utfall, er det ikke et eksperiment, ifølge definisjon 1.1.1.

Regel 2: Komplementærsetningen

Komplementsetningen er svært viktig og gir sammenhengen mellom en begivenhet og den *motsatte* begivenheten:

Setning 1.2.3 — komplementærsetningen

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell og la A være en begivenhet. Da gjelder

$$P(\bar{A}) = 1 - P(A) \quad (1.95)$$

Bevis. Anta vi gjennomfører eksperimentet E totalt N ganger. La N_A være antall ganger begivenhet A inntreffer. Når eksperimentet gir et utfall som *ikke* er i A , da må det nødvendigvis være i komplementet $\bar{A}$. Siden A ikke inntreffer $N - N_A$ ganger må $N_{\bar{A}} = N - N_A$.

Den eksperimentelle sannsynligheten for begivenheten $S = \bar{A}$ er dermed gitt ved

$$f_{\bar{A}} = \frac{N_{\bar{A}}}{N} = \frac{N - N_A}{N} = \left(\frac{N}{N} - \frac{N_A}{N} \right) = 1 - \frac{N_A}{N} = 1 - f_A \quad (1.96)$$

Fra definisjonen av sannsynlighet i lign.(1.53) med $S = \bar{A}$:

$$P(\bar{A}) = \lim_{N \rightarrow \infty} f_{\bar{A}} = \lim_{N \rightarrow \infty} (1 - f_A) = \lim_{N \rightarrow \infty} 1 - \lim_{N \rightarrow \infty} f_A = 1 - P(A) \quad (1.97)$$

■

Regel 3: Den tomme begivenheten inntreffer ikke

Den tomme mengden inntreffer ikke som en følge av komplementsetningen:

Setning 1.2.4 — den tomme begivenheten inntreffer ikke

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell. Da gjelder

$$P(\emptyset) = 0 \quad (1.98)$$

Bevis. Siden $\emptyset = \overline{\Omega_E}$, kan vi bruke komplementsetningen:

$$P(\emptyset) = P(\overline{\Omega_E}) = 1 - P(\Omega_E) = 1 - 1 = 0 \quad (1.99)$$

■

■ **Eksempel 1.26** I **terningeksempelet** 1.24:

Spørsmål: Hva er sannsynligheten for å få *minst* en 1er?

Svar: La A være begivenheten for å få minst en 1er. Da er A gitt ved:

$$A = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & \cdot & \cdot & \cdot & \cdot & \cdot \\ (3,1), & \cdot & \cdot & \cdot & \cdot & \cdot \\ (4,1), & \cdot & \cdot & \cdot & \cdot & \cdot \\ (5,1), & \cdot & \cdot & \cdot & \cdot & \cdot \\ (6,1), & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\}$$

Vi har dermed at antall gunstige utfall er 11, dvs. $|A| = 11$. Sannsynligheten er dermed: $P(A) = \frac{|A|}{|\Omega_E|} = \frac{11}{36}$.

Spørsmål: Hva er sannsynligheten for å *ikke* få en 1er?

Svar: Begivenheten å *ikke* få en 1er, er *komplementet* til A , dvs. $\bar{A}$. Vi kan dermed bruke komplementsetningen:

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{11}{36} = \frac{36}{36} - \frac{11}{36} = \frac{25}{36} \quad (1.100)$$

Vi kan alternativt bruke telling. Siden $\bar{A}$ er gitt ved:

$$\bar{A} = \left\{ \begin{array}{cccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ \cdot & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ \cdot & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ \cdot & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ \cdot & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\}$$

har vi at antall gunstige utfall er 25, dvs. $|\bar{A}| = 25$. Sannsynligheten er dermed $P(\bar{A}) = \frac{|\bar{A}|}{|\Omega_E|} = \frac{25}{36}$. ■

Regel 4: Den generelle addisjonssetningen

Sannsynligheten for unionen mellom to begivenheter er ikke nødvendigvis lik summen av sannsynlighetene:

Setning 1.2.5 — den generelle addisjonssetningen

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell og la A og B være to vilkårlige begivenheter. Da gjelder

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.101)$$

Bevis. Anta vi gjennomfører eksperimentet E totalt N ganger. Vi definerer størrelsene:

$$N_A = \text{antall ganger begivenhet } A \text{ inntreffer} \quad (1.102)$$

$$N_B = \text{antall ganger begivenhet } B \text{ inntreffer} \quad (1.103)$$

$$N_{A \cup B} = \text{antall ganger begivenhet } A \text{ eller } B \text{ inntreffer} \quad (1.104)$$

$$N_{A \cap B} = \text{antall ganger begivenhet } A \text{ og } B \text{ inntreffer} \quad (1.105)$$

Da har vi følgende naturlige sammenheng:

$$N_{A \cup B} = N_A + N_B - N_{A \cap B} \quad (1.106)$$

Forklaring: La u være et utfall i $A \cap B$. Siden u er både i A og i B , så har vi telt dette utfallet to ganger - en gang i N_A og en gang i N_B . Derfor må vi trekke ifra $N_{A \cap B}$ i lign.(1.106).

Den eksperimentelle sannsynligheten for begivenheten $S = A \cup B$ er dermed gitt ved

$$f_{A \cup B} = \frac{N_{A \cup B}}{N} = \frac{N_A + N_B - N_{A \cap B}}{N} = f_A + f_B - f_{A \cap B} \quad (1.107)$$

Fra definisjonen av sannsynlighet i lign.(1.53) med $S = \bar{A}$:

$$P(A \cup B) = \lim_{N \rightarrow \infty} f_{A \cup B} = \lim_{N \rightarrow \infty} (f_A + f_B - f_{A \cap B}) \quad (1.108)$$

$$= \lim_{N \rightarrow \infty} f_A + \lim_{N \rightarrow \infty} f_B - \lim_{N \rightarrow \infty} f_{A \cap B} = P(A) + P(B) - P(A \cap B) \quad (1.109)$$

■



K Vi har inkludert dette beviset for å forklare hvorfor man trekker ifra $P(A \cap B)$ i lign.(1.101), nemlig at man teller et utfall i snittet *to ganger* hvis ikke.

■ **Eksempel 1.27** I **terningeksempelet** 1.24:

Spørsmål: Hva er sannsynligheten for at **summen** av terningene er lik 7? Kall denne begivenheten A

Svar: Begivenheten A er gitt ved:

$$A = \left\{ \begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & (1,6), \\ \cdot & \cdot & \cdot & \cdot & \cdot & (2,5), & \cdot \\ \cdot & \cdot & \cdot & (3,4), & \cdot & \cdot & \cdot \\ \cdot & \cdot & (4,3), & \cdot & \cdot & \cdot & \cdot \\ \cdot & (5,2), & \cdot & \cdot & \cdot & \cdot & \cdot \\ (6,1) & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\} \quad (1.110)$$

Vi har dermed at antall gunstige utfall er 6, dvs. $|A| = 6$. Sannsynligheten er dermed: $P(A) = \frac{|A|}{|\Omega_E|} = \frac{6}{36} = \frac{1}{6}$.

Spørsmål: Hva er sannsynligheten for at minst en av terningene får en 3er? Kall denne begivenheten B .

Svar: Begivenheten B er gitt ved:

$$B = \left\{ \begin{array}{ccccccc} \cdot & \cdot & (1,3), & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & (2,3), & \cdot & \cdot & \cdot & \cdot \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), & \\ \cdot & \cdot & (4,3), & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & (5,3), & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & (6,3) & \cdot & \cdot & \cdot & \cdot \end{array} \right\}$$

Vi har dermed at antall gunstige utfall er 11, dvs. $|\overline{B}| = 11$. Sannsynligheten er dermed $P(B) = \frac{|B|}{|\Omega_E|} = \frac{11}{36}$.

Spørsmål: Hva er sannsynligheten for at **summen** av terningene er lik 7 **og** minst en av terningene får en 3er?

Svar: Dette er begivenheten $A \cap B$, som er gitt ved

$$A \cap B = \left\{ \begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & (3,4), & \cdot & \cdot & \cdot \\ \cdot & \cdot & (4,3) & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\}$$

Vi har dermed at antall gunstige utfall er 2, dvs. $|A \cap B| = 2$. Sannsynligheten er dermed $P(A \cap B) = \frac{|A \cap B|}{|\Omega_E|} = \frac{2}{36} = \frac{1}{18}$.

Spørsmål: Hva er sannsynligheten for at **summen** av terningene er lik 7 **eller** minst en av terningene får en 3er? Bruk den generelle addisjonssetningen.

Svar: Siden $A \cap B$ har to utfall, må vi passe på at disse ikke telles to ganger. Den generelle addisjonssetningen gir:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.111)$$

$$= \frac{6}{36} + \frac{11}{36} - \frac{2}{36} = \frac{6+11-2}{36} = \frac{15}{36} \quad (1.112)$$

Vi kan alternativt bruke telling. Siden $A \cup B$ er gitt ved:

$$A \cup B = \left\{ \begin{array}{cccccc} \cdot & \cdot & (1,3), & \cdot & \cdot & (1,6), \\ \cdot & \cdot & (2,3), & \cdot & (2,5), & \cdot \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ \cdot & \cdot & (4,3), & \cdot & \cdot & \cdot \\ \cdot & (5,2), & (5,3), & \cdot & \cdot & \cdot \\ (6,1), & \cdot & (6,3) & \cdot & \cdot & \cdot \end{array} \right\}$$

Vi har dermed at antall gunstige utfall er 15, dvs. $|A \cup B| = 15$. Sannsynligheten er dermed $P(A \cup B) = \frac{|A \cup B|}{|\Omega_E|} = \frac{15}{36}$. ■



Figur 1.16: Blå og rød terning.

Regel 5: Den spesielle addisjonssetningen

Dersom to begivenheter er **disjunkte**, så er sannsynligheten for unionen lik summen av sannsynlighetene:

Setning 1.2.6 — den spesielle addisjonssetningen

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell og la A og B være to **disjunkte** begivenheter. Da gjelder

$$P(A \cup B) = P(A) + P(B) \quad \text{hvis } A \text{ og } B \text{ er disjunkte, dvs } A \cap B = \emptyset \quad (1.113)$$

Bevis. Den spesielle addisjonssetningen kommer som en direkte følge av den generelle addisjonssetningen og at $P(\emptyset) = 0$:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(\emptyset) \quad (1.114)$$

$$= P(A) + P(B) - 0 = P(A) + P(B) \quad (1.115)$$

■

■ **Eksempel 1.28** I **terningeksempelet** 1.24:

Spørsmål: Hva er sannsynligheten for at summen av terningene er lik 7 **eller** lik 3 hhv.?

Svar: La A være begivenheten for at summen er lik 7 og B er begivenheten for at summen er lik 3. Da er

$$P(A) = \frac{6}{36}, \quad P(B) = \frac{2}{36} \quad (1.116)$$

Legg merke til at $P(B) = \frac{|B|}{|\Omega_E|} = \frac{2}{36}$ siden det kun er to utfall som gir sum 3.

$$A \cup B = \left\{ \begin{array}{cccccc} \cdot & (1, 2) & \cdot & \cdot & \cdot & (1, 6), \\ (2, 1) & \cdot & \cdot & \cdot & (2, 5), & \cdot \\ \cdot & \cdot & \cdot & (3, 4), & \cdot & \cdot \\ \cdot & \cdot & (4, 3), & \cdot & \cdot & \cdot \\ \cdot & (5, 2), & \cdot & \cdot & \cdot & \cdot \\ (6, 1) & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\} \quad (1.117)$$

Siden A og B er **disjunkte**, så gir den spesielle addisjonssetningen:

$$P(A \cup B) = P(A) + P(B) = \frac{6}{36} + \frac{2}{36} = \frac{8}{36} \quad (1.118)$$

■

Regel 6: Differansen mellom to begivenheter

Differansen mellom to begivenheter oppfører seg akkurat som navnet tilsier:

Setning 1.2.7 — differansesetningen

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell og la $A \subseteq B$. Da gjelder

$$P(A - B) = P(A) - P(B) \quad (1.119)$$

Bevis. Siden $A \subseteq B$, så må $A - B$ og B være **disjunkte** (siden vi har trukket vekk B). Den spesielle addisjonssetningen gir dermed

$$P(A) = P((A - B) \cup B) = P(A - B) + P(B) \quad (1.120)$$

$$\Updownarrow$$

$$P(A - B) = P(A) - P(B) \quad (1.121)$$

■

■ **Eksempel 1.29** I **terningeksempelet** 1.24:

Spørsmål: Hva er sannsynligheten for at summen av terningene er lik 7 **unntatt** hvor en av terningene er lik 3 hhv.?

Svar: La A være begivenheten for at summen er lik 7 og B er begivenheten for at summen er lik 7 hvor en av terningene er lik 3. Da er

$$P(A) = \frac{6}{36}, \quad P(B) = \frac{2}{36} \quad (1.122)$$

Legg merke til at $P(B) = \frac{|B|}{|\Omega_E|} = \frac{2}{36}$ siden det er bare to utfall med sum lik 7 som har 3 i seg.

$$A - B = \left\{ \begin{array}{cccccc} \cdot & \cdot & \cdot & \cdot & \cdot & (1,6), \\ \cdot & \cdot & \cdot & \cdot & (2,5), & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & (5,2), & \cdot & \cdot & \cdot & \cdot \\ (6,1) & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\}$$

Siden $A \subseteq B$, så gir differansesetningen:

$$P(A - B) = P(A) - P(B) = \frac{6}{36} - \frac{2}{36} = \frac{4}{36} \quad (1.123)$$

■

Regel 7: Tvillingsetningene

Tvillingsetningene er fine å ha dersom vi kjenner sannsynlighetene til $\bar{A}$ og $\bar{B}$ istedet for A og B .

Setning 1.2.8 — tvillingsetningene

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell og la A og B være to vilkårlige begivenheter. Da gjelder

$$P(A \cap B) = 1 - P(\bar{A} \cup \bar{B}) \quad \text{og} \quad P(A \cup B) = 1 - P(\bar{A} \cap \bar{B}) \quad (1.124)$$

Bevis. For to mengder A og B gjelder alltid:

$$\bar{A} \cap \bar{B} = \overline{A \cup B} \quad (1.125)$$

Dermed fås:

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) \quad (1.126)$$

hvor komplementærsetningen, dvs. lign. (1.2.3), ble brukt ved siste likhet. ■

■ Eksempel 1.30 I **terningeksempelet** 1.24:

Spørsmål: La A være begivenheten hvor summen er større enn 3 og B være begivenheten hvor summen er mindre enn 11.

Finn sannsynligheten for $A \cap B$ ved hjelp av tvillingsetningen for snitt.

Svar: De omvendte begivenhetene er gitt ved:

$$\bar{A} = \text{summen er mindre enn 3} \quad (1.127)$$

$$\bar{B} = \text{summen er større enn 11} \quad (1.128)$$

Sannsynlighetene for $\bar{A}$ og $\bar{B}$ er enkle å finne siden de har få gunstige utfall:

$$P(\bar{A}) = p_{(1,1)} = \frac{1}{36}, \quad P(\bar{B}) = p_{(6,6)} = \frac{1}{36} \quad (1.129)$$

Tvillingsetningen for union kan dermed benyttes:

$$\begin{aligned} P(A \cap B) &= 1 - P(\bar{A} \cup \bar{B}) = 1 - P(\bar{A}) - P(\bar{B}) \\ &= 1 - \frac{1}{36} - \frac{1}{36} = \frac{34}{36} \end{aligned} \quad (1.130)$$

hvor vi har brukt at $\bar{A}$ og $\bar{B}$ er disjunkte. ■

Regel 8: Total sannsynlighet

Total sannsynlighet er ofte brukt, når vi kan dele opp utfallsrommet i **to kategorier**.

Setning 1.2.9 — total sannsynlighet

Anta $(E, \Omega_E, P(S))$ er en sannsynlighetsmodell og la A og B være to vilkårlige begivenheter. Da gjelder

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) \quad (1.131)$$

Bevis. Siden $\Omega_E = B \cup \bar{B}$ kan vi anse at begivenheten B deler opp utfallsrommet i **to kategorier**:

$$B = \text{kategori } B \quad (\text{utfallene som er i } B) \quad (1.132)$$

$$\bar{B} = \text{kategori } \bar{B} \quad (\text{utfallene som ikke er i } B) \quad (1.133)$$

Anta vi gjennomfører eksperimentet E totalt N ganger. Vi definerer størrelsene:

$$N_A = \text{antall ganger begivenhet } A \text{ inntreffer} \quad (1.134)$$

$$N_{A \cap B} = \text{antall ganger begivenhet } A \text{ inntreffer med utfall i kategori } B \quad (1.135)$$

$$N_{A \cap \bar{B}} = \text{antall ganger begivenhet } A \text{ inntreffer med utfall i kategori } \bar{B} \quad (1.136)$$

Da har vi følgende naturlige sammenheng:

$$N_A = N_{A \cap B} + N_{A \cap \bar{B}} \quad (1.137)$$

Forklaring: Siden $B \cap \bar{B} = \emptyset$, er kategoriene disjunkte, dvs. et utfall kan ikke tilhøre begge kategoriene samtidig. Dermed vet vi at ethvert utfall som telles i N_A enten tilhører $N_{A \cap B}$ eller $N_{A \cap \bar{B}}$, men ikke begge.

Den eksperimentelle sannsynligheten for begivenheten $S = A$ er dermed gitt ved

$$f_A = \frac{N_A}{N} = \frac{N_{A \cap B} + N_{A \cap \bar{B}}}{N} = f_{A \cap B} + f_{A \cap \bar{B}} \quad (1.138)$$

Fra definisjonen av sannsynlighet i lign.(1.53) med $S = \bar{A}$:

$$P(A) = \lim_{N \rightarrow \infty} f_A = \lim_{N \rightarrow \infty} (f_{A \cap B} + f_{A \cap \bar{B}}) \quad (1.139)$$

$$= \lim_{N \rightarrow \infty} f_{A \cap B} + \lim_{N \rightarrow \infty} f_{A \cap \bar{B}} = P(A \cap B) + P(A \cap \bar{B}) \quad (1.140)$$



■ **Eksempel 1.31** I **terningeksempelet** 1.24:

Spørsmål: La B være begivenheten hvor summen av øynene er et oddetall. La A være begivenheten hvor antall øyne er delelig med 3.

Bruk loven om total sannsynlighet for å finne $P(A)$.

Svar: Vi har at A må innholde alle kombinasjoner hvor summen er lik 3, 6, 9 eller 12, siden dette er summene som er delelig med 3:

$$A = \left\{ \begin{array}{cccccc} \cdot & (1,2), & \cdot & \cdot & (1,5), & \cdot \\ (2,1), & \cdot & \cdot & (2,4) & \cdot & \cdot \\ \cdot & \cdot & (3,3), & \cdot & \cdot & (3,6), \\ \cdot & (4,2), & \cdot & \cdot & (4,5), & \cdot \\ (5,1), & \cdot & \cdot & (5,4), & \cdot & \cdot \\ \cdot & \cdot & (6,3), & \cdot & \cdot & (6,6) \end{array} \right\}$$

Hvis vi deler opp utfallene i A til de som tilhører kategori B og kategori $\bar{B}$, dvs. hvor summen er et oddetall og partall hhv., får vi:

$$A = \left\{ \begin{array}{cccccc} \cdot & (1,2), & \cdot & \cdot & (1,5), & \cdot \\ (2,1), & \cdot & \cdot & (2,4) & \cdot & \cdot \\ \cdot & \cdot & (3,3), & \cdot & \cdot & (3,6), \\ \cdot & (4,2), & \cdot & \cdot & (4,5), & \cdot \\ (5,1), & \cdot & \cdot & (5,4), & \cdot & \cdot \\ \cdot & \cdot & (6,3), & \cdot & \cdot & (6,6) \end{array} \right\}$$

hvor **blå** farge betyr kategori B og **grønn** farge betyr kategori $\bar{B}$. Vi har dermed ved telling av antall utfall i hver kategori:

$$P(A \cap B) = \frac{6}{36} \quad (\text{de blå utfallene}) \quad (1.141)$$

$$P(A \cap \bar{B}) = \frac{6}{36} \quad (\text{de grønne utfallene}) \quad (1.142)$$

Loven om total sannsynlighet gir dermed

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = \frac{6}{36} + \frac{6}{36} = \frac{12}{36} = \frac{1}{3} \quad (1.143)$$

■

■ Eksempel 1.32 — gravid og kjønn

Eksperiment E : Anta vi trekker en tilfeldig person blant alle kvinner og menn. Vi observerer så kjønn og hvorvidt personen er gravid eller ikke.

Utfallsrom Ω_E : $\Omega = \{ (m, g), (m, \bar{g}), (k, g), (k, \bar{g}) \}$
 hvor m = mann, k = kvinne, g = gravid og $\bar{g}$ = ikke gravid.

Spørsmål: Anta at det er lik sannsynlighet for å trekke en mann eller kvinne, og anta at sannsynligheten for å trekke en gravid person er 0.03.

Finn sannsynligheten at vi trekker en mann **eller** en gravid person.

Svar: Vi definerer først begivenhetene

$$M = \text{en mann trekkes} \quad (1.144)$$

$$K = \text{en kvinne trekkes} \quad (1.145)$$

$$G = \text{en gravid person trekkes} \quad (1.146)$$

Fra opplysningene kan vi konkludere at

$$P(M) = 0.5 \quad (1.147)$$

$$P(K) = 0.5 \quad (1.148)$$

$$P(G) = 0.03 \quad (1.149)$$

Siden menn ikke kan bli gravide, må M og G være disjunkte. Den spesielle addisjonssetningen gir dermed

$$P(M \cup G) = P(M) + P(G) = 0.5 + 0.03 = 0.53 \quad (1.150)$$

■



Figur 1.17: Gravid.

■ Eksempel 1.33 — busslogistikk

En del ansatte ved Vestbase i Kristiansund tar bussen til jobb. En av disse, Ole Hansen, tar bussrute 1, men må skifte til rute 2 underveis.

Eksperiment E : Ole Hansen tar første rute 1 og observerer om bussen er ledig eller ikke, deretter tar han rute 2 og observerer om den bussen er ledig.

Utfallsrom Ω_E : $\Omega = \{ (l_1, l_2), (l_1, \bar{l}_2), (\bar{l}_1, l_2), (\bar{l}_1, \bar{l}_2) \}$
 hvor l_1 = bussrute 1 er ledig, l_2 = bussrute 2 er ledig. Tilsvarende betyr $\bar{l}_1$ og $\bar{l}_2$ ikke ledig på ruten hhv.

Anta Ole har anslått (gjennom lang erfaring) at:

Sannsynligheten for at det er ingen ledige busser på rute 1 er 0.6. (1.151)

Sannsynligheten for at det er ingen ledige busser på rute 2 er 0.7. (1.152)

Sannsynligheten for at det er ledig på begge rutene samtidig er 1 er 0.2. (1.153)

Spørsmål: Hva er sannsynligheten for at Ole Hansen finner ledige sitteplasser på **minst en** av rutene?

Svar: Vi definerer først begivenhetene

$$L_1 = \text{rute 1 er ledig} \quad (1.154)$$

$$L_2 = \text{rute 2 er ledig} \quad (1.155)$$

Fra opplysningene kan vi konkludere at

$$P(\bar{L}_1) = 0.6 \quad P(L_1) = 1 - P(\bar{L}_1) = 0.4 \quad (1.156)$$

$$P(\bar{L}_2) = 0.7 \quad P(L_2) = 1 - P(\bar{L}_2) = 0.3 \quad (1.157)$$

$$P(L_1 \cap L_2) = 0.2 \quad (1.158)$$

Begivenheten at Ole finner ledige plasser på minst en av rutene, er det samme som at rute 1 er ledig **eller** rute 2 er ledig (eller begge). Dette er unionen mellom L_1 og L_2 , så den generelle addisjonssetningen gir

$$P(L_1 \cup L_2) = P(L_1) + P(L_2) - P(L_1 \cap L_2) \quad (1.159)$$

$$= 0.4 + 0.3 - 0.2 = 0.5 \quad (1.160)$$

Spørsmål: Hva er sannsynligheten for at Ole Hansen finner ledige sitteplasser på rute 1, men **ikke** på rute 2?

Svar: Denne oppgaven er vanskelig, siden vi må uttrykke begivenheten det spørres om ved hjelp av begivenhetene vi allerede kjenner sannsynlighetene for.

Begivenheten at Ole Hansen finner ledige sitteplasser på rute 1, men **ikke** på rute 2, er gitt ved **differansen** mellom L_1 og $L_1 \cap L_2$, siden vi da trekker ifra de utfallene hvor rute 2 er ledig (og dermed sitter igjen med de som ikke er ledig på rute 2).

Differansesetningen gir dermed

$$P(L_1 - L_1 \cap L_2) = P(L_1) - P(L_1 \cap L_2) \quad (1.161)$$

$$= 0.4 - 0.2 = 0.2 \quad (1.162)$$

Spørsmål: Hva er sannsynligheten for at Ole Hansen **verken** finner ledige sitteplasser på rute 1 **eller** rute 2?

Svar: Dett er begivenheten $\bar{L}_1 \cap \bar{L}_2$, dvs. hvor både rute 1 **og** rute 2 ikke er ledige.

Tvillingsetningen for union gir dermed (snudd)

$$P(\bar{L}_1 \cap \bar{L}_2) = 1 - P(L_1 \cup L_2) = 1 - 0.5 = 0.5 \quad (1.163)$$

$$= 0.4 - 0.2 = 0.2 \quad (1.164)$$



Figur 1.18: Buss i Kristiansund.