

Litt sannsynlighetsregning

– mest for deg som forsøker å lære risikoanalyse

Marvin Rausand

September 2021

1 Innledning

Dette notatet er skrevet som et tillegg til «Risikoanalyse - teori og metoder» 2.utg. (Fagbokforlaget, 2022). Notatet gir en kortfatta innføring i sannsynlighetsregning. Hvis du trenger en mer omfattende innføring, bør du lese ei av de mange lærebøkene i sannsynlighetsregning og/eller statistikk som er utgitt.

Risiko uttrykkes ofte som en funksjon av sannsynlighet og konsekvens. For å forstå risikobegrepet og for å kunne utføre en risikoanalyse, trenger du en viss forståelse for hva sannsynlighet er og hvordan du kan beregne sannsynligheten til uønska hendelser.

2 Viktige begrep

Dette avsnittet innfører og forklarer noen av de grunnleggende begrepene som brukes når du skal beregne sannsynligheten for hendelser.

2.1 Utfall

For å bestemme sannsynligheten for en hendelse, kan du noen ganger betrakte et fenomen i ei viss tid og studere resultatene fra ei rekke «forsøk». Eksempler på slike forsøk er fly som tar av og lander på en flyplass, biler som kjører på en viss vegstrekning, og operasjoner av en viss type som utføres på et sjukehus. Hvert forsøk resulterer i et *utfall*¹ som vi betegner med bokstaven e .

2.2 Stokastisk forsøk

Sjøl om du kan beskrive alle mulige utfall av et forsøk, kan du likevel ikke med sikkerhet forutsi hvilket utfall som vil inntreffe før forsøket er gjennomført. Dersom forsøket kan gjentas under omtrent like forhold, kalles det et *stokastisk forsøk*² – også i tilfelle der

¹Engelsk: Event.

²Ordet stokastisk kommer fra gresk og betyr tilfeldig.

du bare kan *forestille* deg at det er mulig med noenlunde like gjentakelser, men hvor det i virkeligheten er umulig.

2.3 Utfallsrommet

Mengden av alle de mulige utfallene fra et stokastisk forsøk kaller vi *utfallsrommet*,³ og betegner dette med S . Utfallsrommet kan ha et tellbart eller ikke-tellbart antall utfall. Når utfallene er tellbare, har vi et *diskret* utfallsrom. Når utfallet kan ta en hvilken som helst verdi i et begrensa eller ubegrensa intervall, er utfallsrommet *kontinuerlig*. Et utfallsrom med et begrensa antall, n , utfall kan du skrive utfallsrommet som $S = \{e_1, e_2, \dots, e_n\}$.

Eksempel 2.1

Et stokastisk forsøk består i å kaste en mynt og observere om utfallet blir kron (K) eller mynt (M). Utfallsrommet er da

$$S = \{e_1, e_2\} = \{K, M\}$$

der $e_1 = K$ betyr at utfallet av kastet er kron og $e_2 = M$ at det er mynt. □

2.4 Hendelser

En hendelse, E , er en bestemt mengde av mulige utfall i utfallsrommet S . Vi skriver $E \subset S$ der \subset betyr delmengde. De fleste hendelser består av mange utfall, men et enkelt utfall e oppfattes også som en hendelse.

Et forsøk ender alltid i ett og bare ett utfall e . Når dette utfallet er med i hendelsen E , sier vi at hendelsen E inntreffer.

Eksempel 2.2

Vi er interessert i å observere resultatet av en fotballkamp mellom lag A og lag B , målt i antall mål for hvert lag. I dette tilfellet er utfallsrommet

$$S = \{(0, 0), (0, 1), (1, 0), (1, 1), (0, 2), (2, 0), (1, 2), (2, 1), \dots\}$$

Utfallet $(2, 0)$ betyr at lag A vinner kampen «to-null». Lag A kan vinne kampen på mange måter. Hvis vi får et utfall e som er med i mengden $E_1 = \{(1, 0), (2, 0), (2, 1), \dots\}$, så har lag A vunnet. Mengden E_1 kan uttrykkes som «lag A vinner kampen» og er en delmengde i S . Når kampen, for eksempel, gir utfallet $e = (1, 0)$, som er med i hendelsen E_1 , sier vi at hendelsen E_1 har inntrefft og «lag A har vunnet kampen».

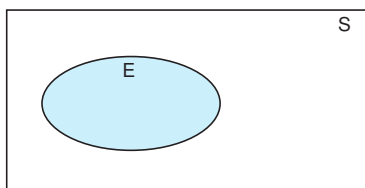
En annen hendelse er «maksimalt to mål i kampen» som er gitt ved mengden $E_2 = \{(0, 0), (0, 1), (1, 0), (1, 1), (0, 2), (2, 0)\}$.

Legg merke til at for eksempel utfallet $e = (1, 0)$ er med i både E_1 og E_2 , slik at begge de to hendelsene inntreffer. □

2.5 Snitt og union

Vi har sett at en hendelse er en *mengde* av alle utfall som oppfyller visse kriterier. Vi skal nå se på hvordan du kan kombinere to eller flere hendelser.

³Engelsk: Sample space.



Figur 1: Venn-diagram med en hendelse, E .

2.5.1 Venn-diagram

Et Venn-diagram, som vist i figur 1, brukes ofte til å illustrere utfallsrommet og de ulike hendelsene. I diagrammet representerer rektangelet utfallsrommet S . De enkelte utfallene av forsøket markeres som punkt i rektangelet. Mengden E inneholder alle utfallene som inngår i hendelsen, E .

2.6 Komplementærhendelser

La E^* være at hendelsen E ikke inntreffer. E^* inneholder alle de utfallene i S som ikke er med i E og kalles komplementærhendelsen til E . I noen tilfelle kan det være enklere å bestemme E^* enn hendelsen E .

Komplementærhendelsen til utfallsrommet S er, ifølge definisjon av S , en *tom mengde*, fordi S inneholder alle de mulige utfallene.

2.7 Snittet av to hendelser

Snittet av de to hendelsene, E_1 og E_2 , skrives med symbolet $E_1 \cap E_2$, og er hendelsen som består av alle utfall som er *felles* for E_1 og E_2 (dvs, som er med i både E_1 og E_2). Se figur 2(a).

Eksempel 2.3

I eksempel 2.2 får vi snittet

$$E_1 \cap E_2 = \{(1, 0), (2, 0)\}$$

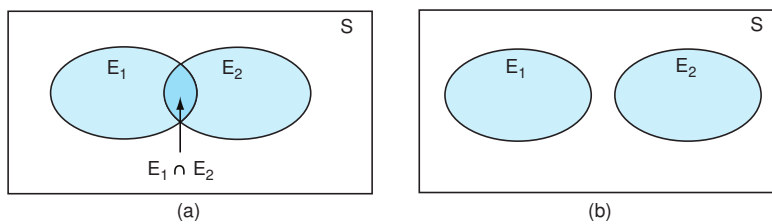
Lag A vinner kampen og det er ikke flere enn to mål. □

2.8 Unionen av to hendelser

Unionen av to hendelser, E_1 og E_2 , skrives med symbolet $E_1 \cup E_2$, og består av alle utfall som tilhører E_1 eller E_2 eller begge, se figur 2(a).

2.9 Disjunkte hendelser

To hendelser, E_1 og E_2 , er disjunkte dersom $E_1 \cap E_2 = \emptyset$, som betyr at det ikke er noen utfall av forsøket som gjør at både E_1 og E_2 inntreffer. To disjunkte hendelser er illustrert i figur 2(b). To disjunkte hendelser sies også å være gjensidig utelukkende, som betyr at hvis én hendelse inntreffer, så kan den andre hendelsen ikke inntreffe samtidig.



Figur 2: Venn-diagram av to hendelser, E_1 and E_2 . I (a) er hendelsene overlappende. I (b) er hendelsene disjunkte.

Legg merke til at enkeltutfallene e i utfallsrommet S er definert slik at de er disjunkte. Et bestemt forsøk vil alltid resultere i ett og bare ett enkeltutfall e i S .

3 Sannsynlighet

Vi benytter sannsynlighetsbegrepet når vi skal si noe om framtida og ønsker å beskrive at vi kan ha flere mulige utfall, men når vi ikke veit hvilket av dem som inntreffer. Sannsynligheten til en hendelse E betegnes ved $\Pr(E)$ og er alltid et tall mellom 0 og 1, eller mellom 0% og 100% dersom vi ønsker å uttrykke sannsynligheten som et prosenttall. Dersom $\Pr(E) = 1$, betyr dette at hendelsen A heilt sikkert vil inntreffe, og dersom $\Pr(E) = 0$, betyr dette at hendelsen A umulig kan inntreffe.

En kortfatta diskusjon av sannsynlighetsbegrepet er gitt i kapittel 2 i boka, hvor det er beskrevet tre ulike tilnærminger til sannsynlighetsbegrepet. Disse var:

1. Den klassiske tilnærmingen (som forutsetter at alle utfallene er like rimelige).
2. Frekventisttilnærmingen (der sannsynligheten til en hendelse E oppfattes som den relative frekvensen av E i en lang rekke forsøk som kan (tenkes) gjennomført under de samme betingelsene).
3. Den bayesianske, tilnærmingen (der sannsynligheten til en hendelse E er et mål for vår tru («degree of belief») eller kunnskap knytta til om E vil inntreffe eller ikke). En bayesiansk sannsynlighet kalles også en *subjektiv* sannsynlighet og/eller en *kunnskapsbasert* sannsynlighet.

Vi gjentar ikke denne beskrivelsen her – og du bør derfor vurdere å lese deler av kapittel 2 i boka før du leser resten av dette notatet.

3.1 Definisjon av sannsynlighet

Sjøl om begrepet sannsynlighet kan oppfattes på ulike måter, vil vi kunne gi sannsynlighetsbegrepet en formell definisjon som kan brukes uansett hvordan du oppfatter begrepet.

La S være utfallsrommet til et stokastisk forsøk. For hver hendelse, E , i utfallsrommet kan du anta at du har et tall, $\Pr(E)$, som tilfredsstiller følgende tre krav:

1. $0 \leq \Pr(E) \leq 1$.

2. $\Pr(S) = 1$.
3. For en hvilket som helst sekvens av disjunkte hendelser E_1, E_2, \dots , slik at $E_i \cap E_j = \emptyset$ for alle $i \neq j$, er

$$\Pr\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \Pr(E_i) \quad (1)$$

$\Pr(E)$ kalles *sannsynligheten* til hendelsen E .

Krav 1 sier at sannsynligheten til en hendelse E alltid må være mellom null og én (1). En sannsynlighet lik null betyr at det er umulig at hendelsen vil inntreffe, mens en sannsynlighet lik én betyr at hendelsen heilt sikkert vil inntreffe (krav 2). Sannsynligheten for at en hendelse befinner seg i utfallsrommet er altså lik én. Krav 3 sier at når du har ei rekke hendelser ikke kan inntreffe samtidig, så er sannsynligheten for at minst en av dem skal inntreffe lik summen av sannsynlighetene til hendelsene (se *addisjonsregelen*).

3.2 Hva er egentlig sannsynlighet?

Sjøl om vi i tilfelle som «kron-mynt»-eksempelet kan finne sannsynligheten forholdsvis enkelt ut fra symmetriegenskapene til forsøksobjektet, må du i en risikoanalyse *anslå*⁴ sannsynlighetene. Slike beregninger kan gjøres på to ulike måter – den klassiske eller den bayesianske.

I den klassiske framgangsmåten anslår du sannsynligheter ut fra data som er samla inn i gjentatte forsøk.⁵ Verdien $p = \Pr(E)$ beregnes fra den relative frekvensen til hendelsen E i disse forsøkene. Tanken er at sannsynligheten $\Pr(E)$ er en egenskap til et fysisk system eller en situasjon.

I den bayesianske framgangsmåten er sannsynligheten *subjektiv* og betraktes som et uttrykk for tru eller overbevisning.⁶ Tidligere erfaring med problemet, fysisk kunnskap og data som er samla fra aktuelle forsøk kan brukes til å anslå sannsynligheten, $\Pr(E)$. Av den grunn brukes også begrepet kunnskapsbasert sannsynlighet. En fordel med den bayesianske metoden er at den kan brukes i situasjoner der stokastiske forsøk ikke er aktuelle. Den subjektive tolkningen betyr også at du kan snakke om sannsynligheter for påstander som ikke har mening i den klassiske tilnærmingen. Ved en bayesiansk tolkning har det for eksempel mening å spørre om sannsynligheten for at Gud eksisterer.⁷

3.3 Uniforme sannsynlighetsmodeller

Den enkleste sannsynlighetsmodellen vi har, er den uniforme sannsynlighetsmodellen som brukes når alle utfallene i utfallsrommet har lik sjanse for å inntreffe.

⁴Engelsk: Estimate.

⁵Noen ganger kan du finne anslåtte sannsynligheter i databaser eller databøker (f.eks. OREDA).

⁶Engelsk: Degree of belief.

⁷Ved å søke på nettet finner du mange artikler og også bøker som diskuterer hvordan du ved en bayesiansk framgangsmåte kan «finne» sannsynligheten for at Gud eksisterer.



Thomas Bayes (1719–1761) studerte teologi og logikk ved universitetet i Edinburgh. Han arbeida som prest, men var spesielt interessert i logikk og grunnlaget for sannsynlighetsregning. Navnet hans er knytta til Bayes regel og til tolkninga hans av sannsynlighetsbegrepet – der sannsynlighet blir oppfatta som et mål for vår tru eller overbevisning i motsetning til den klassiske fortolkningen, der sannsynligheten blir oppfatta som en egenskap ved objektet som studeres. Begrepet «bayesiansk» ble første gang brukt om den subjektive tilnærmingen omtrent i 1950.

La S være et utfallsrom med n mulige utfall. Vi har en uniform sannsynlighetsmodell over S når

$$\Pr(e) = \frac{1}{n} \quad \text{for alle } e \in S \quad (2)$$

La E være en hendelse i S med n_E mulige utfall. Når du har en uniform sannsynlighetsmodell, er sannsynligheten for hendelsen E

$$\Pr(E) = \frac{\text{Antall utfall som gir } E}{\text{Antall mulige utfall}} = \frac{n_E}{n} \quad (3)$$

Eksempel 3.1

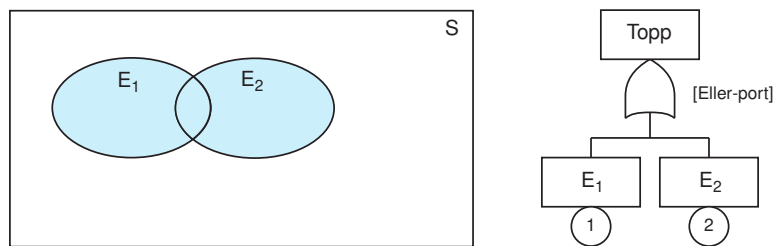
Når du kaster en terning, gjennomfører du et stokastisk forsøk. Utfallsrommet er $S = \{1, 2, 3, 4, 5, 6\}$, og antall mulige utfall er $n = 6$. Dersom terningen er rettferdig, har du en uniform sannsynlighetsmodell og sannsynligheten for å få utfallet $e = 2$ er

$$\Pr(2) = \frac{1}{6}$$

Hendelsen $E = \text{«partall»}$ består av utfallene $\{2, 4, 6\}$. Sannsynligheten til E er derfor

$$\Pr(E) = \frac{n_E}{n} = \frac{3}{6} = 0.50$$

□



Figur 3: Hendelse E_1 eller hendelse E_2 inntreffer.

3.4 Grunnleggende regler i sannsynlighetsregning

3.4.1 Sannsynligheten for komplementærhendelser

La E^* være komplementærhendelsen til hendelsen E . Sannsynligheten E^* er gitt ved

$$\Pr(E^*) = 1 - \Pr(E) \quad (4)$$

Sannsynligheten for at E ikke vil inntreffe er derfor lik én minus sannsynligheten for at hendelsen E vil inntreffe.

3.4.2 Addisjonsregelen

La E_1 og E_2 være to hendelser. Da er

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2) \quad (5)$$

Likninga (5) kalles *addisjonsregelen* i sannsynlighetsregning, og er illustrert i Venn-diagrammet i figur 3. Hvis du legger sammen sannsynlighetene for utfall E_1 og E_2 vil utfallene i snittet $E_1 \cap E_2$ komme med to ganger. Du må derfor trekke fra sannsynligheten for $E_1 \cap E_2$. Addisjonsregelen kan også illustreres ved et feiltre med en eller-port som vist til høyre i figur 3. Topp-hendelsen vil inntreffe når enten inngangshendelsen E_1 eller inngangshendelsen E_2 inntreffer. Dette kan skrives $\text{Topp} = E_1 \cup E_2$.

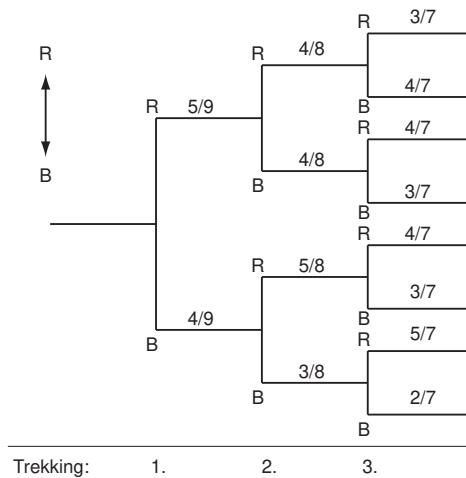
Merk at når E_1 og E_2 er gjensidig utelukkende (dvs. at de ikke kan inntreffe samtidig og er disjunkte), så er

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$$

Dette er illustrert i figur 2(b).

Addisjonsregelen kan enkelt utvides slik at dersom du har en sekvens E_1, E_2, \dots, E_n av hendelser som er gjensidig utelukkende (disjunkte), så er

$$\begin{aligned} \Pr(E_1 \cup E_2 \cup \dots \cup E_n) &= \Pr(E_1) + \Pr(E_2) + \dots + \Pr(E_n) \\ &= \sum_{i=1}^n \Pr(E_i) \end{aligned} \quad (6)$$



Figur 4: Sannsynlighetstre.

3.5 Betinge sannsynlighet

I tida fra sannsynlighetsmodellen ble laga til forsøket ditt blir utført, kan du få ekstra informasjon som du ønsker å ta hensyn til i dine beregninger. Anta at du ønsker å finne sannsynligheten for hendelsen E_2 når du veit at hendelse E_1 allerede har inntruffet. Når E_1 har inntruffet, ønsker du å finne den relative delen av E_2 som er en delmengde av E_1 , kalt $E_1 \cap E_2$. Den *betinge* sannsynligheten for E_2 gitt E_1 defineres da som

$$\Pr(E_2 | E_1) = \frac{\Pr(E_1 \cap E_2)}{\Pr(E_1)} \quad (7)$$

hvis $\Pr(E_1) > 0$.

Eksempel 3.2

Anta at en boks inneholder 5 røde (R) kort og 4 blå (B) kuler. Du trekker tilfeldig ei kule fra boksen. Sannsynligheten for at du trekker ei rød, eventuelt ei blå, kule blir da:

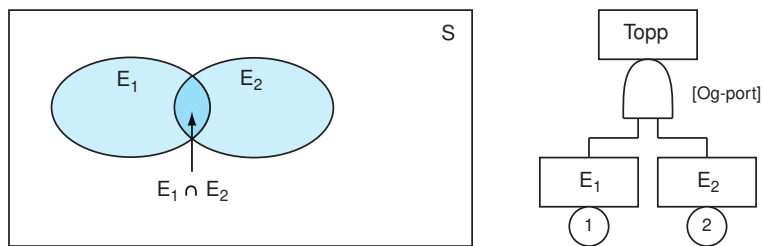
$$\Pr(R) = \frac{5}{9} \quad \Pr(B) = \frac{4}{9}$$

Du trekker så ei ny kule tilfeldig fra boksen. Sannsynligheten for at du nå trekker ei rød (blå) kule vil nå avhenge av hvilken farge det var på kula du trakk i første trekking, og du får

$$\begin{aligned} \Pr(R | R) &= \frac{4}{8} & \Pr(R | B) &= \frac{5}{8} \\ \Pr(B | R) &= \frac{4}{8} & \Pr(B | B) &= \frac{3}{8} \end{aligned}$$

Du trekker så ei tredje kule tilfeldig fra boksen. Sannsynligheten for å trekke ei rød (blå) kule, vil nå avhenge av hvilke farger det var på de kulene du trakk i de to foregående trekkingene. Dette eksemplet kan enkelt illustreres i et *sannsynlighetstre*, som vist i figur 4.

Legg merke til at sannsynlighetstreet tilsvarer et hendelsestre som er omtalt i kapittel 11 i boka. \square



Figur 5: Hendelse E_1 og hendelse E_2 inntreffer.

3.6 Multiplikasjonsregelen

Ved hjelp av betingte sannsynligheter kan du regne ut sannsynligheten for at to eller flere hendelser skal inntreffe samtidig. Fra (7), ser du at

$$\Pr(E_1 \cap E_2) = \Pr(E_2 | E_1) \cdot \Pr(E_1) = \Pr(E_1 | E_2) \cdot \Pr(E_2) \quad (8)$$

Snittet mellom E_1 og E_2 kan illustreres av Venn-diagrammet i figur 5. Feiltreet viser at Topp-hendelsen vil inntreffe dersom inngangshendelsene E_1 og E_2 inntreffer samtidig, slik at $\text{Topp} = E_1 \cap E_2$. Dette kan generaliseres til

$$\begin{aligned} \Pr(E_1 \cap E_2 \cap \dots \cap E_k) &= \\ &= \Pr(E_1) \cdot \Pr(E_2 | E_1) \cdot \Pr(E_3 | E_1 \cap E_2) \dots \Pr(E_k | E_1 \cap \dots \cap E_{k-1}) \end{aligned}$$

3.7 Uavhengige hendelser

Hendelsene E_1 og E_2 sies å være *uavhengige* når de to hendelsene ikke påvirker hverandre, det vil si at informasjon om en hendelse ikke påvirker sannsynligheten for den andre hendelsen. Det betyr at hendelsene E_1 og E_2 er uavhengige hvis, og bare hvis

$$\Pr(E_2 | E_1) = \Pr(E_2) \quad \text{og} \quad \Pr(E_1 | E_2) = \Pr(E_1)$$

Dette er det samme som å skrive at

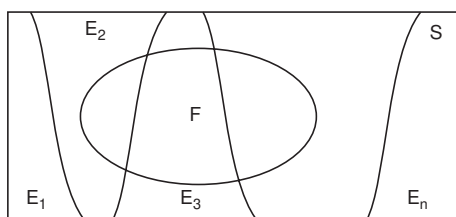
$$\Pr(E_1 \cap E_2) = \Pr(E_1) \cdot \Pr(E_2) \quad (9)$$

Hvis sekvensen E_1, E_2, \dots, E_n av hendelser er *uavhengige*, så er

$$\begin{aligned} \Pr(E_1 \cap E_2 \cap \dots \cap E_n) &= \Pr(E_1) \cdot \Pr(E_2) \dots \Pr(E_n) \\ &= \prod_{i=1}^n \Pr(E_i) \end{aligned} \quad (10)$$

3.8 Oppdeling av utfallsrommet og total sannsynlighet

E_1, E_2, \dots, E_n sies å være en *oppdeling* av utfallsrommet S , dersom disse hendelsene er disjunkte og dersom $S = E_1 \cup E_2 \cup \dots \cup E_n$.



Figur 6: Oppdeling av utfallsrommet.

La F være en hendelse i utfallsrommet S . Sannsynligheten til F kan da skrives som

$$\Pr(F) = \sum_{i=1}^n \Pr(F \cap E_i) = \sum_{i=1}^n \Pr(F | E_i) \cdot \Pr(E_i) \quad (11)$$

Uttrykket (11) kalles regneregelen for *total sannsynlighet* og er illustrert i figur 6.

Eksempel 3.3

Betrakt et prosessystem der tre mulige feilmoder E_1, E_2 og E_3 kan inntreffe. Disse tre feilmodene kan oppfattes som disjunkte og vil omfatte alle mulige feilmoder i denne situasjonen. Når en feilmode inntreffer, vil dette kunne gi en effekt F . Vi antar følgende sannsynligheter:

$$\begin{aligned} \Pr(E_1) &= 0.45 & \Pr(F | E_1) &= 0.35 \\ \Pr(E_2) &= 0.25 & \Pr(F | E_2) &= 0.10 \\ \Pr(E_3) &= 0.30 & \Pr(F | E_3) &= 0.05 \end{aligned}$$

Sannsynligheten for å få effekten F kan da finnes ved å bruke regneregelen for total sannsynlighet

$$\begin{aligned} \Pr(F) &= \Pr(F | E_1) \cdot \Pr(E_1) + \Pr(F | E_2) \cdot \Pr(E_2) + \Pr(F | E_3) \cdot \Pr(E_3) \\ &\approx 0.198 = 19.8\% \end{aligned}$$

□

3.9 Bayes regel

Til nå har vi vist hvordan vi kan beregne sannsynlighetene for at en hendelse skal inntreffe. *Bayes regel* hjelper oss med å regne «motsatt», det vil si at gitt at en hendelse har inntrefft, så kan vi beregne sannsynligheten for at det skyldtes én bestemt årsak.

Hvis E_1, E_2, \dots, E_n er en oppdeling av utfallsrommet S , der $\Pr(E_i) > 0$ for alle $i = 1, 2, \dots, n$, så kan sannsynligheten for hendelsen F i S regnes ut ved hjelp av regelen om *total sannsynlighet*. Sannsynligheten for at hendelsen E_j inntreffer, når du veit at hendelsen F har inntrefft, er gitt ved Bayes regel

$$\Pr(E_j | F) = \frac{\Pr(F \cap E_j)}{\sum_{i=1}^n \Pr(F \cap E_i)} = \frac{\Pr(F | E_j) \cdot \Pr(E_j)}{\sum_{i=1}^n \Pr(F | E_i) \cdot \Pr(E_i)} \quad (12)$$

for alle $j = 1, 2, \dots, n$.

Eksempel 3.4

Et produkt lages av to ulike maskiner A og B . 70 % av produktene er laga av maskin A , mens 30 % av produktene lages av maskin B . Produktene kan være defekte (F) og sannsynligheten for å få et defekt produkt, avhenger av hvilken maskin produktet kommer fra. Erfaring har vist at

$$\Pr(F | A) = 0.02 \quad \text{og} \quad \Pr(F | B) = 0.07$$

Vi velger et tilfeldig produkt, og finner at produktet er defekt (F). Sannsynligheten for at det defekte produktet er laga av maskin B kan da finnes ved å bruke Bayes regel:

$$\begin{aligned} \Pr(B | F) &= \frac{\Pr(F | B) \cdot \Pr(B)}{\Pr(F | B) \cdot \Pr(B) + \Pr(F | A) \cdot \Pr(A)} \\ &= \frac{0.07 \cdot 0.30}{0.07 \cdot 0.30 + 0.02 \cdot 0.70} = 0.60 \end{aligned}$$

□

4 Stokastiske variabler

En stokastisk variabel er en funksjon som «måler» en egenskap ved de ulike utfallene i utfallsrommet. En stokastisk variabel kalles også en tilfeldig variabel. I det følgende bruker vi X som symbol for en stokastisk variabel. Når vi anvender X på et bestemt utfall e , får vi verdien $X(e) = x$.

4.1 Diskrete stokastiske variabler

En stokastisk variabel X er *diskret* når den bare kan anta et tellbart antall verdier. Dette antallet kan være begrensa eller ubegrensa.

4.1.1 Diskret sannsynlighetsfordeling

La X være en diskret stokastisk variabel som kan ha verdiene i mengden $\{x_1, x_2, \dots\}$, og la $\Pr(X = x_i)$ være sannsynligheten for at variabelen X får verdien x_i . Denne sannsynligheten kalles *punktsannsynligheten* for x_i og kan tolkes som sannsynligheten for å få et utfall som er målt (ved X) til å være x_i . Vi har da:

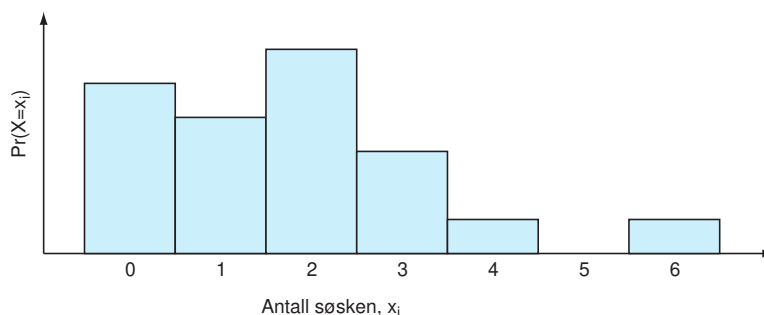
1. $\Pr(X = x_i) \geq 0$ for alle $i = 1, 2, \dots$

2. $\sum_{i=1}^{\infty} \Pr(X = x_i) = 1$

Eksempel 4.1

Anta at en klasse har $n = 20$ studenter og at du ønsker å «måle» hvor mange søsken (inklusive halvsøsken) (X) de ulike studentene har. Etter en spørreunde får du følgende data

Antall søsken:	0	1	2	3	4	5	6
Antall:	5	4	6	3	1	0	1



Figur 7: Histogram.

Tabellen viser at det er 5 studenter som er enebarn, 4 har ett søsken, og så videre. Der-
 som du trekker en student, e , tilfeldig fra denne klassen og «måler» med X hvor mange
 søsken denne studenten har, veit du at verdien av X må være mellom 0 og 6. Punktsann-
 synlighetene for de ulike verdiene er gitt ut fra den uniforme sannsynlighetsmodellen.

Antall søsken (x_i):	0	1	2	3	4	5	6
$\Pr(X = x_i)$:	$\frac{5}{20}$	$\frac{4}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	0	$\frac{1}{20}$

Når du har valgt ut studenten og observert variabelen, veit du verdien x , og du kan
 for eksempel få resultatet $x = 2$. Tilfeldigheten i dette forsøket er knytta til utfallet.
 Vi veit ikke verdien av X fordi vi ikke veit hvilken student som velges ut. Med en
 gang du veit hvem som er valgt ut, kan du i de fleste tilfelle «måle» antall søsken uten
 usikkerhet.

Punktsannsynlighetene til den stokastiske variabelen X illustreres ofte i et *histo-*
gram som vist i figur 7. \square

4.1.2 Fordelingsfunksjon

Fordelingsfunksjonen til en diskret stokastisk variabel X er

$$F(x) = \Pr(X \leq x) = \sum_{x_i \leq x} \Pr(X = x_i) \quad (13)$$

Fordelingsfunksjonen kalles noen ganger den *kumulative* fordelingsfunksjonen.
 Kumulativ betyr å summere.

Sannsynligheten for at den stokastiske variabelen X har en verdi i intervallet $(x_i, x_j]$
 er

$$\begin{aligned} \Pr(x_i < X \leq x_j) &= \Pr(X \leq x_j) - \Pr(X \leq x_i) \\ &= F(x_j) - F(x_i) \end{aligned} \quad (14)$$

for $x_i < x_j$.

4.1.3 Forventningsverdi

Forventningsverdien⁸ til en diskret stokastisk variabel X med verdimengde $\{x_1, x_2, \dots\}$ er

$$E(X) = \mu = \sum_{i=1}^{\infty} x_i \cdot \Pr(X = x_i) \quad (15)$$

Dette betyr at forventningsverdien er summen av alle verdiene variabelen kan ha, multiplisert med de tilhørende sannsynlighetene.

Eksempel 4.2

Betrakt igjen eksempel 4.1. Forventningsverdien til X er

$$\begin{aligned} E(X) &= 0 \cdot \Pr(X = 0) + 1 \cdot \Pr(X = 1) + 2 \cdot \Pr(X = 2) + 3 \cdot \Pr(X = 3) \\ &\quad + 4 \cdot \Pr(X = 4) + 5 \cdot \Pr(X = 5) + 6 \cdot \Pr(X = 6) = 1.75 \end{aligned}$$

Hvis du gjentar forsøket med å velge ut en student tilfeldig mange ganger, vil du i *gjennomsnitt* få 1.75 søsken. Det er opplagt at du ikke kan få forventningsverdien 1.75 i noe enkelt forsøk. \square

La $g(X)$ være en funksjon av X . Forventningsverdien til den stokastiske variabelen $g(X)$ er

$$E[g(X)] = \sum_{i=1}^{\infty} g(x_i) \cdot \Pr(X = x_i) \quad (16)$$

4.1.4 Varians og standardavvik

Variansen til en stokastisk variabel er definert som:

$$\text{Var}(X) = E[(X - \mu)^2] = \sum_{i=1}^{\infty} (x_i - \mu)^2 \Pr(X = x_i) \quad (17)$$

Standardavviket⁹ til X er

$$\text{SD}(X) = \sqrt{\text{Var}(X)} \quad (18)$$

4.2 Kontinuerlige stokastiske variabler

La X være en stokastisk variabel med en verdimengde som ikke er tellbar. Dette er for eksempel tilfellet når vi måler tid eller vekt. Sannsynlighetsfordelingen til X kan beskrives på flere ulike måter. På samme måte som for en diskret variabel, er fordelingsfunksjonen til X gitt ved

$$F(x) = \Pr(X \leq x) \quad \text{for } -\infty < x < \infty$$

Her omfatter verdimengden til X heile tall-linja. I mange tilfelle vil verdimengden være avgrensa.

⁸Engelsk: Expected value eller Mean value.

⁹Engelsk: Standard deviation (SD).

I risikoanalyser kan den stokastiske variabelen representere mange ulike størrelser, som eksplosjonstrykk, spredning av gass, tid til vi får stoppet en gasslekkasje, levetid for tekniske enheter, og så videre. I de neste avsnittene vil vi se spesielt på levetid, eller tid til svikt, for en teknisk enhet. I stedet for X vil vi bruke bokstaven T som betegnelse for den stokastiske variabelen tid til svikt.

5 Pålitelighetsanalyse

I en risikoanalyse vil påliteligheten til utstyr og barrierer ha betydning for resultatet. Det finnes ulike måter å beregne pålitelighet på, og i dette avsnittet beskriver vi noen viktige metoder.¹⁰

Begrepet pålitelighet¹¹ defineres slik:

Pålitelighet: Evnen en teknisk enhet har til å utføre en tiltenkt funksjon under gitte miljø- og driftsforhold over en gitt tidsperiode.

I noen tilfelle defineres pålitelighet som *sannsynligheten* for at en teknisk enhet kan utføre en tiltenkt funksjon under gitte miljø- og driftsforhold over en gitt tidsperiode.

5.1 Svikt og feil

Når en teknisk enhet ikke lenger er i stand til å utføre en tiltenkt funksjon, sier vi at den har svikta. *Svikt* er en hendelse som inntreffer ved et bestemt tidspunkt t . Like før tidspunktet t funksjonerte enheten, mens den like etter t ikke lenger er i stand til å utføre den tiltenkte funksjonen. I noen tilfelle er det lett å observere svikttidspunktet t . I andre tilfelle vil det være en gradvis degradering i enhetens ytelse, slik at det er vanskelig å si akkurat når svikten inntreffer.

Etter at enheten har svikta, sier vi at enheten er i *feiltilstand*, eller at den feiler. Feil er med andre ord en tilstand, mens svikt er en hendelse. På engelsk skiller det klart mellom disse begrepene, der «failure» betegner svikt og «fault» betegner feil. På norsk brukes en mer uryddig terminologi og en skiller ofte ikke mellom svikt og feil.

5.2 Enkle system

Betrakt et system med n komponenter. Hver komponent har to mulige tilstander som beskrives med *tilstandsvariabelen* X_i , for $i = 1, 2, \dots, n$.

$$X_i = \begin{cases} 1 & \text{om komponent } i \text{ er funksjonsdyktig} \\ 0 & \text{om komponent } i \text{ er i feiltilstand} \end{cases} \quad (19)$$

$\mathbf{X} = (X_1, X_2, \dots, X_n)$ kalles *tilstandsvektoren* til systemet.

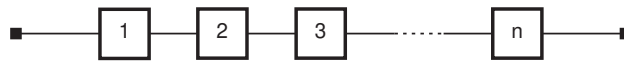
Tilstanden til systemet kan beskrives ved hjelp av den binære funksjonen¹²

$$\phi(\mathbf{X}) = \phi(X_1, X_2, \dots, X_n) \quad (20)$$

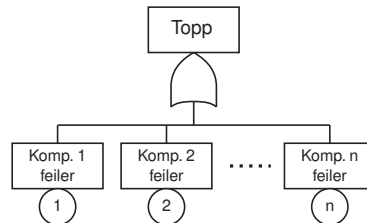
¹⁰En mye grundigere framstilling finner du ved å lese boka: Rausand, M., A. Barror, og A. Høyland (2021): *System Reliability Theory - Models, Statistical Methods, and Applications*. (3.ed.) Wiley

¹¹Engelsk: Reliability.

¹²En binær funksjon/variabel har bare to verdier, her 0 og 1.



Figur 8: Pålitelighetsnettverk for et seriesystem.



Figur 9: Feiltre for et seriesystem.

der

$$\phi(\mathbf{X}) = \begin{cases} 1 & \text{om systemet er funksjonsdyktig} \\ 0 & \text{om systemet er i feiltilstand} \end{cases} \quad (21)$$

Funksjonen $\phi(\mathbf{X})$ kalles *strukturfunksjonen* til systemet.

5.2.1 Seriesystem

Et system, for eksempel et gassdeteksjonssystem, som bare funksjonerer når *alle* n komponentene eller detektorene funksjonerer, kalles et seriesystem. Strukturfunksjonen er da

$$\phi(\mathbf{X}) = X_1 \cdot X_2 \cdots X_n = \prod_{i=1}^n X_i \quad (22)$$

og vi ser at $\phi(\mathbf{X}) = 1$ hvis, og bare hvis $X_i = 1$ for alle $i = 1, 2, \dots, n$.

Seriesystemet kan illustreres ved *pålitelighetsnettverket*¹³ i figur 8.

La hendelsen E_i være at komponent i er i feiltilstand, for $i = 1, 2, \dots, n$, og la E_S betegne at systemet er i feiltilstand. Siden systemet er i feiltilstand når minst én komponent er i feiltilstand, har vi at

$$E_S = E_1 \cup E_2 \cup \cdots \cup E_n$$

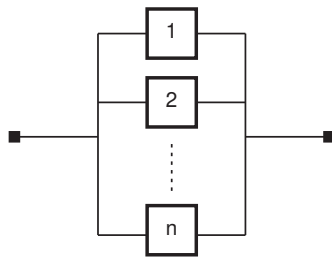
Systemsvikt kan også illustreres ved hjelp av et feiltre, som vist i figur 9, der Topp-hendelsen vil inntreffe dersom inngangshendelse E_1 eller inngangshendelse E_2 eller ... inntreffer. Feiltrær er nærmere omtalt i kapittel 10.

5.2.2 Parallellsystem

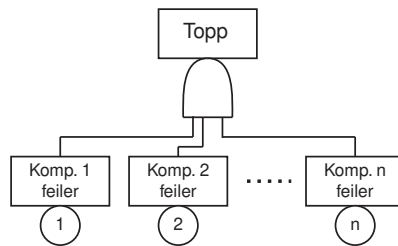
Et system, for eksempel et gassdeteksjonssystem, som funksjonerer når minst en av n komponentene eller detektorene funksjonerer, kalles et parallellsystem. Strukturfunksjonen er da

$$\phi(\mathbf{X}) = 1 - (1 - X_1)(1 - X_2) \cdots (1 - X_n) = 1 - \prod_{i=1}^n (1 - X_i) \quad (23)$$

¹³Engelsk: Reliability block diagram.



Figur 10: Pålitelighetsnettverk for et parallellsystem.



Figur 11: Feiltre for et parallellsystem.

og vi ser at $\phi(\mathbf{X}) = 0$ hvis, og bare hvis, $X_i = 0$ for alle $i = 1, 2, \dots, n$.

Parallellsystemet kan illustreres ved *pålitelighetsnettverket* i figur 10.

Siden parallellsystemet er i feiltilstand når alle komponentene samtidig er i feiltilstand, har vi at

$$E_S = E_1 \cap E_2 \cap \dots \cap E_n$$

der E_i , som ovenfor, betyr at komponent i er i feiltilstand, for $i = 1, 2, \dots, n$.

Systemsvikt kan også illustreres ved hjelp av et feiltre, som vist i figur 11, der Topp-hendelsen inntreffer dersom inngangshendelsene E_1 og hendelse E_2 og ... inntreffer.

5.3 Tid til svikt

*Tid til svikt*¹⁴ for en enhet er tida som går fra enheten settes i drift, til den svikter første gang. Vi bruker $t = 0$ som starttidspunkt. Vi bruker betegnelsen enhet som et samlebegrep for komponent og system. Fordi vi ikke kan forutsi hvor lang tid som går før enheten svikter, kan vi oppfatte tida til svikt, T , som en stokastisk variabel.

Tilstanden til enheten ved tidspunktet t kan beskrives med tilstandsvariabelen $X(t)$, som også er en stokastisk variabel.

$$X(t) = \begin{cases} 1 & \text{om enheten er funksjonsdyktig ved tidspunktet } t \\ 0 & \text{om enheten er i feiltilstand ved tidspunktet } t \end{cases}$$

Sammenhengen mellom tilstandsvariabelen $X(t)$ og tida til svikt T er vist i figur 12.

Legg merke til at tid til svikt T ikke alltid måles i kalendertid. Den kan også måles med andre tidsbegrep, for eksempel som:

¹⁴Engelsk: Time to failure.



Figur 12: Tilstandsvariabelen og tid til svikt for en enhet.

- Antall ganger en bryter koples inn/ut
- Antall km en bil kjører
- Antall omdreininger for et rullelager

I disse eksemplene vil tid til svikt ofte være en diskret variabel. En diskret variabel kan imidlertid tilnærmes ved en kontinuerlig variabel. I det følgende vil vi anta at tid til svikt T er en kontinuerlig variabel. Vi antar at enheten settes i drift ved tidspunkt $t = 0$ og at verdimengden til T er den positive akse, det vil si at vi alltid har $T \geq 0$.

5.3.1 Fordelingsfunksjonen

Fordelingsfunksjonen F er definert som

$$F(t) = \Pr(T \leq t) \quad \text{for } t > 0 \quad (24)$$

Vi ser at $F(t)$ er sannsynligheten for at enheten vil svikte i tidsintervallet $(0, t]$.

5.3.2 Sannsynlighetstettheten

Sannsynlighetstettheten¹⁵ $f(t)$ er

$$\begin{aligned} f(t) &= \frac{d}{dt} F(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < T \leq t + \Delta t)}{\Delta t} \end{aligned} \quad (25)$$

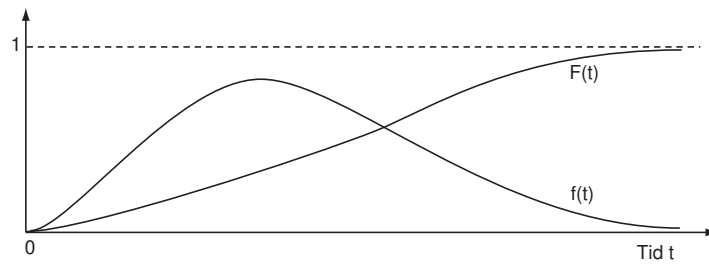
Dette betyr at når Δt er liten, så er

$$\Pr(t < T \leq t + \Delta t) \approx f(t) \cdot \Delta t \quad (26)$$

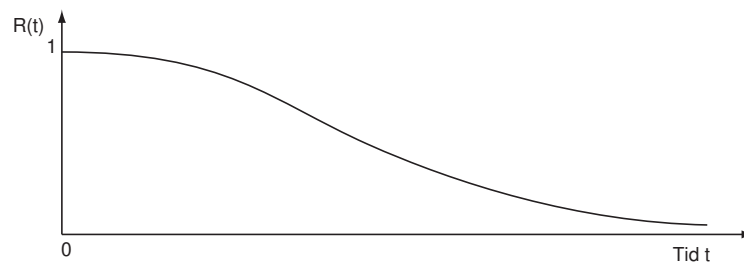
Fordelingsfunksjonen $F(t)$ og sannsynlighetstettheten $f(t)$ er illustrert i figur 13.

Fordelingsfunksjonen begynner med verdien 0 for den minste verdien og slutter med 1 for den største verdien. Sannsynlighetstettheten viser altså sannsynligheten for å få en verdi i et bestemt intervall.

¹⁵Engelsk: Probability density function.



Figur 13: Fordelingsfunksjonen $F(t)$ og sannsynlighetstettheten $f(t)$.



Figur 14: Overlevelsesfunksjonen $R(t)$.

5.3.3 Overlevelsesfunksjonen

Overlevelsesfunksjonen¹⁶ til en enhet er gitt ved

$$R(t) = 1 - F(t) = \Pr(T > t) \quad \text{for } t > 0 \quad (27)$$

$R(t)$ uttrykker sannsynligheten for at en enhet ikke vil svikte i tidsintervallet $(0, t]$, eller sannsynligheten for at enheten overlever tidsintervallet $(0, t]$ og fortsatt er funksjonsdyktig ved tidspunktet t .

5.3.4 Sviktintensitet/feilrate

Sannsynligheten for at en enhet vil svikte i tidsintervallet $(t, t + \Delta t]$ når vi veit at enheten funksjonerte ved tidspunktet t , er gitt ved den betingte sannsynligheten

$$\Pr(t < T \leq t + \Delta t \mid T > t) = \frac{\Pr(t < T \leq t + \Delta t)}{\Pr(T > t)} = \frac{F(t + \Delta t) - F(t)}{R(t)}$$

Ved å dividere denne sannsynligheten på lengden av tidsintervallet, Δt , og la $\Delta t \rightarrow 0$, får vi sviktintensiteten $z(t)$ til enheten

$$\begin{aligned} z(t) &= \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < T \leq t + \Delta t \mid T > t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \frac{1}{R(t)} = \frac{f(t)}{R(t)} \end{aligned} \quad (28)$$

¹⁶Engelsk: Survivor function.

Sviktintensiteten¹⁷ $z(t)$ blir også kalt *feilraten* til enheten. I det følgende vil vi bruke begrepet feilrate sjøl om dette begrepet er litt inkonsekvent.

Når Δt er liten, ser vi fra (28) at

$$\Pr(t < T \leq t + \Delta t \mid T > t) \approx z(t) \cdot \Delta t \quad (29)$$

Legg merke til forskjellen mellom sannsynlighetstettheten $f(t)$ og feilraten $z(t)$. Anta at vi starter med en ny enhet ved tidspunktet $t = 0$ og at vi ved tidspunktet $t = 0$ spør «hva er sannsynligheten for at denne enheten vil svikte i intervallet $(t, t + \Delta t]$?» Ifølge (26) er denne sannsynligheten tilnærmet lik sannsynlighetstettheten $f(t)$ ved tidspunktet t multiplisert med lengden av intervallet Δt . Betrakt deretter en enhet som har overlevd heilt til tidspunktet t , og at vi ved tidspunktet t spør «hva er sannsynligheten for at denne enheten vil svikte i det neste intervallet $(t, t + \Delta t]$?» Denne (betinga) sannsynligheten er ifølge (29) tilnærmet lik feilraten $z(t)$ ved tidspunktet t multiplisert med intervallengden Δt .

5.3.5 Forventa tid til svikt

La T være tida til svikt for en enhet. *Forventa tid til svikt*, MTTF ¹⁸ er da

$$\text{MTTF} = E(T) = \int_0^{\infty} t f(t) dt \quad (30)$$

Fordi $f(t) = -R'(t)$, har vi at

$$\text{MTTF} = - \int_0^{\infty} t R'(t) dt$$

Ved hjelp av delvis integrasjon, får vi

$$\text{MTTF} = - [tR(t)]_0^{\infty} + \int_0^{\infty} R(t) dt$$

Hvis $\text{MTTF} < \infty$, kan det vises at $[tR(t)]_0^{\infty} = 0$. I så fall er

$$\text{MTTF} = \int_0^{\infty} R(t) dt \quad (31)$$

Det er ofte enklere å beregne MTTF ut fra (31) enn ved å bruke (30).

5.3.6 Median levetid

Median levetid t_m defineres ved

$$R(t_m) = 0.50 \quad (32)$$

Medianen deler fordelingen i to halvdel. Enheten vil svikte før tidspunktet t_m med 50 % sannsynlighet, og etter t_m med 50 % sannsynlighet.

¹⁷Engelsk: Failure rate function. Betegnelsen Hazard rate function brukes også.

¹⁸Engelsk: Mean Time To Failure (MTTF).

5.3.7 Varians

Variansen til T er

$$\text{Var}(T) = \sigma^2 = E[(T - \mu)^2] = \int_0^{\infty} (t - \mu)^2 f(t) dt \quad (33)$$

5.3.8 Uavhengige variabler

De to kontinuerlige variablene T_1 og T_2 er uavhengige hvis $f(t_2 | t_1) = f_2(t_2)$ og $f(t_1 | t_2) = f_1(t_1)$, som er det samme som å si at T_1 og T_2 er uavhengige dersom

$$f(t_1, t_2) = f_1(t_1) \cdot f_2(t_2) \quad \text{for alle } t_1 \text{ og } t_2$$

Denne definisjonen kan lett utvides til å gjelde for mer enn to variabler.

6 Noen vanlig brukte fordelinger

6.1 Diskrete fordelinger

6.1.1 Binomisk fordeling

Den binomiske fordelingen brukes når

1. Vi har n uavhengige forsøk
2. Hvert forsøk har to mulige utfall E og E^*
3. Sannsynligheten $\Pr(E) = p$ er den samme i alle forsøk

Denne situasjonen kalles en *binomisk situasjon*, og forsøkene kalles av og til *Bernoulli-forsøk*. La X være antallet av de n forsøkene som gir utfallet E . Da er X en diskret stokastisk variabel med punktsannsynlighet

$$\Pr(X = x) = \binom{n}{x} p^x (1 - x)^{n-x} \quad \text{for } x = 0, 1, \dots, n \quad (34)$$

der

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \quad (35)$$

er *binomialkoeffisienten*.

Fordelingen (34) kalles den *binomiske* fordelingen (n, p) , og vi skriver noen ganger $X \sim \text{bin}(n, p)$. Her er n , antall forsøk, vanligvis en kjent konstant, mens p er en *parameter* i fordelingen. Parameteren p er vanligvis ukjent og ikke direkte observerbar. Det er ikke mulig å «måle» en parameter. En parameter kan bare anslås basert på observerte data. Dersom vi bruker den bayesianske framgangsmåten, kan parameteren også anslås basert på tru og erfaring.

Forventningsverdien og variansen til X er

$$E(X) = np \quad (36)$$

$$\text{Var}(X) = np(1 - p) \quad (37)$$

Eksempel 6.1

Anta at du har ansvaret for å teste ei brannpumpe med jamne mellomrom. I disse testene forsøker du å starte pumpa og lar den kjøre ei kort stund. Du observerer «svikt i oppstart» (hendelse E) dersom brannpumpa ikke kan startes i løpet av et bestemt tidsintervall. Anta at du har gjennomført $n = 150$ tester, og at du betrakter disse testene som uavhengige. I $X = 2$ tester har hendelsen E blitt registrert. Fra (36) veit du at $p = E(X)/n$ og at det derfor er naturlig å anslå sannsynligheten for svikt i oppstart ved

$$\hat{p} = \frac{x}{n} = \frac{2}{150} \approx 0.0133 = 1.33\%$$

□

Eksempel 6.2

Et 2-av-3-system er et system som funksjonerer når minst 2 av 3 komponenter funksjonerer. Anta at komponentene er uavhengige, og la X være antall komponenter som funksjonerer. La p være sannsynligheten for at en bestemt komponent funksjonerer. Denne situasjonen kan betraktes som en binomisk situasjon med $n = 3$ forsøk, og X har derfor ei binomisk fordeling. Sannsynligheten p_S for at 2-av-3-systemet funksjonerer er

$$\begin{aligned} p_S &= \Pr(X \geq 2) = \Pr(X = 2) + \Pr(X = 3) \\ &= \binom{3}{2} p^2 (1-p)^{3-2} + \binom{3}{3} p^3 (1-p)^{3-3} \\ &= 3p^2(1-p) + p^3 = 3p^2 - 2p^3 \end{aligned}$$

□

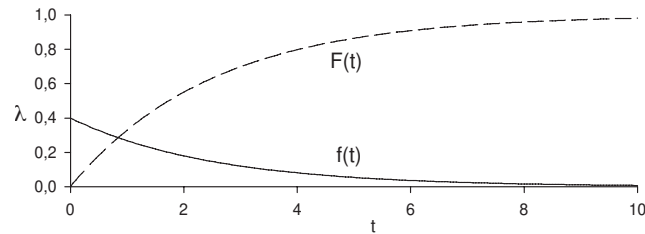
6.1.2 Poisson-fordelingen og Poisson-prosessen

En Poisson-prosess brukes til å modellere antall hendelser E som inntreffer i løpet av et gitt tidsintervall. Hendelsen E kan, for eksempel, være en svikt eller en uønska hendelse. Følgende forutsetninger stilles:

1. Antall hendelser E i et tidsintervall er uavhengig av antallet som inntreffer i et hvilket som helst anna ikke-overlappende tidsintervall. Dette betyr at Poisson-prosessen ikke har noen «hukommelse».
2. Sannsynligheten for at hendelsen E vil inntreffe i løpet av et svært kort tidsintervall er proporsjonal med lengden av tidsintervallet, og avhenger ikke av antallet hendelser som inntreffer utafør dette tidsintervallet.
3. Sannsynligheten for at mer enn én hendelse vil inntreffe i løpet av et svært kort tidsintervall er neglisjerbar.

Vi antar at prosessen starter ved tidspunktet $t = 0$. La $N(t)$ være antall hendelser E som inntreffer i tidsintervallet $(0, t]$. Sannsynlighetsfordelingen til $N(t)$ kalles *Poisson-fordelingen* og er gitt ved

$$\Pr(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad \text{for } n = 0, 1, \dots \quad (38)$$



Figur 15: Eksponensialfordelingen ($\lambda = 0.4$).

der $\lambda > 0$ er en parameter og $e = 2.71828 \dots$

Forventningsverdien av $N(t)$ er

$$E(N(t)) = \lambda t \quad (39)$$

og

$$\lambda = \frac{E(N(t))}{t}$$

Parameteren λ er derfor forventet antall hendelser pr. tidsenhet, og kalles *frekvensen* til Poisson-prosessen.

Variansen til $N(t)$ er

$$\text{Var}(N(t)) = \lambda t \quad (40)$$

6.2 Kontinuerlige fordelinger

6.2.1 Eksponensialfordelingen

La oss anta at tida til svikt T for en enhet er eksponensialfordelt med parameter λ . Sannsynlighetstettheten til T er da gitt ved

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{for } t > 0, \lambda > 0 \\ 0 & \text{ellers} \end{cases} \quad (41)$$

Fordelingsfunksjonen er

$$F(t) = \Pr(T \leq t) = 1 - e^{-\lambda t} \quad \text{for } t > 0 \quad (42)$$

Sannsynlighetstettheten og fordelingsfunksjonen til eksponensialfordelingen er illustrert i figur 15.

Påliteligheten (overlevelsesfunksjonen) blir

$$R(t) = \Pr(T > t) = \int_t^{\infty} f(u) du = e^{-\lambda t} \quad \text{for } t > 0 \quad (43)$$

Forventa tid til svikt (MTTF)

$$\text{MTTF} = \int_0^{\infty} R(t) dt = \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda} \quad (44)$$

Variansen til T er

$$\text{Var}(T) = \frac{1}{\lambda^2} \quad (45)$$

og feilraten er

$$z(t) = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \quad (46)$$

Vi ser at feilraten til en enhet med eksponensialfordelt levetid er konstant, det vil si uavhengig av tida. En brukt enhet har derfor samme sviktsannsynlighet som en ny enhet.

Resultatene i (44) og (46) er i god overensstemmelse med bruken av begrepene i dagligtalen. Hvis en enhet i gjennomsnitt har $\lambda = 4$ svikt/år, så er enhetens MTTF 1/4 år.

Anta at en enhet har eksponensiell tid til svikt T . For en slik enhet

$$\Pr(T > t + x \mid T > t) = \frac{\Pr(T > t + x)}{\Pr(T > t)} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \Pr(T > x)$$

Dette betyr at sannsynligheten for at en enhet vil funksjonere ved tidspunktet $t + x$, gitt at den funksjonerer ved tidspunktet t , er lik sannsynligheten for at en ny enhet har tid til svikt som er lengre enn x . Derfor er den gjenværende levetida til en enhet som funksjonerer ved tidspunkt t , uavhengig av t , og vi sier at eksponensialfordelingen mangler hukommelse.

Oppsummert betyr en antakelse om eksponensialfordelt levetid at

- En brukt enhet er stokastisk sett *så god som ny*. Derfor er det ingen grunn til å bytte ut en enhet som funksjonerer.
- For beregningen av overlevelsesfunksjonen, MTTF, og så videre, er det tilstrekkelig å samle inn data om antall timer i drift og antall svikt. Alderen til disse enhetene er ikke av interesse i denne sammenhengen.

Eksponensialfordelingen er den mest brukte levetidsfordelingen i anvendte risiko- og pålitelighetsanalyser. Grunnen er dens matematiske enkelthet og at den fører til realistiske levetidsmodeller for visse typer enheter.

6.2.2 Normal-fordelingen

Den mest brukte fordelingen i statistikkfaget er normalfordelingen. En stokastisk variabel T er normalfordelt med forventning ν og varians τ^2 når sannsynlighetstettheten til T er

$$f(t) = \frac{1}{\sqrt{2\pi} \cdot \tau} e^{-(t-\nu)^2/2\tau^2} \quad \text{for } -\infty < t < \infty \quad (47)$$

For enkelhets skyld skriver vi noen ganger $T \sim \mathcal{N}(\nu, \tau^2)$. Sannsynlighetstettheten $f(t)$ er symmetrisk rundt en vertikal akse gjennom $t = \nu$, og har t -aksen som horisontal asymptote. Kurven har sine vendepunkt i $t = \nu \pm \tau$. Arealet som begrenses av kurven og den horisontaleaksen er (som for alle sannsynlighetsfordelinger) lik 1.

I det spesielle tilfellet når $\nu = 0$ og $\tau^2 = 1$ kalles fordelingen *standard normalfordeling*, $\mathcal{N}(0, 1)$.

Hvis $T \sim \mathcal{N}(\nu, \tau^2)$ og $\tau^2 > 0$, så er den stokastiske variabelen

$$U = \frac{T - \nu}{\tau} \sim \mathcal{N}(0, 1)$$

Fordelingsfunksjonen til en stokastisk variabel U med standard normalfordeling betegnes vanligvis $\Phi(u) = \Pr(U \leq u)$. Den tilhørende sannsynlighetstettheten er

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \quad (48)$$

Fordelingsfunksjonen til en generell normalfordelingen $T \sim \mathcal{N}(\nu, \tau^2)$ kan skrives som

$$\begin{aligned} F(t) &= \Pr(T \leq t) = \Pr\left(\frac{T - \nu}{\tau} \leq \frac{t - \nu}{\tau}\right) \\ &= \Pr\left(U \leq \frac{t - \nu}{\tau}\right) = \Phi\left(\frac{t - \nu}{\tau}\right) \end{aligned} \quad (49)$$

Eksempel 6.3

La T være tida til svikt for en teknisk enhet og anta at T er normalfordelt med forventning $\nu = 20\,000$ timer og varians τ^2 der $\tau = 5000$ timer. Sannsynligheten for at enheten vil svikte i tidsintervallet fra $t_1 = 17\,000$ timer til $t_2 = 21\,000$ timer er:

$$\begin{aligned} \Pr(t_1 < T \leq t_2) &= \Pr\left(\frac{t_1 - \nu}{\tau} < \frac{T - \nu}{\tau} \leq \frac{t_2 - \nu}{\tau}\right) \\ &= \Phi\left(\frac{t_2 - \nu}{\tau}\right) - \Phi\left(\frac{t_1 - \nu}{\tau}\right) \\ &= \Phi(0.200) - \Phi(-0.600) \approx 0.305 \end{aligned}$$

□

La T_1, T_2, \dots, T_n være uavhengig og identisk fordelt $\mathcal{N}(\nu, \tau^2)$. Det kan vises at

$$\sum_{i=1}^n T_i \sim \mathcal{N}(n\nu, n\tau^2)$$

Summen av uavhengige normalfordelte stokastiske variabler er derfor også normalfordelt og

$$\frac{\sum_{i=1}^n T_i - n\nu}{\sqrt{n}\tau} = \frac{\frac{1}{n} \sum_{i=1}^n T_i - \nu}{\tau} \sqrt{n} \sim \mathcal{N}(0, 1)$$

Ved å bruke betegnelsen $\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i$ kan det siste uttrykket skrives som

$$\frac{\bar{T} - E(\bar{T})}{\sqrt{\text{Var}(\bar{T})}} \sim \mathcal{N}(0, 1) \quad (50)$$

6.2.3 Sentralgrensesetningen

La X_1, X_2, \dots være ei rekke av uavhengige, identisk fordelte stokastiske variabler, hver med forventningsverdi ν og varians τ^2 . La \bar{X} være det empiriske gjennomsnittet av de n første variablene i denne sekvensen, slik at

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Fordelingen er gitt ved

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{\text{Var}(\bar{X})}} = \frac{\bar{X} - \nu}{\tau} \sqrt{n}$$

som nærmer seg en standard normalfordeling når $n \rightarrow \infty$. Det betyr at

$$\Pr \left(\frac{\bar{X} - \nu}{\tau} \sqrt{n} \leq y \right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-t^2/2} dt \quad (51)$$

Dette resultatet kalles *sentralgrensesetningen* og (51) er gyldig for *alle* fordelinger for X_i .

Eksempel 6.4

La X ha ei binomisk fordeling med parametrene n og p . X kan betraktes som en sum av n uavhengige stokastiske variabler som er binomisk fordelte $(1, p)$. Vi kan derfor bruke sentralgrensesetningen til å konkludere med at fordelingen til

$$\frac{X - E(X)}{\sqrt{\text{Var}(X)}} = \frac{X - np}{\sqrt{np(1-p)}}$$

nærmer seg ei standard normalfordeling når $n \rightarrow \infty$. Normaltilnærmingen betraktes som god når n og p er slik at $np(1-p) \geq 10$. \square