STOKASTIK
Sannolikhetsteori och statistikteori med tillämpningar

Sven Erick Alm och Tom Britton

Typsatt med liber1ab

2007-10-08

Innehåll

KAPITEL 2

Sannolikhetsteorins grunder

I detta kapitel kommer vi att gå igenom grunderna för sannolikhetsteorin. Inledningsvis kommer merparten exempel vi använder vara enkla, klassiska slumpexperiment så som slantsingling, tärningskast och kortdragning. Vi använder dessa eftersom de är renodlade och kräver liten eller ingen ytterligare förklaring. I senare avsnitt i boken kommer våra exempel att bli mer varierade och intressanta.

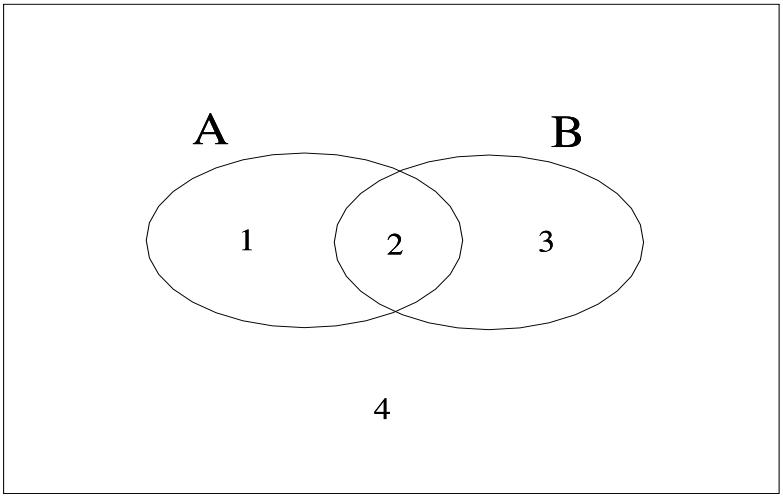
2.1 Utfallsrum och mängdlära

Vi skall nu definiera vad som menas med utfall, händelser och utfallsrum, samt gå igenom några viktiga begrepp från mängdläran. Dessa kunskaper kommer vi att använda oss av i nästa avsnitt då vi presenterar slumpexperiment och beräknar sannolikheter för utfall av slumpexperiment. I grunden ligger hela tiden ett slumpexperiment eller slumpförsök. Med detta menas en situation där något kommer att inträffa, men vi inte med säkerhet i förhand kan säga vad.

DEFINITION 2.1 (UTFALL, HÄNDELSER OCH UTFALLSRUM)

Resultatet av ett slumpförsök kallas ett *utfall*. Mängden av möjliga utfall från ett visst slumpförsök kallas *utfallsrum*. En viss specificerad mängd utfall kallas för en *händelse* – således är enskilda utfall, liksom hela utfallsrummet också händelser. Enskilda utfall betecknas med u_1, $u_2, \ldots$, händelser betecknas med versaler A, B, $\ldots$ och utfallsrummet med Ω. Utfallsrum med ändligt eller uppräkneligt oändligt många utfall kallas *diskreta* utfallsrum medan övriga kallas *kontinuerliga* utfallsrum.

Det är värt att påpeka att utfall och händelser inte är ”tal” utan *element* respektive *mängder* av element. Man kan således inte addera eller subtrahera händelser med varandra men man kan däremot betrakta unioner och snitt av händelser, och dessa unioner och snitt är i sin tur också händelser. Unioner betecknas med $\cup$ och snitt med $\cap$. Händelsen $A \cup B$, som läses ”A union B”, utgörs av alla utfall som ingår i någon av händelserna A eller B, eller bägge (område 1, 2 och 3 i Venndiagrammet i Figur 2.1). Händelsen $A \cap$



Figur 2.1. *Ett Venndiagram med utfallsrum Ω och två händelser A och B. Hur område 1, 2, 3 och 4 (det som ligger utanför ringarna) uttrycks i termer av A och B beskrivs i texten.*

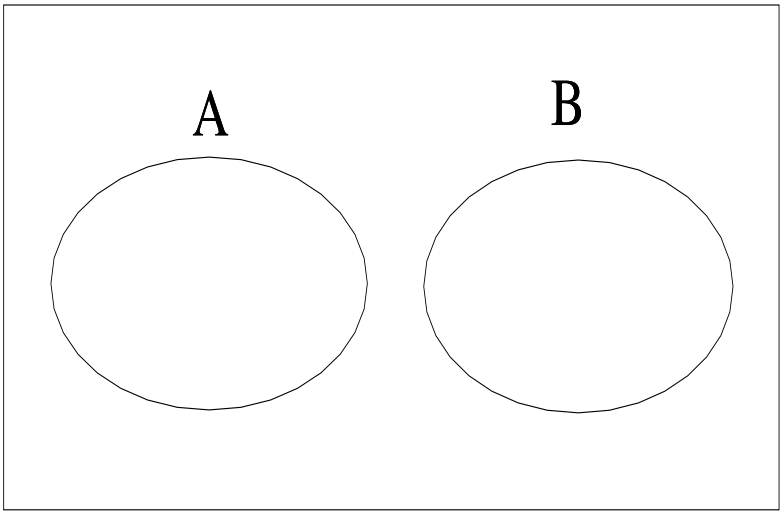
B, snittet av A och B, består däremot bara av utfallen som ingår i bägge händelserna (område 2 i Figur 2.1). För flera händelser $A_1, \ldots, A_n$ består händelserna $\cup_{i=1}^{n} A_i$ och $\cap_{i=1}^{n} A_i$ på motsvarande sätt av de utfall som ingår i någon, respektive alla, A_i-händelserna.

EXEMPEL 2.1 (*Årsdatum*)

Låt $\Omega = \{1/1, 2/1, \ldots, 31/1, 1/2, \ldots, 31/12\}$ vara dagarna på året under ett år utan skottår. Om vi låter $A = \{1/1, \ldots, 30/6\}$ vara det första halvåret medan $B = \{1/6, \ldots, 31/8\}$ vara dagarana i de tre sommarmånaderna blir $A \cup B = \{1/1, \ldots, 31/8\}$, dvs alla dagar i januari till augusti, medan $A \cap B = \{1/6, \ldots, 30/6\}$ endast består av dagarna i juni. Definiera även $A_1 = \{1/1, 1/2, \ldots, 1/12$, dvs den första dagen i varje månaderna och

på motsvarande sätt $A_i = \{i/1, i/12 \ldots, i/12\}$ den i:te dagen i respektive månad, $i = 1, \ldots, 31$. Då blir $\bigcup_{i=1}^{5} A_i = \{1/1, \ldots, 5/1, 1/2, \ldots, 5/12\}$ de första fem dagarna i respektive månad. Mängden $\bigcap_{i=1}^{5} A_i = \emptyset$ (nedan förklaras beteckningen $\emptyset$) innehåller däremot inga utfall eftersom månadernas första dag inte har någon gemensam dag med månadernas andra dag osv.

Ibland vill man betrakta ”komplementet” till en händelse, och med detta menas ”de utfall som inte ingår i händelsen”. Komplementet till händelsen A betecknas A^c, läses som ”A-komplement”, och består således av utfallen som inte finns i A, dvs. $A^c = \{u \in \Omega;\ u \notin A\}$ (område 3 och 4 i Figur 2.1). En annan typ av händelse är ”A men inte B”. Denna händelse har därför fått en egen beteckning, nämligen $A \setminus B$ (område 1 i Figur 2.1). Egentligen är beteckningen överflödig eftersom $A \setminus B = A \cap B^c$, men den är trots detta praktisk att ha till hands. En speciell händelse är ”inget utfall” vilket brukar betecknas med $\emptyset$ och kallas för tomma mängden. Till exempel gäller att $\Omega^c = \emptyset$. Två händelser sägs vara *oförenliga*, eller *disjunkta*, om de inte har några gemensamma utfall (se Figur 2.2). För sådana par av händelser gäller att $A \cap B = \emptyset$. Slutligen definierar vi begreppet delmängd. En händelse A är en



Figur 2.2. *Ett utfallsrum* Ω *med två* oförenliga *(disjunkta) händelser* A *och* B*, dvs.* $A \cap B = \emptyset$.

delmängd av händelsen B om alla utfall i A också ligger i B. Detta skrivs som $A \subset B$. Det gäller för övrigt att $A \subset B$ om och endast om $A \cap B = A$.

Sammanfattningsvis gäller alltså:

- att A inte inträffar skrivs som A^c,
- att minst en av A och B inträffar skrivs $A \cup B$,
- att både A och B inträffar skrivs $A \cap B$,
- att A men inte B inträffar skrivs $A \setminus B$.

EXEMPEL 2.2 (*Kortdragning*)

För slumpexperimentet att dra ett kort ur en kortlek består utfallsrummet av $\Omega = \{S1, \ldots, S13, H1, \ldots, H13, K1, \ldots, K13, R1, \ldots, R13\}$ där utfallet S1, betyder spader ess, utfallet H13 hjärter kung osv. Händelsen A =”klöver” omfattar således utfallen $A = \{K1, \ldots, K13\}$, händelsen B =”femma” definieras av $B = \{S5, H5, K5, R5\}$ och händelsen C =”klädda kort” (dvs. knekt, dam eller kung) omfattar $C = \{S11, S12, S13, H11, \ldots, R13\}$. För dessa händelser gäller $A \cup B = \{K1, \ldots, K13, S5, H5, R5\}$, $A \cap B = \{K5\}$, $A^c = \{S1, \ldots, S13, H1, \ldots, H13, R1, \ldots, R13\}$ samt $A \setminus B = \{K1, \ldots, K4, K6, \ldots K13\}$.
Händelserna B och C är oförenliga, ett kort kan ju inte vara en femma och klätt på samma gång, så $B \cap C = \emptyset$.

EXEMPEL 2.3 (*Temperaturmätning*)

Låt oss studera ”slumpexperimentet” att man mäter temperaturen en viss tid på en viss plats. Här definieras utfallsrummet lämpligen som $\Omega = \mathcal{R}$, dvs. alla reella tal (där talen motsvararar temperaturen angiven i grader Celsius). Det är möjligt att snäva in utfallsrummet, t.ex. är det ju teoretiskt omöjligt att det är kallare än -273.15. (Om det gäller utomhustemperatur i skuggan i Sverige kan man nog snäva in Ω ännu mer: kallare än -100 känns inte aktuellt och varmare än 50 dröjer väl ännu några år innan växthuseffekten ger upphov till – men vi lämnar dessa justeringar därhän.) Om A =”minusgrader” gäller att $A = (-\infty, 0)$, medan händelsen “mellan 10 och 20 grader kallt” blir $B = (-20, -10)$. För A gäller $A^c = [0, +\infty)$, dvs. ”plusgrader” (om man räknar in 0 i ”plus”). Det gäller vidare att $A \cup B = (-\infty, 0)$ och $A \cap B = (-20, -10)$. De två sistnämnda resultaten är en direkt följd av att $B \subset A$ vilket alltid medför att $A \cup B = A$ och $A \cap B = B$.

Utfallsrum som har ändligt många, eller uppräkneligt oändligt många, utfall definierades tidigare som diskreta medan övriga kallas kontinuerliga. Utfallsrummet bestående av resultatet från en kortdragning i Exempel 2.2 är diskret, liksom t.ex. utfallsrummet som består av alla positiva heltal, medan utfallsrummet för temperaturen i Exempel 2.3 är kontinuerligt. Om kontinuerliga variabler endast uppmäts med förpreciserad noggrannhet, vilket nästan alltid är fallet, är emellertid även dylika utfallsrum diskreta. Om t.ex. temperaturen anges med en decimals noggrannhet blir utfallsrum $\Omega = \{\ldots, -0.2, -0.1, 0.0, 0.1, 0.2, \ldots\}$ vilket är ett diskret utfallsrum.

ÖVNING 2.1

Betrakta utfallsrummet Ω bestående ett företags ekonomiska resultat (avrundat och mätt i tusentals kronor). Låt A beteckna händelsen att företaget gör ett positivt resultat. Låt B beteckna händelsen att företaget gör ett bättre resultat än föregående år då man gjorde ett vinstresultat på 1.400 miljoner kronor.

a) Definiera Ω, A och B.
b) Bestäm $A \cup B$ och $A \cap B$.
c) Bestäm A^c och $A \setminus B$.

ÖVNING 2.2

En pilkastningstävling går till så att deltagarna får kasta tills de för första gången träffar ”bulls eye” (den innersta lilla cirkeln på piltavlan). Den vinner som klarar detta på minst antal kast. Bestäm utfallsrummet av möjliga utfall, samt händelserna A att det sker efter högst 10 kast samt B att det sker på ett jämnt antal kast.

a) Definiera Ω, A och B.
b) Bestäm $A \cup B$ och $A \cap B$.
c) Bestäm A^c och $A \setminus B$.

ÖVNING 2.3

Betrakta årets dagar ett år som inte är skottår, t ex 2/9 och 31/7.

a) Definiera utfallsrummet Ω.
b) Bestäm händelsen S bestående av september månads dagar och O bestående av oktober månads dagar.

c) Bestäm V dagarna i stjärntecknet Vågen (Vågen inträffar mellan 24/9 och 23/10).
d) Uttryck följande händelser i termer av S, O och V, samt i termer av de enskilda utfallen: september månads dagar då inte Vågens stjärntecken inträffar, dagarna i oktober då Vågen inträffar, dagarna då det är september eller Vågen inträffar.

2.2 Sannolikheter på utfallsrum

Nu när vi preciserat vad som menas med utfallsrum ska vi definiera slumpförsök på sådana och sannolikheter på utfallsrum. Ett *slumpförsök* på ett utfallsrum består av ett försök som resulterar i ett av utfallen i utfallsrummet, och man kan på förhand inte veta exakt vilket av utfallen som kommer att inträffa. I stället beskrivs slumpförsöket genom att precisera sannolikheten för (alla) händelser i utfallsrummet. Sannolikheten för händelsen A brukar skrivas $P(A)$ inspirerat av engelskans ”probability”. Man ställer dock vissa krav på funktionen $P(\cdot)$ för att den skall få kallas sannolikhetsfunktion. Följande högst rimliga krav på sannolikheter infördes av den ryske matematikern Andrei Kolmogorov (1903-1987):

DEFINITION 2.2 (KOLMOGOROVS AXIOMSYSTEM)

En reell funktion, P, på händelser i utfallsrummet Ω är en *sannolikhetsfunktion* om den uppfyller följande tre villkor (axiom):

1. $0 \leq P(A) \leq 1$ för alla händelser $A \subset \Omega$,
2. $P(\Omega) = 1$,
3. om $A \cap B = \emptyset$ så gäller $P(A \cup B) = P(A) + P(B)$.

Om utfallsrummet är oändligt ersätts villkor 3 med

3’. Om $A_1, A_2, \ldots$ är en oändlig följd av parvis oförenliga händelser (dvs. $A_i \cap A_j = \emptyset$ för alla $i \neq j$) så gäller $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

ANMÄRKNING 2.1

I Axiom 1 räcker det att $P(A) \geq 0$. Att $P(A) \leq 1$ följer av de övriga axiomen, se Övning 2.10.

Alla tre villkoren i Kolmogorovs axiomsystem är självklara för alla utan att man reflekterat över det. Det första villkoret är en ren konvention – någon som pratar om negativa sannolikheter eller sannolikheter större än 1 betraktas med rätta med skepsis. Även det andra villkoret är en konvention – det har blivit praxis att den helt säkra händelsen, dvs. den som innehåller alla möjliga utfall, ges sannolikheten 1. Det tredje villkoret slutligen, säger att sannolikheten för två oförenliga händelser är lika med summan av sannolikheterna för var och en av händelserna. Förutsättningen att A och B skall vara oförenliga i villkor 3 är viktigt – i annat fall gäller inte utsagan.

Den vanligaste tolkningen av en sannolikhet, t.ex. $P(A) = 0.3$, är att om man upprepar slumpförsöket många gånger så kommer den relativa frekvensen för en händelse A ligga nära 0.3. Även för detta sätt att se på sannolikheter är utsagorna i Kolmogorovs axiomsystem självklara: den relativa frekvensen ligger ju alltid mellan 0 och 1, den relativa frekvensen för utfall i Ω är förstås 1 (alla utfall ligger ju i Ω så den relativa frekvensen av utfall i Ω blir 1). Slutligen blir den relativa frekvensen av unionen av oförenliga händelser lika med summan av de respektive relativa frekvenserna: inga utfall ingår ju i flera händelser, så antalet utfall i unionen blir lika med summan av antal utfall i respektive händelse.

EXEMPEL 2.4 (*Kortdragning, sannolikheter*)

I Exempel 2.2 betraktades försöket att dra ett kort slumpmässigt ur en kortlek. Det betyder att varje kort, dvs. varje utfall, har samma sannolikhet 1/52. (Denna vanliga slumpstruktur kallas likformig sannolikhetsfördelning, och tas upp i Definition 2.3 i nästa avsnitt.) Sannolikheten för en händelse blir därför antalet utfall i händelsen dividerat med 52. I exemplet definierades händelserna A =”klöver”, B =”femma” och C =”klädda”. Genom att räkna antal utfall i respektive händelse inser man snabbt att $P(A) = 13/52 = 1/4$, $P(B) = 4/52 = 1/13$ och $P(C) = 12/52 = 3/13$. Vi konstaterade att händelserna B och C var oförenliga ($B \cap C = \emptyset$). Således gäller $P(B \cup C) = P(B) + P(C) = 1/13 + 3/13 = 4/13$. $A \cap B =$ ”klöver fem”, så $P(A \cap B) = 1/52$.

Det är ofta klargörande att föreställa sig en sannolikhetsfunktion som att en enhet ”sannolikhetsmassa” smetas ut över ett Venndiagram (t.ex. Figur 2.1, sidan 2). Värdet $P(A)$ kan då ses som hur stor del av sannolikhetsmassan som ligger i händelsen A. Med denna tolkning är de tre villkoren i Kolmogorovs axiomsystem också självklara. Även följande sats är självklar med denna bildtolkning.

SATS 2.1

Låt A och B vara godtyckliga händelser i utfallsrummet Ω. Då gäller

1. $P(A^c) = 1 - P(A)$,
2. $P(\emptyset) = 0$,
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Det första resultatet säger att mängden sannolikhetsmassa utanför A är 1 (dvs. all sannolikhetsmassa) minus den som finns i A. Eftersom tomma mängden inte innehåller något utfall kan den ju inte ha någon sannolikhetsmassa, vilket är resultat 2. Om vi skall räkna ut hur mycket sannolikhetsmassa som finns i unionen av A och B kan vi göra detta genom att addera mängden sannolikhetsmassa som finns i A (område 1 och 2 i Figur 2.1) med mängden sannolikhetsmassa i B (område 2 och 3 i Figur 2.1). Vi ser då att vi räknat sannolikhetsmassan i område 2, dvs. $P(A \cap B)$, två gånger varför vi måste subtrahera detta tal.

Man kan även visa resultaten i Sats 2.1 mer formellt från axiomsystemets villkor. Vi gör detta för det tredje resultatet och lämnar övriga två resultat som övningar.

BEVIS, SATS 2.1, RESULTAT 3

Mängden $A \cup B$ kan skrivas som $A \cup (B \setminus A)$ där de två mängderna är oförenliga, dvs $A \cap (B \setminus A) = \emptyset$. Att så är fallet ses lätt i Figur 2.1 på sidan 2 där A utgör område 1 och 2 medan $B \setminus A$ är område 3. Från tidigare vet vi att $B \setminus A = B \cap A^c$, så villkor 3 i Kolmogorovs axiomsystem ger oss att

$$P(A \cup B) = P(A) + P(B \cap A^c).$$

Vidare kan händelsen B delas upp i de två disjunkta delarna $(B \cap A)$ och $(B \cap A^c)$, så från samma villkor får vi att $P(B) = P(B \cap A) + P(B \cap A^c)$, dvs. att $P(B \cap A^c) = P(B) - P(B \cap A)$. Om vi substituerar detta i föregående uttryck erhåller vi $P(A \cup B) = P(A) + P(B) - P(B \cap A)$, dvs. satsens tredje resultat.

EXEMPEL 2.5 (*Kortdragning, sannolikheter, forts.*)

Händelsen A betyder ”klöver” vilket gör att A^c betyder ”ej klöver” och sannolikheten för denna blir enligt satsen $1 - P(A) = 1 - 1/4 = 3/4$.

Detta överensstämmer med det andra sättet att räkna ut denna sannolikhet, nämligen genom att räkna utfallen i "icke-klöver" som är 39 och dividera detta med 52. Sannolikheten för "inget utfall" blir förstås $P(\emptyset) = 0$. Sannolikheten för händelsen $A \cup B$, som alltså utgörs av utfallen med klöver och/eller siffran 5, blir enligt satsen $P(A \cup B) = P(A) + P(B) - P(B \cap A) = 1/4 + 1/13 - 1/52 = 4/13$. Om vi i stället betraktar vilka utfall som ingår i $A \cup B$ är dessa $A \cup B = \{\text{K1}, \ldots, \text{K13}, \text{S5}, \text{H5}, \text{R5}\}$ som består av 16 utfall, varför händelsen får sannolikheten $16/52 = 4/13$ vilket alltså överensstämmer med det svar som satsen gav oss.

ÖVNING 2.4

Betrakta försöket att kasta en vanlig tärning, dvs. där utfallsrummet är $\Omega = \{1, \ldots, 6\}$ och alla utfall har samma sannolikhet, som alltså måste vara 1/6 vardera. Låt A vara händelsen att tärningen visar ett jämnt antal prickar och B vara händelsen att antalet prickar är delbart med 3. Ange vilka utfall som utgör händelserna A, B, $A \cap B$ och $A \cup B$ och beräkna motsvarande sannolikheter.

ÖVNING 2.5

Antag att för ett slumpförsök med två händelser A och B gäller $P(A) = 0.4$, $P(B) = 0.5$ och $P(A \cup B) = 0.6$. Beräkna $P(A \cap B)$.

ÖVNING 2.6

I ett lotteri finns tre vinstlotter, högsta vinsten H vinner man med sannolikheten $P(H) = 0.001$, näst högsta vinsten N vinner man med sannolikhet $P(N) = 0.01$ och tredje priskategori T vinner man med sannolikheten $P(T) = 0.1$. Bestäm sannolikheten för att överhuvud taget vinna, och sannolikheten att inte vinna (och således förlora satsat belopp).

ÖVNING 2.7

Bevisa resultat 1 i Sats 2.1 utifrån Kolmogorovs axiomsystem, de s.k. sannolikhetsaxiomen. (L)

ÖVNING 2.8

Bevisa resultat 2 i Sats 2.1 utifrån Kolmogorovs axiomsystem, de s.k. sannolikhetsaxiomen.

ÖVNING 2.9 (*Booles olikhet*)

Bevisa Booles olikhet, dvs. att för två godtyckliga händelser A och B gäller $P(A \cup B) \leq P(A) + P(B)$.

ÖVNING 2.10

Visa att den högra olikheten, $P(A) \leq 1$, i Axiom 1 följer av den vänstra olikheten, $P(A) \geq 0$, och de övriga axiomen. (L)

ÖVNING 2.11 (*Stjärnor*)

Antag att ett slumpexperiment består i att centrera ett stjärnkikarsikte mot en slumpvis vald stjärna. Låt A_n, $n = 1, 2, \ldots$, beteckna händelsen att man totalt i kikarsiktet ser exakt n stjärnor (eftersom kikarsiktet var centrerat mot en stjärna ser vi åtminstone en stjärna, dvs A_0 är inte aktuellt). Antag att $P(A_n) = c/n^2$ för någon konstant c.

a) Bestäm c. (L)
b) Beräkna sannolikheten för händelsen $B =$ ”högst 3 stjärnor syns i kikarsiktet”.
c) Beräkna $P(A_1^{\mathrm{c}})$ och ange i ord vad händelsen innebär.

ÖVNING 2.12

Visa att om $A_1, \ldots, A_n$ är disjunkta händelser så gäller att $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$. (L)

ÖVNING 2.13

Härled ett uttryck för $P(A \cup B \cup C)$ liknande det i Sats 2.1, sidan 8, för $P(A \cup B)$. Härled även en allmän form för $A_1 \cup A_2, \cdots \cup A_n$ för n händelser $A_1, \ldots, A_n$. (L)

ÖVNING 2.14

Formulera och visa Booles olikhet, se Övning 2.9,
a) för tre händelser, (L)
b) för n händelser, $n \geq 2$. (L)

2.3 Tolkning och exempel på sannolikheter

De tidigast förekommande sannolikheterna som studerades hade så kallad likformig sannolikhetsfördelning.

DEFINITION 2.3 (LIKFORMIG SANNOLIKHETSFÖRDELNING)

Ett slumpexperiment med ändligt utfallsrum säges ha *likformig sannolikhetsfördelning* om alla utfall har samma sannolikhet.

ANMÄRKNING 2.2

Vi har redan stött på likformig sannolikhetsfördelning i exemplen med kortdragning (Exempel 2.4 och dess fortsättning Exempel 2.5).

SATS 2.2 (KLASSISKA SANNOLIKHETSDEFINITIONEN)

För ett slumpexperiment med likformig sannolikhetsfördelning gäller att sannolikheten för en händelse är lika med antalet utfall i händelsen dividerat med antalet händelser i utfallsrummet, dvs. antalet gynnsamma utfall dividerat med antalet möjliga utfall. Om händelsen A innehåller $n(A)$ utfall och utfallsrummet totalt har $n(\Omega)$ utfall gäller alltså att $P(A) = n(A)/n(\Omega)$.

ANMÄRKNING 2.3

*Denna sats kallas klassiska sannolikhets*definitionen *av historiska skäl. Man talade då inte om olika sannolikhetsfördelningar varför $P(A) = n(A)/n(\Omega)$ kunde gälla som definition av sannolikheten för A. Numer finns många olika sannolikhetsfördelningar, som vi kommer se längre fram i boken, varför denna "definition" blir en följd av definitionen för likformig sannolikhetsfördelning.*

BEVIS

Antag att det finns n utfall $u_1, \ldots, u_n$ i utfallsrummet. Från Kolmogorovs sannolikhetsaxiom (Definition 2.2 på sidan 6) gäller att $P(\Omega) = 1$, och från generaliseringen av Kolmogorovs tredje axiom (Övning 2.12) gäller $P(\Omega) = \sum_{i=1}^{n} P(u_i)$. Eftersom alla $P(u_i)$ är identiska vid likformig sannolikhetsfördelning måste således $P(u_i) = 1/n$. Om en händelse A innehåller k utfall blir således $P(A) = \sum_{i;u_i \in A} P(u_i) = \sum_{i;u_i \in A} n^{-1} = k/n$, dvs antalet gynnsamma dividerat med antal utfall i utfallsrummet.

EXEMPEL 2.6 (*Slantsingling*)

Ett symmetriskt mynt singlas 3 gånger. De tre slantsinglingarna kan ge 8 olika resultat: $\Omega = \{(kl, kl, kl),\ (kl, kl, kr),\ (kl, kr, kl),\ (kl, kr, kr),\ (kr, kl, kl),\ (kr, kl, kr),\ (kr, kr, kl),\ (kr, kr, kr)\}$. Eftersom sannolikheten för krona är lika med sannolikheten för klave ($= 1/2$) vid varje enskilt kast har de 8 utfallen samma sannolikhet, vilken alltså måste vara $1/8$. Sannolikheten att det blir 2 klave kan man beräkna genom att se hur många av utfallen som ger två klave, vilket är 3 stycken: (kl, kl, kr), (kl, kr, kl) och (kr, kl, kl). Således blir sannolikheten för att få två klave $3/8$.

Tolkningen av påståendet ovan att ”sannolikheten att få två klave på tre slantsinglingar är 3/8” är nog uppenbar för de flesta. Nämligen att om man utför tre slantsinglingar många gånger, så bör andelen av dessa trippla slantsinglingar som gav upphov till just två klave vara ungefär 3/8. Detta är den s.k.*frekvenstolkningen* av sannolikheter. Frekvenstolkningen är också tillämpbar i många mer praktiska fall. Ett försäkringsbolag som bedömer att risken för att ett fritidshus brinner kommande år är 0.003 tolkar detta som att om försäkringsbolaget tecknar 1000 försäkringskontrakt med liknande fritidshus bör i genomsnitt 3 av dessa brinna det kommande året.

I andra sammanhang är det inte lika självklart hur man skall tolka uttalande om sannolikheter. Den vanligast förekommande tolkningen är nog den frekventistiska. För att ett sannolikhetspåstående skall kunna tolkas frekventistisk skall det vara så att, om ett slumpexperiment upprepas oberoende allt fler gånger, så kommer den relativa frekvensen för att händelsen inträffar att stabilisera sig vid sannolikheten för händelsen. Sannolikheten att få en sexa vid tärningskast är 1/6 just för att, om man gör många tärningskast så kom-

mer andelen sexor ligga mycket nära 1/6 (om den inte gör det finns anledning att tro att tärningen inte är helt symmetrisk).

I fallet med slantsingling, och för den delen kortdragning och många andra enklare slumpexperiment baseras de ofta ursprungligen på någon symmetri som man tar för givet, t.ex. att krona och klave har samma sannolikhet vid slantsingling eller att alla sidor på tärningen har samma sannolikhet att komma upp. Sådana sannolikheter brukar sägas vara *axiomatiska*, dvs sådana som tas för givet att de gäller (men som ibland kan bevisas statistiskt att inte gälla, mer om detta i Avsnitt **??**). Utifrån dylika axiomatiska sannolikheter kan erhålla s.k. *beräknade* sannolikheter, som t.ex. att sannolikheten att få två klavar på tre slantsinglingar är 3/8.

Beräknade sannolikheter förekommer inte sällan i olika riskbedömningar. Om t.ex. Statens kärnkraftsinspektion gör bedömingen att risken för en allvarlig olycka vid en given kärnkraftreaktor under ett år är approximativt 10^{-8}, så baseras detta på avancerade beräkningar (ibland involverande även simuleringar) som baseras på ett antal givna förutsättningar. Dessa förutsättningar, som att ett rör springer läck med sannolikheten 0.01 eller att en elledning går av med sannolikheten 0.005, är i sin tur axiomatiska sannolikheter i sammanhanget, förhoppningsvis väl underbyggda med empiri vilka då även kan kallas *skattade* sannolikheter. När (axiomatiska eller beräknade) är väldigt små, vilket ofta gäller vid riskberäkningar, är den frekventistiska sannolikheten inte lika lätt att tolka. Man kan förvisso i princip föreställa sig att man har 10^8 identiska kärnkraftreaktorer och att då i genomsnitt en bör ha en allvarlig olycka, men det känns ofta svårare att bilda sig en uppfattningen om sådana sannolikheter. När det gäller små sannolikheter är det ofta mer fruktbart att relatera olika sannolikheter. Om två i övrigt likvärdiga kärnkraftreaktorer beräknas ha risk för en allvarlig olycka 10^{-8} respektive 10^{-7} är ju förstås den förra att föredra eftersom denna löper 10 gånger mindre risk för allvarliga olyckor.

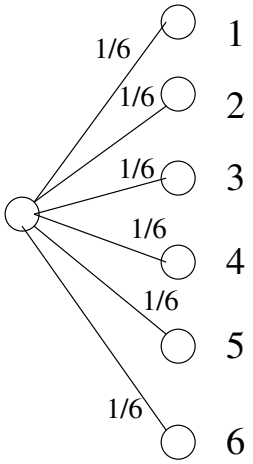
Det är inte alltid möjligt att tänka sig att upprepa ett försök många gånger. En som tror att The Ark vinner Eurovisionsschlagerfestivalen med sannolikhet 0.33 tänker nog inte att om tävlingen upprepades många gånger så skulle The Ark vinna var tredje tävling. Här har man att göra med (mer eller mindre väl underbyggda) *subjektiva sannolikheter*. Ofta kan dessa tolkas i termer av *odds*. Personen som säger att The Ark har 1/3 chans att vinna Eurovisionsschlagerfestivalen bör anse att det korrekta oddset om man satser pengar (på att The Ark vinner) är 3 till 1, dvs att man får tillbaka 3 gånger satsat belopp vid vinst och inget vid förlust.

Som sammanfattning kan man alltså säga att sannolikheter grovt sett kan tolkas antingen som frekventistiska eller subjektiva. Deras numeriska värde

kan i sin tur erhållas antingen axiomatiskt eller beräknas (från andra axiomatiska sannolikheter). Det förekommer även situationer där de numeriska värdena baseras på empiriska studier då man talar om skattade, eller *estimerade* sannolikheter. Hur man erhåller skattade sannolikheter tas upp mer i Kapitel **??**.

2.3.1 Träddiagram

Ett praktiskt sätt att åskådliggöra slumpexperiment, speciellt från diskreta utfallsrum, är genom att använda sig av s.k. *träddiagram*. Oftast ritas träddiagram från vänster till höger, och från startpunkten till vänster ritas de möjliga utfallen ut genom kanter åt höger. På respektive kant skriver man ut sannolikheten för utfallet (se Figur 2.3). Om man vill beräkna sannolikheten för



Figur 2.3. *Träddiagram för slumpexperimentet att kasta en tärning. Utfallen (= antal prickar) står till höger och sannolikheterna för respektive utfall anges vid respektive kant.*

en händelse, dvs. sannolikheten för en viss mängd av utfall, kan denna lätt beräknas med summaregeln:

SATS 2.3 (SUMMAREGELN (FÖR DISKRETA UTFALLSRUM))

Sannolikheten för en händelse är lika med summan av sannolikheterna för utfallen i händelsen.

Från träddiagrammet i Figur 2.3 ser man t.ex. att sannolikheten för att få minst 5 prickar vid kast med tärning är lika med sannolikheten att få 5 prickar plus sannolikheten att få sex prickar, som blir $1/6 + 1/6 = 1/3$. Vi återkommer till träddiagram längre fram då ett slumpexperiment sker i flera steg då det kommer till stor användning.

ÖVNING 2.15

Marie går i en skolklass med 31 elever (17 flickor och 14 pojkar). En person lottas slumpmässigt (likformigt) från klasslistan. Vad är sannolikheten att det blir Marie respektive sannolikheten att det blir en flicka?

ÖVNING 2.16

Avgör om följande sannolikheter främst bör tolkas frekventistisk eller subjektivt.

a) Sannolikheten att det regnar under en godtycklig dag i juli i Jokkmokk är 0.29.
b) Djurgården vinner allsvenskan med sannolikheten 1/3.
c) En viss kemisk process överhettas med sannolikheten 0.005.

ÖVNING 2.17

Bedöm om följande sannolikheter främst bör tolkas frekventistisk eller subjektivt,verkar vara axiomatiska, beräknade eller skattade.

a) Ett häftstift som kastas på ett på bord hamnar med spetsen upp med sannolikheten 0.613.
b) Sannolikheten att personen du sitter bredvid på på bussen fyller år idag är 1/365.
c) Sannolikheten att få fyrtal på första given (fyra av samma valör när man får 5 spelkort) är $624/2598960 \approx 0.000240$.

ÖVNING 2.18

Rita ett träddiagram för slumpexperimentet som består i att man köper en lott från ett lotteri där chansen för högvinst är 1/500 och chansen för lågvinst är 1/10; resterande lotter ger ingen vinst alls, s.k. nitlotter. Beräkna även chansen för att man vinner över huvud taget med hjälp av summaregeln.

2.4 Kombinatorik

I Exempel 2.6 visades att det fanns tre olika sätt på vilket man kunde få två klave på tre försök, dvs. det var tre sätt på vilket de två ”klavesinglingarna” kunde väljas ut bland de tre slantsinglingarna.

Vi ska nu reda ut på hur många sätt man kan göra olika val. Vårt första steg i den riktningen är den s.k. *multiplikationsprincipen*. Den säger att om vi har två val att göra, och i första valet har j alternativ och i andra valet har k alternativ så finns det totalt $j \cdot k$ kombinationer av val att göra.

En annan typ av val man ofta stöter på i sannolikhetsteori och annan matematik är s.k. val utan återläggning. Det handlar då om att beräkna hur många gånger man kan välja ut k ”element” bland n ”element” (i Exempel 2.6 var $k = 2$, $n = 3$ och det som valdes var positionerna bland de tre kasten där man fick klave). Svaret beror på om vi tar hänsyn till i vilken ordning de k elementen väljs ut eller inte.

Om man tar hänsyn till ordningen i vilket de k elementen väljs ut kan man enligt multiplikationsprincipen välja det första elementet på n sätt, det andra kan man därefter välja på $n - 1$ och så vidare fram till det k:te som kan väljas på $n - k + 1$ sätt. Totalt kan detta således ske på $n(n-1)\cdots(n-k+1)$ sätt. Oftare är man dock inte intresserad av i vilken ordning de k elementen har valts utan bara vilka element som valts. Om vi har valt ut k element kan detta ha skett i ett antal olika ordningsföljder. Det först utvalda kan ha varit vilket som helst av de k elementen, det därpå följande utvalda kan ha varit vilket som helst av de $k - 1$ återstående, osv. Det finns således $k(k-1)\cdots 1$ olika ordningsföljder som kan ha givit de k utvalda elementen. Denna produkt skrivs ofta som $k! = k(k-1)\cdots 1$ och utläses ”k-fakultet” (t.ex. är $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$). För talet noll måste man införa en separat definition som är $0! = 1$ (man kan argumentera för att detta är den naturliga definition men vi går inte vidare in på detta).

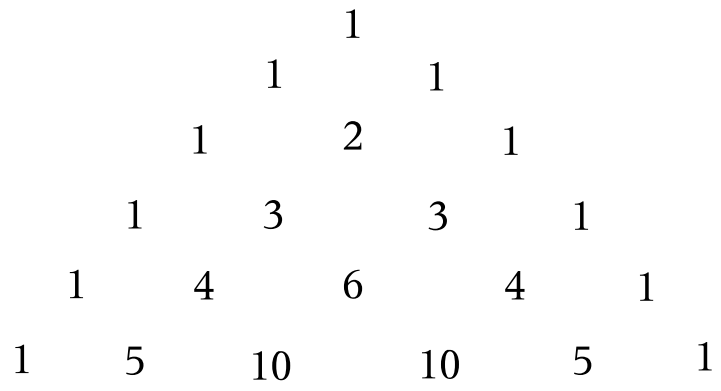
Vill man inte ta hänsyn till ordningsföljden ger alla dessa $k!$ ordningsföljder samma k utvalda element. Slutsatsen är således att antalet sätt med vilket man kan välja ut k element bland n, utan hänsyn till ordning, är lika med $n(n-1)\cdots(n-k+1)/k!$ vilket är detsamma som $n!/(k!(n-k)!)$. Eftersom detta är ett vanligt förekommande begrepp har det fått en egen beteckning, $\binom{n}{k}$ som läses ”n över k” och kallas för en binomialkoefficient. På engelska läses detta ”n choose k” vilket bättre beskriver vad binomialkoefficienten innebär. Av samma skäl hade kanske ”n välj k” varit en bättre utläsning av notationen, men n över k är redan etablerat. Vi formulerar resultaten ovan i en sats.

SATS 2.4

1. Om vi har två val att göra, och i första valet har j alternativ och i andra valet har k alternativ så finns det totalt $j \cdot k$ kombinationer av val att göra.
2. Antalet sätt att ordna n element är lika med $n! = n(n-1)(n-2)\cdots 1$.
3. Antalet sätt att välja ut k element bland n ($n \geq k \geq 0$) utan hänsyn till ordning ges av binomialkoefficienten

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

Ett snyggt sätt att illustrera binomialkoefficienterna upptäcktes av den franske matematikern Blaise Pascal och kallas därför Pascals triangel, se Figur 2.4. Talen i figuren erhålls genom att skriva ettor nedåt längs kanterna



Figur 2.4. *De första 6 raderna av Pascals triangel.*

och därefter fylla i de inre elementen uppifrån genom att på varje plats skriva in summan av de två elementen snett ovanför (till höger respektive vänster). T.ex. blir den femte radens tredje tal 6 eftersom talen snett ovanför är 3 respektive 3 vilket summerat blir 6. Från tabellen får man binomialkoefficienterna på följande sätt. Om vi t.ex. vill se hur stort $\binom{4}{2}$ är går vi till 4+1=5:e

raden ovanifrån (n=0 har en egen rad) och går till den 2+1=3:e ($k = 0$ har en egen position) positionen från vänster. Där ser vi att det står 6, så $\binom{4}{2} = 6$.

EXEMPEL 2.7 (*Lotto*)

Spelet ”Lotto” går ut på att 6 bollar dras slumpmässigt ur en urna innehållande 36 bollar numrerade från 1 till 36 (vi bortser från tilläggsnummer). Den erhållna lottoraden ges av numren på de dragna bollarna och den presenteras i växande ordning oavsett ordningen i vilken bollarna drogs. Antalet möjliga rader är således $\binom{36}{6} = 36{\cdot}35 \cdots 31/6! = 1\,947\,792$. Om varje uppsättning bollar, dvs. varje möjlig rad, har samma sannolikhet betyder det att sannolikheten att få alla rätt på Lotto $1/1\,947\,792 \approx 0.000\,000\,513$ om man tippar en rad.

ÖVNING 2.19

En reklamförsäljare som ringer upp presumtiva kunder väljer telefonnummer med hjälp av en slumptalsgenerator som slumpar nya nummer oberoende av varandra. Vi antar för enkelhets skull att alla telefonnummer är sexsiffriga och att slumptalsgeneratorn väljer första siffran bland 1–9 (dvs. 0 kan inte väljas).

a) Vad är sannolikheten att telefonnumret slutar på 0?
b) Vad är sannolikheten att två på varandra slumpade nummer har samma startsiffra?

ÖVNING 2.20

För att undvika att interner blir alltför nära vänner grupperas de ofta olika vid olika tillfällen. På hur många sätt kan man dela in 6 fångar i

a) två lika stora grupper?
b) två grupper (av samma eller olika storlekar)?

ÖVNING 2.21

Vid en gruppuppgift i högskolan skall 8 studenter delas in i två lika stora grupper.

a) På hur många sätt kan detta ske?

b) Antag att man även vid nästa lektionstillfälle gör en ny gruppindelning helt slumpmässigt och oberoende av gruppindelningen gången innan. Vad är sannolikheten att du som student har samma tre studenter i gruppen bägge gångerna? Vad är sannolikheten att det blir två studenter från förra gruppen, en från förra gruppen, respektive ingen student från förra gruppen?

ÖVNING 2.22

Bevisa likheten $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$. (L)

2.5 Betingning och oberoende

Vi ska nu presentera de två viktiga begreppen *betingning* och *oberoende* som bägge rör (minst) två händelser.

Ibland är man intresserad av sannolikheten för en viss händelse B betingat av att man vet att en annan händelse A har inträffat. Detta skrivs som $P(B \mid A)$ och läses som sannolikheten för B betingat av (eller givet) att A har inträffat.

EXEMPEL 2.8 (*Kortdragning*)

Antag att vi skall dra två kort från en kortlek, och låt A vara händelsen att första kortet är ett ess och B händelsen att andra kortet är ett ess. Då är $P(B \mid A) = 3/51$ eftersom, om första kortet blev ett ess vilket ju var händelsen A som vi betingar på, det finns 51 kort kvar vid andra dragningen varav 3 är ess.

Nedan följer en definition av betingad sannolikhet i termer av redan kända begrepp. Låt oss först motivera definitionen. Vi vet alltså att A har inträffat och undrar vad den därav betingade sannolikheten för att B skall inträffa är (vi antar även att $P(A) > 0$)). Denna sannolikhet måste ju vara större ju ”mer” av händelsen A som finns i B. Om t.ex. A och B är oförenliga, dvs. $A \cap B = \emptyset$ och således $P(A \cap B) = 0$, så måste vi ju ha $P(B \mid A) = 0$. Om å andra sidan hela A finns i B, dvs $A \cap B = A$ och $P(A \cap B) = P(A)$, så skall det gälla att $P(B \mid A) = 1$ eftersom B måste inträffa om A inträffar. Definitionen nedan uppfyller dessa krav.

DEFINITION 2.4 (BETINGAD SANNOLIKHET)

Antag att för händelen A gäller $P(A) > 0$. Den *betingade sannolikheten* för händelsen B, betingat av att A har inträffat, skrivs $P(B \mid A)$ och definieras som

$$P(B \mid A) := \frac{P(B \cap A)}{P(A)}.$$

EXEMPEL 2.9 (*EU-inställning*)

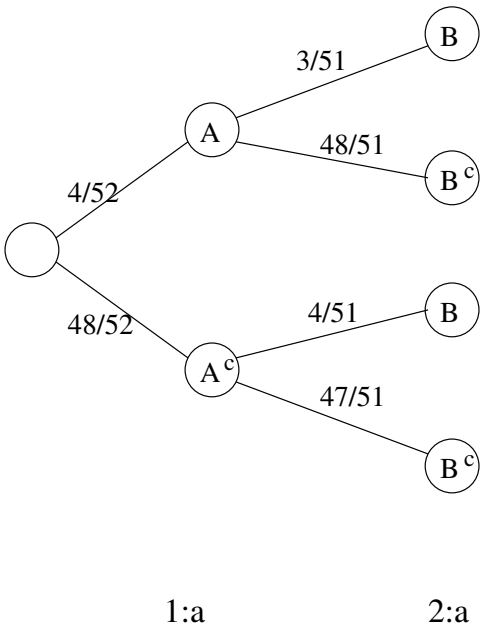
Vid en opinionsmätning ställs två frågor: ”Vill Du ha en folkomröstning om EU:s nya grundlag?” respektive ”Stödjer Du EU:s nya grundlag?”. Andelen som vill ha folkomröstning visar sig vara 37% och andelen som vill ha folkomröstning och samtidigt stöder EU:s nya grundlag är 9%. Förutsatt att opinionsmätningen är representativ betyder detta att om man väljer en individ slumpmässigt ur befolkningen och den visar sig vara positiv till folkomröstning så är den därav betingade sannolikheten att individen är positiv till EU:s nya grundlag lika med $P(E \mid F) = P(E \cap F)/P(F) = 0.09/0.37 = 0.24$, där E är händelsen att individen är positivt till EU:s nya grundlag och F händelsen att personen är positiv till att ha folkomröstning.

EXEMPEL 2.10 (*Fortkörning*)

Vid en hastighetskontroll vid en skola visar det sig att 30% av fordonen överskrider tillåten hastighet och att 4% överskrider tillåten hastighet med mer än 30 km/h (vilket resulterar i indraget körkort). Om vi låter A vara händelsen att ett fordon körs för fort och B händelsen att man kör mer än 30 km/h för fort får man således att sannolikheten för att en fortkörning resulterar i indraget körkort till $P(B \mid A) = P(B \cap A)/P(A) = P(B)/P(A) = 0.04/0.3 = 0.133$. Observera här att händelsen B ligger inuti händelsen A, vilket skrivs som $B \subset A$ (kör man mer än 30 km/h för fort så kör man ju för fort) så $B \cap A = B$ och $P(B \cap A) = P(B)$.

Betingade sannolikheter dyker upp på ett naturligt sätt i träddiagram med slumpexperiment i flera steg. Om vi t.ex. först tänker oss att ett slumpexperiment avgör om A eller dess komplement A^c inträffar, och därefter att ett experiment avgör om B inträffar eller ej, då kan detta illustreras med ett

träddiagram där betingade sannolikheter dyker upp. Låt som i Exempel 2.8 försöket bestå i att dra två kort ur en kortlek och A beteckna händelsen att det första kortet är ett ess och B händelsen att det andra kortet är ett ess (se Figur 2.5). $P(A)$ blir 4/52 eftersom vi drar ett kort på måfå och det finns 4 ess bland



Figur 2.5. *Träddiagram för försöket att dra två kort efter varandra (utan återläggning). A är händelsen att 1:a kortet är ett ess och B händelsen att 2:a kortet är ett ess.*

52 kort. Om vi vet att vi fick ett ess i första dragningen så är sannolikheten att få ett ess i andra 3/51 eftersom det finns 3 ess kvar bland 51 kort, dvs. $P(B \mid A) = 3/51$. Om vi å andra sidan *inte* fick ett ess i första dragningen (dvs. A^c har inträffat) så är sannolikheten att få ett ess i andra dragningen $P(B \mid A^c) = 4/51$. Sannolikheten för att både händelse A och händelse B inträffar (dvs. $P(A \cap B)$) blir produkten av sannolikheterna längs respektive väg, dvs. $P(A)P(B \mid A) = 4/52 \cdot 3/51 = 1/221 \approx 0.00452$. Detta följer förvisso av definitionen för betingad sannolikhet, men vi formulerar det ändå som en sats för träddiagram.

SATS 2.5 (PRODUKTREGELN)

I ett träddiagram med flera nivåer ges sannolikheten för en väg av produkten av sannolikheterna längs vägen.

Slutligen, sannolikheten att andra kortet är ett ess, dvs. $P(B)$, är inte lika självklar vad den skall vara. Händelsen B kan ju inträffa antingen i kombination med att A inträffar eller att A^c inträffar. Sannolikheten $P(A \cap B)$ har vi redan räknat ut och $P(A^c \cap B)$ blir på motsvarande sätt $P(A^c)P(B \mid A^c)$. Vi utnyttjar så summaregeln för att komma fram till att $P(B) = P(A)P(B \mid A) + P(A^c)P(B \mid A^c) = 4/52 \cdot 3/51 + 48/52 \cdot 4/51$. Denna relation gäller i allmänhet och inte bara för detta exempel. Motsvarande relation om man delar upp utfallsrummet i godtyckligt många delar, i stället för bara A och A^c som här, gäller också och kallas för lagen om total sannolikhet, se Sats 2.6 längre fram på sidan 25.

När man är intresserad av sannolikheter för flera olika händelser i ett slumpexperiment är det ofta av intresse om händelserna beror av varandra. Om de inte gör det sägs händelserna vara oberoende, ett begrepp vi nu skall definiera. I vardagligt tal tolkar man uttrycket att ”två händelser sker *oberoende* av varandra” som att händelserna inte har med varandra att göra, dvs. att vetskapen om att den ena händelsen inträffat inte påverkar vad vi tror om den andra händelsen.

EXEMPEL 2.11 (*Lotto och regn i Peru*)

Betrakta händelserna L =att vinna på Lotto en viss dag, och R = att det regnar i Peru samma dag. Dessa händelser har uppenbarligen inget med varandra att göra varför vi säger att de är oberoende. Med detta menar vi att sannolikheten att vinna på Lotto är densamma densamma vare sig det regnar i Peru eller ej, och omvänt att sannolikheten för regn i Peru är densamma vare sig vi vinner på Lotto eller ej. Det betyder alltså att $P(L \mid R) = P(L)$ och $P(R \mid L) = P(R)$.

Detta exempel motiverar följande definition av oberoende.

DEFINITION 2.5 (OBEROENDE HÄNDELSER)

Två händelser A och B är *oberoende* om $P(A \mid B) = P(A)$ förutsatt att $P(B) > 0$, och $P(B \mid A) = P(B)$ förutsatt att $P(A) > 0$.

ANMÄRKNING 2.4

Av definitionen följer att om A och B är oberoende så är även A och B^c, A^c och B samt A^c och B^c oberoende. Beviset av detta överlåter vi till läsaren (Övning 2.24).

Man kan även definiera oberoende utan att använda sig av betingade sannolikheter vilket dock inte lika tydligt stämmer med vardagligt bruk av ordet oberoende. Matematiskt blir definition något enklare eftersom man inte behöver förutsätta att vissa sannolikheter är större än noll, vilket gjordes ovan beroende på att betingade sannolikheter definieras som ett bråk vilket förutsätter att nämnaren är positiv.

DEFINITION 2.6 (ALTERNATIV DEFINITION: OBEROENDE HÄNDELSER)

Två händelser A och B är *oberoende* om

$$P(A \cap B) = P(A)P(B).$$

ANMÄRKNING 2.5

Under förutsättning att händelserna ifråga har positiv sannolikhet är denna definition ekvivalent med den ursprungliga. Förutsatt att händelserna är oberoende enligt den alternativa definition får vi nämligen att $P(A|B) = P(A \cap B)/P(B) = P(A)P(B)/P(B) = P(A)$. Den omvända relation visas analogt.

ANMÄRKNING 2.6

En mängd händelser $A_1, A_2, \ldots$ sägs vara parvis oberoende *om för alla par $i \neq j$ gäller att $P(A_i \cap A_j) = P(A_i)P(A_j)$.*
Mängden händelser sägs vara fullständigt oberoende *om det för alla distinkta delmängder $\{A_{i_1}, \ldots, A_{i_k}\}$ med $i_1 < \cdots < i_k$ gäller att $P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$. Motsvarande definitioner kan även göras med betingade sannolikheter om händelserna antas ha positiv sannolikhet.*

Att händelser kan vara parvis oberoende utan att vara fullständigt oberoende visas med följande exempel.

EXEMPEL 2.12 (*Parvis men inte fullständigt oberoende*)

Betrakta försöket att kasta två tärningar, en röd och en svart. Låt A vara händelsen att den röda tärningen visar udda antal prickar, B händelsen att den svarta tärningen visar udda antal prickar, och slutligen C händelsen att summan blir udda. Om bägge tärningarna visar udda blir ju summan jämn, så $P(A \cap B \cap C) = 0$. Eftersom de tre händelserna var och en har positiv sannolikhet blir produkten av sannolikheterna positiv, $P(A)P(B)P(C) > 0$, så händelserna är uppenbarligen inte fullständigt oberoende. De är emellertid parvis oberoende vilket vi nu skall visa. Händelserna A och B är ju uppenbart oberoende eftersom de två tärningarnas utfall ej beror av varandra, dvs $P(A \cap B) = P(A)P(B) = 3/6 \cdot 3/6 = 1/4$. Vi beräknar nu $P(C \cap A) = P(C \mid A)P(A)$. Denna blir ju lika med $P(C)P(A)$ om vi kan visa att $P(C \mid AA) = P(C)$, dvs att sannolikheten för att summan är udda betingat på att den röda tärningen visar udda antal prickar är densamma som den obetingade sannolikheten att summan är udda. Men om den röda visar udda antal prickar blir summan jämn om den svarta visar jämt antal prickar, alltså sker detta med sannolikhet 3/6=1/2. Vad är då den obetingade sannolikheten att summan blir udda, dvs $P(C)$? Totalt kan den röda och svarta träningen resultera i 36 olika utfall: $(1,1), (1,2), \ldots, (1,6), (2,1), \ldots, (6,6)$, och alla utfall har samma sannolikhet 1/36. Man kan lätt förvissa sig om att 18 av utfallen resulterar i udda antal prickar (t.ex. ser man att för varje utfall av röda tärningen har 3 av 6 utfall udda summa) vilket implicerar att $P(C) = 18/36 = 1/2$. Vi har därmed visat att $P(C \cap A) = P(C)P(A)$ och $P(C \cap B) = P(C)P(B)$ visas helt analogt – färgen på tärning spelar ingen roll. Således är de tre händelserna parvis oberoende men inte fullständigt oberoende.

Ett vanligt fall av oberoende händelser är om ett slumpförsök upprepas, men där utfallen inte har med varandra att göra. Om man t.ex. mäter blodtrycket på två slumpvis valda patienter eller om man mäter flyghastigheten hos två starar som inte flyger tillsammans kan utfallen av de två blodtrycken anses oberoende, liksom även stararnas hastigheter. Händelser förknippade med olika upprepningar av respektive försök blir då oberoende.

EXEMPEL 2.13 (*Färgblindhet och diabetes*)

En stor medicinsk studie slår fast att förekomsten av färgblindhet (F) är oberoende av diabetes (D). Om andelen färgblinda i populationen är $P(F) = 0.042$ och andelen diabetiker är $P(D) = 0.02$, betyder det att andelen färgblinda bland diabetikerna är lika med $P(F \mid D) = P(F) =$

0.042 och att andelen med bägge symptomen är lika med $P(D \cap F) = P(D)P(F) = 0.00048$.

ANMÄRKNING 2.7 (*Oberoende är inte samma sak som oförenliga!*)
Det händer ofta att begreppen oberoende *och* oförenliga *blandas samman. Dessa betyder emellertid helt olika saker och är snarare varandras motsats. Två händelser är oförenliga (disjunkta) om $A \cap B = \emptyset$, medan A och B är oberoende om $P(B \mid A) = P(B)$. I det förra fallet är A och B inte oberoende ty $P(B \mid A) = P(B \cap A)/P(A) = 0/P(A) = 0$, alltså inte lika med $P(B)$.*

2.5.1 Lagen om total sannolikhet

Ibland kan det vara svårt att räkna ut en sannolikhet medan den skulle ha varit lättare att beräkna om man haft någon ytterligare information. Man kan då använda sig av lagen om total sannolikhet.

SATS 2.6 (LAGEN OM TOTAL SANNOLIKHET)

Låt $A_1, \ldots A_n$ vara oförenliga händelser sådana att $P(A_i) > 0,\ i = 1, \ldots n$, och anta att händelserna tillsammans utgöra hela utfallsrummet (dvs. $A_i \cap A_j = \emptyset$, $i \neq j$, och $\cup_{i=1}^{n} A_i = \Omega$). Då gäller

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i)P(A_i).$$

BEVIS

Från satsens förutsättningar gäller att $B = \cup_{i=1}^{n}(B \cap A_i)$, vi delar helt enkelt upp händelsen B beroende på vilken A_i-händelse utfallen ligger i. Eftersom A_i-mängderna är oförenliga är även de mindre mängderna $B \cap A_i$ oförenliga. Vi kan då applicera punkt 3' i Kolmogorovs axiomsystem (Definition 2.2, sidan 6), dvs.
$P(B) = P(\cup_{i=1}^{n}(B \cap A_i)) = \sum_{i=1}^{n} P(B \cap A_i)$.
Slutligen ger definitionen för betingad sannolikhet att $P(B \cap A_i) = P(B \mid A_i)P(A_i)$ varmed satsen är bevisad.

Om vi i Figur 2.5 på sidan 21 vill beräkna sannolikheten att andra kortet blir ett ess, dvs. $P(B)$, så ser vi att man kan ”nå” B på två sätt, antingen genom att A (dvs. första kortet blir ett ess) inträffar och därefter B, eller att A^c inträffar (dvs. första kortet blir inte ess) och därefter B. Sannolikheterna för dessa två varianter blir $4/52 \cdot 3/51$ respektive $48/52 \cdot 4/51$ vilket tillsammans blir 4/52. Vid närmare eftertanke är det inte så konstigt att sannolikheten att få ett ess i andra dragningen är densamma som i första så länge vi inte känner till vilket kort som drogs först. Om sannolikheten t.ex. vore större att få ess andra gången kunde man ju i så fall bara flytta på det översta kortet i leken utan att titta på det, och därefter ha förhöjd sannolikhet att få ess när man drog andra kortet.

EXEMPEL 2.14 (*TBC-test*)

Antag att förekomsten av TBC-smitta i en viss delbefolkning är 20%. Det snabbtest man kan utföra för att testa förekomst av TBC-smitta är inte perfekt. Sensitiviteten, dvs. sannolikheten att en smittad person ger ett positivt test (vilket man vill) är 0.9 (och 0.1 att det blir negativt utslag), medan specificiteten, dvs. sannolikheten att en icke-smittad ger ett negativt test (vilket man också vill) är 0.7 (och 0.3 att det blir positivt utslag). Sannolikheten att en slumpvis utvald person ger ett positivt test kan då erhållas med hjälp av lagen om total sannolikhet. Låt ”+” beteckna händelsen att personen testar positivt, och ”-” beteckna händelsen att personen testar negativt, och låt S vara händelsen att personen verkligen är smittad av TBC. De storheter som givits är då $P(S) = 0.2$, $P(+ \mid S) = 0.9$, $P(- \mid S) = 0.1$, $P(- \mid S^c) = 0.7$ och $P(+ \mid S^c) = 0.3$: Sannolikheten att vara smittad är 0.2, sannolikheten att testet ger positivt utslag hos smittade är 0.9, och att testet visar negativt bland icke-smittade har sannolikhet 0.7. Det gäller vidare att $P(S^c) = 1 - P(S) = 0.8$. Den eftersökta sannolikheten att en slumpvis vald person testar positivt, dvs. $P(+)$, blir därför enligt lagen om total sannolikhet

$$P(+) = P(+ \mid S)P(S) + P(+ \mid S^c)P(S^c) = 0.9 \cdot 0.2 + 0.3 \cdot 0.8 = 0.42.$$

I detta exempel har vi använt oss av att betingade sannolikheter också uppfyller kriterierna för att vara sannolikhetsmått – t.ex. sa vi att eftersom $P(+ \mid S) = 0.9$ så måste $P(- \mid S) = 0.1$. Att så är fallet visas mer stringent i Avsnitt 2.6.

2.5.2 Bayes sats

Ibland kan man vara intresserad av sannolikheten för en viss händelse betingat av en annan händelse, när man känner till den omvända betingade sannolikheten och de obetingade sannolikheterna. Man har då användning av Bayes sats.

SATS 2.7 (BAYES SATS)

Under samma villkor som för lagen om total sannolikhet, Sats 2.6 på sidan 25, gäller

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}.$$

ANMÄRKNING 2.8

För en viss typ av statistisk slutledningsprincip är denna sats fundamental vilket även antyds av dess namn Bayesiansk statistik*. Denna statistiska princip, som för övrigt fått ett enormt uppsving sedan slutet av 1900-talet i och med datorernas ökade beräkningskapacitet, beskrivs i Avsnitt* **??** *på sidan* **??**.

BEVIS, SATS 2.7

Enligt definitionen för betingad sannolikhet gäller att $P(A_i \mid B) = P(A_i \cap B)/P(B)$. Om vi använder definitionen för betingad sannolikhet igen för täljaren (men 'baklänges') får vi $P(A_i \cap B) = P(A_i)P(B \mid A_i)$ vilket ger samma uttryck som i satsens täljare. Att $P(B)$ blir detsamma som satsens nämnare följer av lagen om total sannolikhet.

EXEMPEL 2.15 (*TBC-test, forts från Exempel 2.14*)

Vi skall nu beräkna sannolikheten att en person som testar positivt verkligen är smittad, vilket man kan göra med Bayes sats. Den sannolikhet vi vill beräkna är alltså $P(S \mid +)$, dvs. sannolikheten att en person som testar positiv ("+") är smittad ("S"). Vi får

$$\begin{aligned} P(S \mid +) &= \frac{P(S)P(+ \mid S)}{P(S)P(+ \mid S) + P(S^{\mathrm{c}})P(+ \mid S^{\mathrm{c}})} = \frac{0.2 \cdot 0.9}{0.2 \cdot 0.9 + 0.8 \cdot 0.3} \\ &= 9/21 \approx 0.429. \end{aligned}$$

Alltså, trots att testet är ganska bra – det visar ju rätt med sannolikhet 0.9 respektive 0.7, är det bara ca 43% sannolikhet att man verkligen är sjuk om testet givit positivt utfall. Orsaken till detta är att majoriteten av befolkningen (80%) inte är sjuka, men bland dessa ger testet trots allt positivt utslag ibland.

Ett generellt tips som underlättar problemlösning som rör flera händelser och kanske även betingningar, är att tydligt definiera lämpliga händelser, gärna med passande beteckningar. Vilka händelser som är lämpliga att definiera beror förstås på vilket problem man studerar men också vilken sannolikhet man är ute efter. Om man definierat lämpliga händelser och även lyckats uttrycka eftersökt (betingad) sannolikhet med hjälp av dessa händelser så har man ofta kommit en mycket god väg mot lösningen av problemet.

ÖVNING 2.23

En symmetrisk tärning kasta. Låt A vara händelsen att tärningen visar ett udda antal prickar, B händelsen att antal prickar blir minst 4 och C händelsen att antalet prickar är delbart med 3.

a) Är A och B oberoende? (Motivera ditt svar)
b) Är A och C oberoende? (Motivera ditt svar)
c) Är B och C oberoende? (Motivera ditt svar)

ÖVNING 2.24

Visa att om A och B är oberoende så är även A och B^c oberoende, A^c och B oberoende samt A^c och B^c oberoende. (L)

ÖVNING 2.25

Två typer av fel vid produktionen av mobiltelefoner uppstår oberoende av varandra. Om det ena felet förekommer med frekvensen 1 på 10 000 och det andra med frekvensen 3 på 10 000, hur ofta uppstår bägge felen i en och samma mobiltelefon?

ÖVNING 2.26

Beräkna sannolikheten att TBC-testet i Exempel 2.14 ger rätt utslag. (L)

ÖVNING 2.27

Betrakta försöket att kasta två tärningar. Antag att summan av tärningskasten blev 4. Beräkna den därav betingade sannolikheten att

a) den första tärningen gav en 3:a,
b) den andra tärningen gav 2 eller mindre,
c) bägge tärningarna visade ett udda antal prickar.

(L)

ÖVNING 2.28

En hatt innehåller 100 lappar varav 5 ger en vinst. Först drar Elvira en lapp varefter Filippa drar en. Låt E vara händelsen att Elvira drar en vinstlapp och F händelsen att Filippa drar en vinstlapp.

a) Ange $P(E)$, $P(F \mid E)$ och $P(F \mid E^{\mathrm{c}})$.
b) Beräkna $P(F)$.
c) Är E och F oberoende händelser?

ÖVNING 2.29

Antag att en uppföljningsstudie av forna studenter ger följande resultat. Bland tidigare humanistiska studenter är 61% privatanställda, 31% offentliganställda och 8% arbetslösa. Motsvarande procentsiffror för samhällsvetenskapliga studenter är 54%, 40% respektive 6%, för naturvetarstudenter 67%, 29% respektive 4%, och slutligen 73%, 23% respektive 4% för teknologstudenter. Antag vidare att studentkullarnas storlekar består av 20% humaniststudenter, 37% samhällsvetenskapliga studenter, 21% naturvetarstudenter och 22% teknologstudenter (vi bortser från övriga studentkategorier). En tidigare student (bland ovan angivna studentkategorier) väljs på måfå.

a) Inför lämpliga beteckningar för de olika händelserna ovan.
b) Klargör vilka sannolikheter frekvenserna ovan svarar mot.
c) Vad är sannolikheten att den valda studenten är privatanställd?
d) Givet att studenten var arbetslös, vad är då sannolikheten att han/hon var teknolog?

ÖVNING 2.30

Antag att var femte bilolycka med dödlig utgång har en onykter förare. Om detta skall användas som argument för att det är farligt att köra onykter (snarare än att Sveriges befolkning har alkoholproblem) beror på hur stor andel av förarna längs Sveriges vägar som är onyktra.

a) Definiera lämpliga händelser för att vara onykter i trafiken och att dö i trafikolycka.
b) Uttryck frekvensen ovan med hjälp av sannolikheter för ett lämpligt slumpexperiment.
c) Vilken sannolikhetsrelation bör gälla för att utsagan ovan ska kunna tolkas som att onykterhet inte påverkar risken att dö i bilolycka huruvida man kör onykter eller ej, dvs. att dödsrisken är oberoende av nykterhet/onykterhet?
d) Vilken sannolikhetsrelation bör gälla för att utsagan ovan ska kunna tolkas som att det är farligt att köra onykter?

ÖVNING 2.31 (*Stjärnor, forts från Övning 2.11*)

Antag att ett slumpexperiment består i att centrera ett kikarsikte mot en slumpvis vald stjärna. Låt A_n, $n = 1, 2, \ldots$, beteckna händelsen att man i kikarsiktet ser exakt n stjärnor. Antag att $P(A_n) = 6/(\pi^2 n^2)$. Antag vidare att stjärnfall sker oberoende för olika stjärnor och att sannolikheten att en given stjärna har ett stjärnfall under en timme är 10^{-6}.

a) Beräkna numeriskt sannolikheten att se något stjärnfall om man tittar en timme i kikarsiktet. (L)
b) Beräkna numeriskt sannolikheten att bara se en stjärna, och beräkna sannolikheten för denna händelse givet att man sett ett stjärnfall.

2.6 Sannolikhetsmått

I Avsnitt 2.2 definierade vi sannolikheter genom att utgå från Kolmogorovs axiomsystem (Definition 2.2 på sidan 6). Utgående från dessa axiom visade vi olika egenskaper hos sannolikheter, som t.ex. komplementsatsen och additionssatsen, se Sats 2.1 på sidan 8.

Vi kan uttrycka detta mer allmänt genom att studera funktioner definierade på delmängder till någon grundmängd Ω.

DEFINITION 2.7 (SANNOLIKHETSMÅTT)

En funktion P definierad på delmängder till en grundmängd Ω som uppfyller Kolmogorovs axiomsystem sägs vara ett *sannolikhetsmått* på Ω.

För varje sannolikhetsmått gäller de resultat som vi härlett utifrån axiomen, t.ex. komplementsatsen $P(A^c) = 1 - P(A)$ och additionssatsen $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Vi kan utnyttja detta för att verifiera några egenskaper hos betingade sannolikheter som vi faktiskt redan utnyttjat.

SATS 2.8

Låt P vara ett sannolikhetsmått på Ω och A_1 vara en given händelse med $P(A_1) > 0$. Då är Q, som för varje händelse A definieras som

$$Q(A) := P(A \mid A_1) = \frac{P(A \cap A_1)}{P(A_1)},$$

ett sannolikhetsmått på Ω.

BEVIS

Vi vet att P är ett sannolikhetsmått, dvs. uppfyller axiomen, och vi ska visa att även Q gör det.
1) Visa att $Q(A) \geq 0$ för alla $A \subset \Omega$.

$$Q(A) = P(A \mid A_1) = \frac{P(A \cap A_1)}{P(A_1)} \geq 0,$$

eftersom $P(A \cap A_1) \geq 0$ enligt Axiom 1 och $P(A_1)$ enligt satsens antagande.
2) Visa att $Q(\Omega) = 1$.

$$Q(\Omega) = P(\Omega \mid A_1) = \frac{P(\Omega \cap A_1)}{P(A_1)} = \frac{P(A_1)}{P(A_1)} = 1.$$

3) Låt A och B vara godtyckliga händelser med $A \cap B = \emptyset$. Visa att $Q(A \cup B) = Q(A) + Q(B)$.

$$\begin{aligned} Q(A \cup B) &= P(A \cup B \mid A_1) = \frac{P((A \cup B) \cap A_1)}{P(A_1)} \\ &= \frac{P((A \cap A_1) \cup (B \cap A_1))}{P(A_1)} = \frac{P(A \cap A_1) + P(B \cap A_1)}{P(A_1)} \\ &= P(A \mid A_1) + P(B \mid A_1) = Q(A) + Q(B). \end{aligned}$$

Här har vi utnyttjat att $A \cap B = \emptyset$ medför att $(A \cap A_1) \cap (B \cap A_1) = A \cap B \cap A_1 = \emptyset$.
Även villkor 3') i Definition 2.2 kan på samma sätt visas vara uppfyllt.

Genom att kombinera satsen med Sats 2.1 på sidan 8 får vi följande resultat.

FÖLJDSATS 2.1

Antag att $P(A_1) > 0$ och låt A och B vara godtyckliga händelser. Då gäller
(i) $P(A^c \mid A_1) = 1 - P(A \mid A_1)$,
(ii) $P(A \cup B \mid A_1) = P(A \mid A_1) + P(B \mid A_1) - P(A \cap B \mid A_1)$.

EXEMPEL 2.16 (*Upprepad betingning*)

Med hjälp av ovanstående kan vi också reda ut vad som gäller vid upprepad betingning. Eftersom $Q_1(A) = P(A \mid A_1)$ är ett sannolikhetsmått kan vi, för A_2 sådan att $Q_1(A_2) > 0$ definiera

$$Q_2(A) := Q_1(A \mid A_2) = \frac{Q_1(A \cap A_2)}{Q_1(A_2)}.$$

Då gäller att

$$\begin{aligned} Q_2(A) = Q_1(A \mid A_2) &= \frac{Q_1(A \cap A_2)}{Q_1(A_2)} = \frac{P(A \cap A_2 \mid A_1)}{P(A_2 \mid A_1)} \\ &= \frac{P(A \cap A_2 \cap A_1)/P(A_1)}{P(A_2 \cap A_1)/P(A_1)} = \frac{P(A \cap A_2 \cap A_1)}{P(A_2 \cap A_1)} \\ &= P(A \mid A_1 \cap A_2). \end{aligned}$$

Observera att $Q_1(A_2) > 0$ medför att $P(A_2 \cap A_1) > 0$ eftersom $Q_1(A_2) = P(A_2 \cap A_1)/P(A_1)$.

Låt oss illustrera detta med ett enkelt exempel. Antag att vi drar tre kort slumpmässigt ur en kortlek och vill veta sannolikheten att alla är spader. Sannolikheten kan beräknas kombinatoriskt som

$$\frac{\binom{13}{3}}{\binom{52}{3}} = \frac{286}{22\,100} = \frac{11}{850}.$$

Ett alternativ är att utnyttja upprepad betingning.
Inför händelserna A_i = ”kort nummer i är en spader”. Vi vill då veta $P(A_1 \cap A_2 \cap A_3)$.
Låt $Q_1(A) := P(A \mid A_1)$ och $Q_2(A) := Q_1(A \mid A_2)$. Då gäller

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1) \cdot P(A_2 \cap A_3 \mid A_1) \\ &= P(A_1) \cdot Q_1(A_2 \cap A_3) \\ &= P(A_1) \cdot Q_1(A_2) \cdot Q_1(A_3 \mid A_2) \\ &= P(A_1) \cdot Q_1(A_2) \cdot Q_2(A_3) \\ &= P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2). \end{aligned}$$

Detta kan man också få fram direkt genom att göra betingningen i en annan ordning.

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1 \cap A_2) \cdot P(A_3 \mid A_1 \cap A_2) \\ &= P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2). \end{aligned}$$

Numeriskt får vi $P(A_1) = 13/52 = 1/4$, $P(A_2 \mid A_1) = 12/51 = 4/17$ och $P(A_3 \mid A_1 \cap A_2) = 11/50$, så att den sökta sannolikheten blir

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{4} \cdot \frac{4}{17} \cdot \frac{11}{50} = \frac{11}{850}$$

som tidigare.

ÖVNING 2.32

Utnyttja upprepad betingning för att beräkna sannolikheten att, när man drar tre kort ur en kortlek, få ett ess, en kung och en dam
a) i nämnd ordning,
b) i godtycklig ordning.

ÖVNING 2.33

En kortlek delas slumpmässigt i fyra lika stora delar (en bridgegiv). Beräkna sannolikheten att alla delarna innehåller ett ess. (L)

ÖVNING 2.34 (*TBC-test, forts från Exempel 2.15*)

Definiera Q som den betingade sannolikheten givet att man testats positivt, dvs $Q(A) = P(A \mid +)$ för händelser A. Beräkna $Q(S)$ respektive $Q(S^c)$, dvs de betingade sannolikheterna för att vara sjuk respektive inte sjuk.

2.7 Blandade problem

201. Antag att Du singlar slant 5 gånger. Vad är sannolikheten att Du
    a) inte får någon klave?
    b) får 1 klave?
    c) får 2 klave?
    (L)

202. I en urna ligger 2 röda och 5 svarta bollar. Man plockar upp tre bollar (utan återläggning). Utfallet att den första bollen är röd, den andra svart och den tredje svart betecknas med rss, och analogt för övriga utfall.
    a) Specificera utfallsrummet Ω.
    b) Ange sannolikheterna för respektive utfall.
    c) Beräkna sannolikheten för händelsen A att man får minst en röd boll.

203. Ett litet Tombola-hjul består av 20 likstora sektorer numrerade 1 till 20. När hjulet snurras stannar det på de olika numren med samma sannolikhet.
    a) Specificera utfallsrummet Ω.
    b) Precisera händelsen A= ”numret är delbart med tre” och bestäm $P(A)$.
    c) Precisera händelsen B= ”numret är delbart med fem” och bestäm $P(B)$.
    d) Precisera händelsen $A \cup B$ och bestäm $P(B)$.
    e) Precisera händelsen $A \cap B$ och bestäm $P(B)$.

204. En tentamen i matematisk statistik ger maximalt 40 poäng (och 0 som minimum) och endast hela poängtal delas ut. Tentan betygsätts enligt följande: F: 0–5p, Fx: 6–15p, E: 16–20, D: 21–25, C: 26–32, B: 33–37, A:

38–40. Vid ordinarie tentamenstillfällen brukar betygen fördela sig ungefär enligt frekvenserna: F: 20%, Fx: 10%, E: 10%, D: 15% C: 20% B: 15%, A: 10%. En student väljs slumpmässigt och studentens tentamensresultet vid det ordinarie tentamenstillfället noteras.

a) Bestäm utfallsrummet Ω.
b) Definiera händelser för vart och ett av betygen och ange sannolikheten för respektive händelse.

205. En duktig bågskytt träffar innersta ringen med sannolikheten 0.9 i varje försök och försöken kan anses vara oberoende. Hon bestämmer sig för att avsluta dagens träning men vill först sätta en skott i innersta ringen. Låt slumpexperimentet ange hur många försök hon behöver.

a) Bestäm utfallsrummet Ω.
b) Beräkna sannolikheter för att hon behöver 1, 2 respektive 3 försök.
c) Vad är sannolikheten att hon behöver minst 4 försök?

206. Antag att händelsen A är en delmängd av händelsen B, dvs $A \subset B$. Visa att detta medför att $P(A) \leq P(B)$. (L)

207. Antag att en häst river olika hinder oberoende av varandra, och att varje hinder rivs med sannolikheten 0.2. Beräkna sannolikheten att hästen river för första gången vid hinder nummer

a) 1,
b) 2,
c) 5.

208. Ett lotteri har tre sorters vinster: förstapris som man vinner med sannolikhet 0.001, andrapris som man vinner med sannolikhet 0.01, och tredjepris som man vinner med sannolikhet 0.1. Man kan inte vinna flera priser på samma lott. Antag att en person vann på lotteriet. Beräkna sannolikheten att personen vann förstapriset.

209. En viss komponent i ett flygplan är mycket viktigt för flygplanets funktion. Vid konstruktion väljer man mellan att ha två ”parallella” sådana där det räcker att en av de två fungerar, eller bara ha en komponent men med högre kvalitet. Antag att komponenter blir trasiga med sannolikheten $p = 0.0001$ oberoende av varandra och att den högkvalitativa komponenten minskar felsannolikheten med en faktor 100, dvs. att felsannolikheten då blir $p_H = 0.000\,001$. Vilken av de två varianterna ger minst risk för att fel uppstår? Vilken relation skall gälla mellan felsannolikheterna p och p_H för att parallellkoppling skall vara bättre respektive sämre?

210. *(Hellre fria än fälla)*
De flesta rättssystem bygger på principen ”hellre fria än fälla”. Likväl är

det oundvikligt att oskyldiga ibland döms. Antag att det bland de åtalade är 70% som verkligen är skyldiga och att sannolikheten att en skyldig döms är 60%, medan sannolikheten att en oskyldig döms är så liten som 0.3%. Hur stor andel av de dömda kommer i så fall att vara oskyldigt dömda?

211. *(Bilen och getterna)*
Ett numera klassiskt sannolikhetsproblem handlar om ett tävlingsspel som faktiskt fanns tidigare på TV i USA och går under benämningen ”bilen och getterna”. Spelet går till som följer. Den tävlande har tre dörrar att välja bland, bakom en finns en bil och bakom de övriga två finns var sin get. Den tävlande vet inte vad som döljer sig bakom respektive dörr men det

Figur 2.6. *Bakom vilken dörr finns det en bil?*

vet tävlingsledaren. Den tävlande kommer att få peka på en av dörrarna. Innan dörren öppnats, och vare sig den tävlande pekade på bilen eller ej, kommer tävlingsledaren att öppna en av de andra två dörrarna bakom vilket det finns en get, och därefter kommer hon att erbjuda den tävlande att ändra sitt val till den kvarvarande oöppnade dörren. (Observera att det alltid är möjligt för tävlingsledaren att, bland de två kvarvarande dörrarna, öppna en dörr med en get bakom.) Den tävlande får sedan öppna den dörr

han valt och får innehållet där bakom. Frågan är nu om man bör acceptera möjligheten att ändra sitt val, om man bör behålla sitt val, eller om det inte spelar någon roll. Beräkna sannolikheten att vinna bilen om man

a) behåller sin ursprungliga dörr,
b) byter dörr,
c) lottar mellan de två dörrarna som finns kvar när tävlingsledaren öppnat en getdörr.

Anm. Detta sannolikhetsproblem väckte stor uppmärksamhet när det togs upp i en populärvetenskaplig amerikanskt tidskrift. Bland annat blev den kvinnliga vetenskapsjournalisten som skrev om problemet och dess lösning fullständigt överöst med klagobrev och rättelser, även från professorer i matematik. Det visade sig dock att hon hade rätt. Anledningen till de starka reaktionerna är att svaret är kontraintuitivt – i alla fall innan man förstått lösningen.

212. *(Bilen och getterna, forts.)*
Ett sätt att tydliggöra att de två dörrarna som finns kvar inte är likvärdiga är att generalisera till att man t.ex. har 100 dörrar med en bil bakom en och getter bakom 99 av dörrarna. Låt säga att den tävlande pekar på dörr 1. Därefter öppnar tävlingsledaren alla övriga dörrar utom en, t.ex. dörr 72. De flesta skulle då vara benägna att byta till dörr 72. Beräkna sannolikheten att vinna bilen i detta modifierade spel för de tre strategierna ovan.

KAPITEL 3

Slumpvariabler

3.1 Definition av slumpvariabel

I föregående kapitel stiftade vi bekantskap med slumpförsök och sannolikheter. Ofta ger slumpförsöken upphov till numeriska resultat. Sådana slumpförsök är av speciellt intresse, dels för att de är vanligt förekommande och dels för att man kan jämföra olika utfall, t.ex. genom att beräkna skillnader. När slumpexperiment ger upphov till numeriska värden talar man om *slumpvariabler*, eller *stokastiska variabler* med ett mer formellt språkbruk.

EXEMPEL 3.1 (*Marknadsundersökning*)

I en marknadsundersökning i ett köpcentrum vill man tillfråga förbipasserande som har småbarn. Antalet individer som passerar innan första småbarnsföräldern kommer kan betraktas som en slumpvariabel. Detta antal är per definition heltalsvärt (och icke-negativt). Man säger då att man har en *diskret* slumpvariabel.

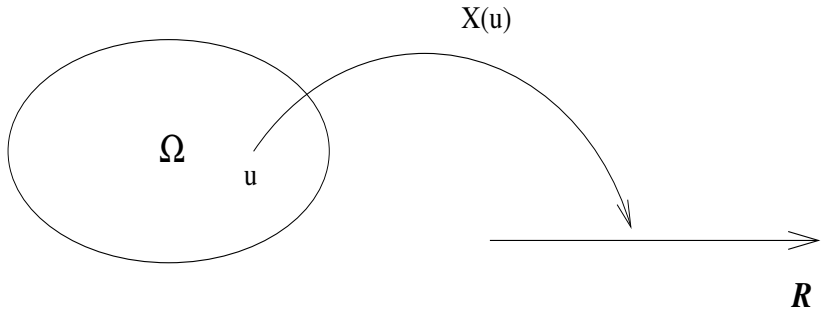
EXEMPEL 3.2 (*Gruvborrning*)

Vid ett gruvfält borras hål på jämnt utspridda platser och man borrar tills dess att halten ädelmetall överstiger 0.001 g/kg berg. Djupet som behöver borras i ett specifikt hål kan betraktas som en slumpvariabel som antar positiva värden. Eftersom i princip alla positiva tal kan antas, och inte bara t.ex. heltal, sägs slumpvariabeln vara *kontinuerlig*.

Vi är nu redo för att definiera begreppet slumpvariabel.

DEFINITION 3.1 (SLUMPVARIABEL)

En *slumpvariabel* (alternativt *stokastisk variabel*) $X(u)$ är en reellvärd funktion definierad på utfallsrummet, $X : \Omega \longmapsto \mathcal{R}$. När slumpförsöket genomförts och ett utfall erhållits så sägs funktionens värde för utfallet vara en *observation* av slumpvariabeln. (Se Figur 3.1 för en illustration av en slumpvariabel och dess utfallsrum.)



Figur 3.1. *Illustration av en slumpvariabel X och tillhörande utfallsrum Ω.*

Slumpvariabler betecknas oftast med stora bokstäver: X, Y, Z och motsvarande observationer med små: x, y, z. När det inte är explicit nödvändigt utelämnas ofta funktionsargumentet och man skriver X i stället för $X(u)$. Vi vill ofta kunna räkna ut sannolikheter av typen $P(X \in A)$ där A kan vara en eller flera punkter, ett intervall eller någon annan delmängd av $\mathcal{R}$. Rent formellt innebär detta att räkna ut sannolikheten för den händelsemängd i utfallsrummet Ω för vilken slumpvariabeln antar värden i A, dvs. $P(\{u \in \Omega; X(u) \in A\})$, men oftast skriver man kort och gott $P(X \in A)$.

I bägge exemplen ovan var slumpvariabeln detsamma som själva utfallet av slumpexperimentet, på samma sätt som utfallsrummet för tärningskast är $\Omega = \{1, 2, \ldots, 6\}$ och slumpvariabeln X som anger antalet prickar är identitetsfunktion, dvs. $X(1) = 1,\ X(2) = 2$ osv. Så behöver dock inte alltid vara fallet vilket visas med följande två exempel.

EXEMPEL 3.3 (*Kast med två tärningar*)

Betrakta försöket att kasta två tärningar och låt vår slumpvariabel Z definieras som summan av antalet prickar på de två tärningarna. Utfallsrummet består av de olika utfallen som tärningskasten kan resultera i:

$\Omega = \{(1,1),(1,2),\ldots,(6,6)\}$ och funktionen Z adderar de två talen (dvs. antal prickar) i utfallet, t.ex. är $Z((3,2)) = 5$. Man inser att flera olika utfall kan ge samma värde på Z – det gäller ju t.ex att $Z((1,4)) = 5$.
I Figur 3.2 illustreras utfallsrummet Ω och funktionen Z som för varje utfall ger ett heltal mellan 2 och 12. Varje utfall har samma sannolikhet 1/36 eftersom tärningskasten är oberoende och varje enkastsutfall har sannolikhet 1/6. Däremot har slumpvariabelns värden inte samma sannolikheter eftersom alla observationsvärden inte svarar mot lika många utfall. Om vi t.ex. vill räkna ut sannolikheten att få minst summan 10, dvs. $P(Z \geq 10)$, kan vi göra detta direkt genom att räkna hur många utfall som ger summan 10, 11 respektive 12: $P(Z \geq 10) = P(Z = 10) + P(Z = 11) + P(Z = 12) = 3/36 + 2/36 + 1/36 = 6/36 = 1/6$. Ett alternativ är att först identifiera vilka utfall u i Ω som uppfyller $Z(u) \geq 10$: $B = \{u \in \Omega; Z(u) \geq 10\} = \{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}$, och eftersom vi har likformig sannolikhetsfördelning på utfallsrummet gäller att sannolikheten för en händelse är lika med antalet utfall i händelsen dividerat med antalet händelser totalt: $P(B) = 6/36 = 1/6$.

Bild saknas

Figur 3.2. *Utfallsrummet Ω för kast med två tärningar och slumpvariabeln Z som anger ögonsumman för respektive utfall. De olika utfallen med given summa $Z = k$ har ringats in.*

EXEMPEL 3.4 (*DNA-sekvensering*)

Tre apparater sekvenserar DNA-kod under en timme vardera. Hur mycket kod (mäts i antal ”base pairs”, bp) som hinner sekvenseras för var och en av apparaterna är slumpmässigt. Ett utfall av slumpexperimentet blir då tre heltalsvärda tal $u = (u_1, u_2, u_3)$ (u_1 är mängd sekvenserad kod av

första apparaten osv.). Om man bara är intresserad av den totala längden sekvenserad kod blir detta en diskret slumpvariabel som för ett utfall $u = (u_1, u_2, u_3)$ definieras som $X = X(u) = u_1 + u_2 + u_3$.

Vi kommer i de följande avsnitten först att gå igenom egenskaper hos diskreta slumpvariabler, därefter kontinuerliga slumpvariabler samt nämna lite om andra typer av slumpvariabler. Efter detta går vi igenom läges- och spridningsmått samt egenskaper hos vanligt förekommande diskreta och kontinuerliga slumpvariabler.

ÖVNING 3.1

Bestäm utfallsrummen Ω i Exempel 3.1 och 3.2 på sidan 39.

ÖVNING 3.2 (*Vinst varje gång*)

I ett vinst-varje-gång-lotteri drar man en kula ur en urna som innehåller 100 kulor numrerade från 1 till 100, och man vinner beloppet som står på den dragna kulan. Definiera en slumpvariabel som anger hur mycket man vinner och dess utfallsrum. Bestäm även sannolikheten att vinsten överstiger 90 kr och uttryck denna storhet med hjälp av sannolikheter, händelser och slumpvariabler.

ÖVNING 3.3

En kvalitetskontrollant bestämmer den exakta vikten på ”100 g” chokladkakor från en viss producent. Kontrollanten skriver upp hur mycket för lite varje chokladkaka väger (om den inte väger för lite skrivs siffran 0 – chokladkakor som väger för mycket utgör inget problem). Specificera utfallsrummet och definiera en slumpvariabel som anger hur mycket för lite en slumpvis vald chokladkaka väger.

3.2 Diskreta slumpvariabler

DEFINITION 3.2 (DISKRET SLUMPVARIABEL)

En slumpvariabel X är *diskret* om den endast kan anta ändligt eller uppräkneligt oändligt antal värden $x_1, x_2, \ldots$.

Antag att vi intresserar oss för en viss diskret slumpvariabel X. Ett sätt att beskriva slumpstrukturen hos slumpvariabeln är med den s.k. sannolikhetsfunktionen.

DEFINITION 3.3 (SANNOLIKHETSFUNKTION)

Sannolikhetsfunktionen $p_X(\cdot)$ för en diskret slumpvariabel X definieras av

$$p_X(x) = P(X = x) = P(X \text{ antar värdet } k), \qquad x = x_1, x_2, \ldots.$$

ANMÄRKNING 3.1

I de allra flesta fallen är observationsvärdena för en diskret slumpvariabel heltal varför vi ofta slarvar lite och skriver $p_X(k)$ *(där* k *syftar på heltal) även när vi pratar allmänt om diskreta slumpvariabler.*

EXEMPEL 3.5 (*Tärningskast*)

Låt X beteckna antalet prickar vid tärningskast med en symmetrisk sexsidig tärning. Då gäller att X har sannolikhetsfunktion $p_X(x) = 1/6$ för $x = 1, \ldots, 6$ och $p_X(x) = 0$ för alla andra x. Sannolikhetsfunktionen illustreras i Figur 3.3. Observera att $p_X(x) = 0$ för alla x utom heltalen 1,2,...,6.

ANMÄRKNING 3.2

När man definierar sannolikhetsfunktioner för diskreta slumpvariabler brukar man oftast bara göra det för värden som har positiv sannolikhet. Implicit betyder det att alla övriga värden har sannolikhet 0. Samma sak gäller för täthetsfunktioner för kontinuerliga slumpvariabler som definieras i Avsnitt 3.4.

Några viktiga egenskaper hos sannolikhetsfunktioner, som följer direkt från Kolmogorovs axiomsystem (2.2) sidan 6, formulerar vi i följande sats.

Bild saknas

Figur 3.3. *Sannolikhetsfunktionen $p_X(k)$ för kast med en symmetrisk sexsidig tärning.*

SATS 3.1

För en diskret slumpvariabel X gäller

1. $0 \leq p_X(k) \leq 1, \forall k$,
2. $\sum_k p_X(k) = 1$,
3. $P(a \leq X \leq b) = \sum_{\{k;a\leq k\leq b\}} p_X(k)$,
4. $P(X \leq a) = \sum_{\{k;k\leq a\}} p_X(k)$,
5. $P(X > a) = \sum_{\{k;k>a\}} p_X(k) = 1 - \sum_{\{k;k\leq a\}} p_X(k) = 1 - P(X \leq a)$.

BEVIS

Låt $A_k = \{u; X(u) = k\}$, dvs. A_k är den mängd i utfallsrummet för vilken slumpvariabeln antar värdet k. Då gäller att $P(X = k) = P(A)$, och enligt Kolmogorovs axiomsystem att $0 \leq P(A) \leq 1$ vilket bevisar det första påståendet.

Vidare gäller att $A_0, A_1, \ldots$ är oförenliga ($X(u)$ antar ju bara ett värde för varje u) och tillsammans utgör de hela utfallsrummet (vi antar att X är heltalsvärd – i annat fall gäller utsagan, men för de observationsvärden X kan anta). Således följer av Kolmogorovs axiomsystem (2.2) att $\sum_k p_X(k) = \sum_k P(A_k) = P(\Omega) = 1$.

Händelsen $\{a \leq X \leq b\}$ är identisk med händelsen $\cup_{\{k;a\leq k\leq b\}} A_k$ eftersom det är just dessa händelser A_k som gör att $a \leq X \leq b$. Eftersom dessa händelser är oförenliga gäller enligt Kolmogorovs axiomsystem alltså $P(a \leq X \leq b) = P(\cup_{\{k;a\leq k\leq b\}} A_k) = \sum_{\{k;a\leq k\leq b\}} p_X(k)$.

De två kvarvarande påståendena visas på liknande sätt och bevisen utelämnas därför.

EXEMPEL 3.6 (*Tändstickor*)

När Swedish Match tillverkar små tändsticksaskar är målsättningen att askarna skall innehålla 50 tändstickor. Dock innehåller inte varje producerad tändsticksask exakt 50 tändstickor. I stället kan antalet tändstickor Z i en slumpvis vald tändsticksask med god approximation beskrivas av en diskret slumpvariabel med sannolikhetsfunktionen: $p_Z(47) = 0.02$, $p_Z(48) = 0.08$, $p_Z(49) = 0.20$, $p_Z(50) = 0.40$, $p_Z(51) = 0.20$, $p_Z(52) = 0.08$, $p_Z(53) = 0.02$ (se Figur 3.4). Av detta följer att $\sum_{z=47}^{53} p_Z(z) = 0.02 + 0.08 + 0.20 + 0.40 + 0.20 + 0.08 + 0.02 = 1$, samt t.ex. att $P(48 \leq Z \leq 52) = \sum_{z=48}^{52} p_Z(z) = 0.08+0.20+0.40+0.20+0.08 = 0.96$.

[Bild saknas]

Figur 3.4. *Sannolikhetsfunktionen $p_Z(k)$ för Z definierad i Exempel 3.6.*

ÖVNING 3.4 (*Barn i svenska hushåll*)

Låt Y beteckna antalet (omyndiga) barn i ett slumpvis valt hushåll i Sverige. Enligt Statistisk årsbok 2005 (utgiven av Statistiska Centralbyrån) fanns det totalt 4 448 746 hushåll i Sverige år 2002, 427 666 av dessa hade 1 barn, 435 612 hade 2 barn och 180 804 hade 3 eller fler barn. Resterande hushåll hade inga barn. Härled $p_Y(y)$ för $y = 0, 1, 2$ och $P(Y \geq 3)$ samt beräkna $P(Y \leq 2)$, $P(1 \leq Y \leq 2)$ och $P(Y > 1)$.

ÖVNING 3.5

Låt X och Y vara slumpvariabler som bägge kan anta värdena 1,2,3,4. Antag att $p_X(1) = 0.2$, $p_X(2) = 0.3$ och $p_X(3) = 0.4$, respektive $p_Y(1) = 0.2c$, $p_Y(2) = 0.3c$, $p_Y(3) = 0.4c$, och $p_Y(4) = 0.5c$.

a) Bestäm $p_X(4)$, samt $P(X \leq 2)$.

b) Bestäm c i sannolikhetsfunktionen för Y, samt beräkna $P(Y > 2)$.

ÖVNING 3.6 (*Tändstickor, forts*)

Beräkna $P(Z \leq 49)$ och $P(Z > 51)$ i Exempel 3.6, sidan 45.

ÖVNING 3.7

En viss sorts apparat kodar en DNA-sekvens av given längd felaktigt (i någon position) med sannolikheten 0.12. Antag att 3 sekvenser av denna längd kodas av olika apparater (av ovan nämnda sort). Låt X vara antalet bland de 3 kodade sekvenserna som kodas helt korrekt. Vilka observationsvärden kan X anta? Härled $p_X(k)$ för dessa k.

3.3 Fördelningsfunktioner

Ganska ofta vill man för slumpvariabler beräkna sannolikheter på formen $P(a < X \leq b)$, $P(X \leq a)$, $P(X > a)$ eller liknande. Om slumpvariabeln är diskret kan man göra detta med hjälp av sannolikhetsfunktionen som beskrivits tidigare. En alternativ metod, som också är mycket användbar för kontinuerliga slumpvariabler, är att införa begreppet *fördelningsfunktion* och uttrycka slumpstrukturen hos slumpvariabeln (inkl. ovan nämnda sannolikheter) med hjälp av denna funktion.

DEFINITION 3.4 (FÖRDELNINGSFUNKTION)

Fördelningsfunktionen $F_X(t)$ för en slumpvariabel X definieras av $F_X(t) = P(X \leq t)$, $-\infty < t < \infty$.

ANMÄRKNING 3.3

Ibland skrivs $F_X(x)$ eller $F_X(s)$ i stället för $F_X(t)$. X:et i index syftar på slumpvariabeln X medan fördelningsfunktionens argumentet kan betecknas hur som helst (på samma sätt som att t.ex. funktionerna $g(t) = t^2$ och $g(x) = x^2$ är desamma).

ANMÄRKNING 3.4

För en diskret slumpvariabel gäller $F_X(t) = \sum_{\{k;k\leq t\}} p_X(k)$, dvs. fördelningsfunktionens värde i punkten x är lika med summan av sannolikhetsfunktionens värden för alla observationsvärden som ej överstiger x. Omvänt kan man beräkna sannolikhetsfunktionen från fördelningsfunktionen: $p_X(k) = F_X(k) - F_X(k-1)$.

EXEMPEL 3.7 (*Tärningskast, fördelningsfunktion*)

Vi har tidigare studerat slumpförsöket att kasta en sexsidig tärning och betraktat slumpvariabeln X som anger antal prickar som visas. Vi fann då att $p_X(k) = 1/6$ för $k = 1, 2, \ldots, 6$. Motsvarande fördelningsfunktion $F_X(t)$ visas i Figur 3.5. Observera att fördelningsfunktionen är konstant mellan heltalen och att den är 0 till vänster om $k = 1$ och 1 till höger om $k = 6$. I heltalen 1,2,...,6 är fördelningsfunktionen det högre av de två värdena i figuren, t.ex. är $F_X(2) = 2/6$ medan $F_X(1.999) = 1/6$

[Bild saknas]

Figur 3.5. *Fördelningsfunktionen $F_X(t)$ för tärningskast med symmetrisk sexsidig tärning.*

EXEMPEL 3.8 (*Tändstickor, fördelningsfunktion*)

I Exempel 3.6, sidan 45, specificerades sannolikhetsfunktionen för antalet tändstickor i en slumpvis vald tändsticksask.
Vi ger några exempel på beräkning av fördelningsfunktionens värden:
$F_Z(48) = P(Z \leq 48) = p_Z(47) + p_Z(48) = 0.02 + 0.08 = 0.1$,
$F_Z(49.4) = P(Z \leq 49.4) = p_Z(47) + p_z(48) + p_Z(49) = 0.02 + 0.08 + 0.20 = 0.30$,
$F_Z(44) = P(Z \leq 44) = 0$ och $F_Z(58.2) = P(Z \leq 58.2) = \sum_{z=47}^{53} p_Z(z) = 0.02 + 0.08 + 0.2 + 0.4 + 0.2 + 0.08 + 0.02 = 1$.
Fördelningsfunktion visas i Figur 3.6. Som synes i figuren gör fördelningsfunktionen hopp i de möjliga observationspunkterna och ligger konstant däremellan (till skillnad från sannolikhetsfunktionen som är noll utom i observationspunkterna).

[Bild saknas]

Figur 3.6. *Fördelningsfunktionen $F_Z(z)$ för Z definierad i Exempel 3.6, sidan 45.*

Några viktiga egenskaper hos fördelningsfunktioner som följer direkt från sannolikhetsaxiomen (2.2) sammanfattas i följande sats.

SATS 3.2 (EGENSKAPER HOS FÖRDELNINGSFUNKTIONER)

Låt $F_X(t)$ vara fördelningsfunktionen för en slumpvariabel X. Då gäller

1. $0 \leq F_X(t) \leq 1, \forall t$,
2. $t \mapsto F_X(t)$ är icke-avtagande och högerkontinuerlig,
3. $\lim_{t \to -\infty} F_X(t) = 0$,
4. $\lim_{t \to \infty} F_X(t) = 1$,
5. $P(a < X \leq b) = F_X(b) - F_X(a)$,

6. $P(X > a) = 1 - F_X(a)$,
7. $P(X < a) = \lim_{h \downarrow 0} F_X(a - h)$.

BEVIS

Vi bevisar bara de två första påståendena. Låt $A_t = \{u; X(u) \leq t\}$. Då följer från Kolmogorovs axiomsystem och definitionen av fördelningsfunktionen att $F_X(t) = P(X \leq t) = P(A_t)$ och att $0 \leq P(A_t) \leq 1$.
Tag nu x och y så att $x < y$. Då gäller att $A_x \subseteq A_y$; alla u för vilka $X(u) \leq x$ har ju även $X(u) \leq y$. Vi kan då dela upp A_y i två disjunkta delar $A_y = A_x \cup A_{(x,y]}$, där $A_{(x,y]} = \{u; x < X(u) \leq y\}$. Enligt Kolmogorovs axiomsystem gäller då

$$F_X(y) = P(A_y) = P(A_x) + P(A_{(x,y]}) \geq P(A_x) = F_X(x),$$

vilket bevisar att $F_X(\cdot)$ är icke-avtagande. Högerkontinuiteten bevisar vi för specialfallet att X är diskret och heltalsvärd – det allmänna fallet är lite krångligare. Fixera t, inte nödvändigtvis heltal. Låt $[t]$ vara heltalsdelen av t (heltalet närmast ”nedanför”, t.ex. är $[4.9] = 4$). Då gäller att $[t+h] = [t]$ om $h > 0$ är tillräckligt liten. Men då gäller $F_X(t + h) = F_X([t + h]) = F_X([t]) = F_X(t)$ vilket visar att $F_X(\cdot)$ är högerkontinuerlig.

ÖVNING 3.8

Låt Y vara en slumpvariabel med fördelningsfunktion

$$F_Y(t) = \begin{cases} 0 & \text{om } t < 0, \\ t^2 & \text{om } 0 \leq t \leq 1, \\ 1 & \text{om } t > 1. \end{cases}$$

a) Rita upp $F_Y(t)$.
b) Beräkna $P(Y \leq 0.5)$.
c) Beräkna $P(0.5 < Y \leq 0.9)$.

ÖVNING 3.9 (*Barn i svenska hushåll, forts*)

Bestäm fördelningsfunktionen för antalet barn i svenska hushåll som beskrevs i Övning 3.4, sidan 45.

ÖVNING 3.10

Bevisa påstående 5 i Sats 3.2 på sidan 48, dvs. att $P(a < X \leq b) = F_X(b) - F_X(a)$. (L)

3.4 Kontinuerliga slumpvariabler

Vi har fram till nu huvudsakligen behandlat diskreta slumpvariabler, dvs. variabler som bara kan anta ett ändligt eller uppräkneligt oändligt antal värden – typiskt alla eller delar av heltalen. I detta avsnitt behandlar vi den andra huvudtypen slumpvariabler, nämligen de kontinuerliga. Dessa slumpvariabler kan typiskt anta alla möjliga värden i något intervall på den reella talaxeln – mer sällsynt förekommer variabler som kan anta värden i flera olika intervall eller krångligare mängder.

EXEMPEL 3.9 (*Barnvikt*)

Låt Y beteckna vikten på ett nyfött barn. Då kan Y anta värden i ett kontinuum av den reella positiva talaxeln (möjligen kan man snäva in intervallet till mellan 0 och 10 kg eller liknande). Den uppmätta vikten Z är däremot inte kontinuerlig – en elektronisk våg brukar begränsa noggrannheten till g, varför vikten blir diskret (och heltalsvärd mätt i gram).

EXEMPEL 3.10 (*Energi i molekyl*)

Låt X beteckna energin i en viss molekyl i ett givet medium, då är X en (icke-negativ) kontinuerlig slumpvariabel.

Liksom i övrig matematik motsvaras summor för diskreta objekt av integraler för kontinuerliga objekt. Vi definierar därför en slumpvariabel som kontinuerlig om följande villkor är uppfyllt.

DEFINITION 3.5 (KONTINUERLIG SLUMPVARIABEL OCH TÄTHETSFUNKTION)

En slumpvariabel X sägs vara *kontinuerlig* om det finns en funktion $f_X(x)$ så att för ”alla” mängder A gäller

$$P(X \in A) = \int_A f_X(t)\, dt.$$

Funktionen $f_X(\cdot)$ kallas för slumpvariabelns *täthetsfunktion*. Se Figur 3.7 för ett exempel.

ANMÄRKNING 3.5

Precis som i fallet med fördelningsfunktioner kan beteckningen för täthetsfunktions argument växla, t.ex. förekommer t, s och x.

ANMÄRKNING 3.6

Ekvationen ovan behöver inte gälla för alla *mängder A utan det räcker att ekvationen är satisfierad för intervall och unioner av intervall (s.k. Borelmängder). Det går att konstruera mer komplicerade mängder men vi fördjupar oss inte i detta.*

[Bild saknas]

Figur 3.7. *Täthetsfunktion för en kontinuerlig slumpvariabel*

I det diskreta fallet fanns ett nära samband mellan fördelningsfunktionen och sannolikhetsfunktionen (Anmärkning 3.4, sidan 47). För det kontinuerliga fallet finns i stället ett samband mellan fördelningsfunktionen och täthetsfunktionen vilket vi formulerar i en sats.

SATS 3.3

För en kontinuerlig slumpvariabel Y med täthetsfunktion $f_Y(\cdot)$ och fördelningsfunktion $F_Y(\cdot)$ gäller

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)\, dt \tag{3.1}$$

och omvänt

$$f_Y(y) = F'_Y(y) = \lim_{h \to 0} \frac{F_Y(y+h) - F_Y(y)}{h},$$

för de punkter där $f_Y(\cdot)$ är kontinuerlig. Det gäller även att

$$\int_{-\infty}^{\infty} f_Y(t)\, dt = 1.$$

BEVIS

Det första resultatet följer enkelt genom att använda definitionen av mängden $A_y = \{-\infty < Y \leq y\} = \{u; -\infty < Y(u) \leq y\}$. Det andra resultatet är en direkt följd av vad som brukar kallas integralkalkylens huvudsats. Slutligen, att $\int_{-\infty}^{\infty} f_Y(t)\, dt = 1$ följer direkt av att $\int_{-\infty}^{\infty} f_Y(t)\, dt = \lim_{y \to \infty} F_Y(y)$ och för en fördelningsfunktion gäller alltid att $\lim_{y \to \infty} F_Y(y) = 1$ (Sats 3.2, sidan 48).

ANMÄRKNING 3.7

Varje funktion $f(x)$ som uppfyller att $f(x) \geq 0$ för alla x och att $\int_{-\infty}^{\infty} f(x)\, dx = 1$, duger som täthetsfunktion till en slumpvariabel.

ANMÄRKNING 3.8

Egentligen har vi i Definition 3.5 definierat en absolutkontinuerlig *slumpvariabel. Det finns nämligen, som vi kommer att visa i Avsnitt 3.6.2, fördelningsfunktioner som är kontinuerliga, men inte kan framställas på formen (3.1). Motsvarande slumpvariabler är så ovanliga att vi använder den kortare benämningen* kontinuerlig *i stället för absolutkontinuerlig utan risk för missförstånd.*

EXEMPEL 3.11 (*Exponentialfördelad slumpvariabel*)

Låt X vara en slumpvariabel med täthetsfunktion $f_X(x) = 2e^{-2x}$, $x > 0$ (implicit är alltså $f_X(x) = 0$ för $x \leq 0$). Denna typ av slumpvariabel, som vi kommer att studera närmare i Avsnitt 3.8.2, sägs vara exponentialfördelad. Täthetsfunktionen illustreras i Figur 3.8. Vi kan till att börja med

förvissa oss om att det är en täthetsfunktion, dvs. att $\int_0^\infty f_X(x)\,dx = 1$. Men, $\int_0^\infty 2e^{-2x}\,dx = [-e^{-2x}]_0^\infty = -0-(-1) = 1$ vilket alltså gör att $f_X(\cdot)$ är en täthetsfunktion.
Låt oss beräkna $P(X \leq 1)$ samt $P(0.3 < X \leq 1.2)$. Vi beräknar först fördelningsfunktionen.

$$F_X(x) = \int_0^x 2e^{-2t}\,dt = [-e^{-2t}]_0^x = -e^{-2x} - (-1) = 1 - e^{-2x}.$$

Från denna får vi $P(X \leq 1) = F_X(1) = 1 - e^{-2} \approx 0.865$, samt $P(0.3 < X \leq 1.2) = F_X(1.2) - F_X(0.3) \approx 0.650$.

[Bild saknas]

Figur 3.8. *Täthetsfunktionen* $f_X(x) = 2e^{-2x},\ x > 0$.

Täthetsfunktionen för en slumpvariabel X är stor i områden där X relativt sett har stor sannolikhet att anta värden och liten i områden där X troligen inte antar sitt värde. I Figur 3.9 illustrerar vi en täthetsfunktion $f_X(x)$ och att $P(a < X < b) = \int_a^b f_X(x)\,dx = F_X(b) - F_X(a)$ är arean under täthetsfunktionen i intervallet (a, b).

Det är dock inte så att täthetsfunktionens värde kan tolkas direkt som en sannolikhet på samma sätt som man kan göra med sannolikhetsfunktionen för en diskret slumpvariabel. I själva verket gäller för en kontinuerlig slumpvariabel X att $P(X = x) = 0$ för all värden på x! Sannolikheten att en kontinuerlig slumpvariabel antar ett bestämt enskilt värde är således 0 – detta följer av att integralen över ett allt mindre intervall går mot 0. Det gäller ju nämligen att $P(X = x) \leq P(x - h \leq X \leq x + h) = \int_{x-h}^{x+h} f_X(t)\,dt$ och denna integral går mot 0 då h går mot 0.

Eftersom sannolikheten att en kontinuerlig slumpvariabel X antar ett enskilt givet värde är 0 gäller det att man inte behöver skilja på ”$<$” och ”$\leq$”

[Täthetsfunktion och $P(a \leq X \leq b)$ inritad]

Figur 3.9. *En täthetsfunktion $f_X(x)$. I figuren är $P(a < X < b) = \int_a^b f_X(x)\,dx$ markerad som arean av det skuggade området.*

samt på ”> och ”≥”, och således gäller t.ex. $P(a \leq X \leq b) = P(a < X < b)$ vilket gör att man inte behöver vara lika noggrann med dessa som i det diskreta fallet där likheten inte gäller.

I vissa sammanhang brukar man för kontinuerliga slumpvariabler prata om intensiteter, speciellt gäller detta för tidpunkter till händelser vilket således rör positiva slumpvariabler. Man menar med detta sannolikheten för att något skall inträffa i ett litet intervall betingat av att det ännu inte inträffat. Mer precist har vi följande definition.

DEFINITION 3.6 (INTENSITET)

Låt X vara en kontinuerlig slumpvariabel med täthetsfunktion $f_X(t)$ och fördelningsfunktion $F_X(t)$. Då definieras *intensiteten* $\lambda_X(t)$ för X som

$$\lambda_X(t) = \frac{f_X(t)}{1 - F_X(t)}.$$

ANMÄRKNING 3.9

Uttrycket i nämnaren härrör från att man betingar på $X > t$.

EXEMPEL 3.12

I Exempel 3.11 studerades en s.k. exponentialfördelad slumpvariabel X med täthetsfunktion $f_X(x) = 2e^{-2x}$, $x \geq 0$, och fördelningsfunktion $F_X(t) = 1 - e^{-2t}$. Från definitionen får vi således att intensiteten blir

$\lambda_X(t) = 2$. Om X symboliserar tiden till en viss händelse (t.ex. att en maskinkomponent går sönder) tolkas detta som att risken att komponenten, som ännu ej gått sönder, går sönder inom en kort tid h är $2h$, och att detta gäller oberoende av hur gammal komponenten är. Det speciella med exponentialfördelningen är f.ö. just att intensiteten är konstant vilket brukar formuleras som att den är minneslös. Vad gäller komponenten betyder det att den inte åldras eftersom den inte får ökad risk att gå sönder ju äldre den är.

Vi har nu stött på såväl diskreta som kontinuerliga slumpvariabler och sett att de behandlas olika matematiskt. Framför allt gäller att det som skrivs som summor för diskreta slumpvariabler i stället blir integraler för kontinuerliga slumpvariabler. Integraler definieras f.ö. som gränsvärden av summor. Av detta kan man ana att en diskret slumpvariabel med många olika möjliga utfall ofta kan approximeras av någon kontinuerlig slumpvariabel. Vi illustrerar detta med ett exempel.

EXEMPEL 3.13

Vid kvalitetskontroll av en datorserver görs mätningar av tid att expediera elektronisk post. Den sanna tidsåtgången kan beskrivas av en (positiv) kontinuerlig slumpvariabel X med någon viss fördelning $f_X(t)$. Mätningarna görs bara med viss noggrannhet och den uppmätta tidsåtgången kan därför beskrivas av en diskret slumpvariabel Y med sannolikhetsfunktion $p_Y(y)$. Om vi t.ex. antar att noggrannheten är mikrosekunder och vi anger tiden i millisekunder, så är täthetsfunktionen för X och sannolikhetsfunktionen för Y relaterade på följande vis:

$$p_Y(y) = \int_{y-0.0005}^{y+0.0005} f_X(t)dt,$$

för $y \geq 0$ angivet med tre decimaler, t.ex. $y = 1.279$. Som exempel följer det av det approximativa sambandet mellan summor och integraler att

$$P(1.350 \leq X \leq 2.484) = \int_{1.350}^{2.484} f_X(t)dt \approx \sum_{y=1.350}^{2.484} p_Y(y).$$

Trots att X är kontinuerlig och Y är diskret är de två slupvariablerna alltså approximativt lika varför det inte spelar så stor roll vilken man betraktar. Tur är väl det, eftersom de flesta saker man gör mätningar av är diskreta approximationer av kontinuerliga ting.

ÖVNING 3.11

Låt $f_X(x) = 2, 2.5 \leq x \leq 3$ (och $f_X(x) = 0$ för övriga x).
a) Bestäm $F_X(x)$ för alla x.
b) Bestäm $P(X \leq 2.7)$.
c) Bestäm $P(X \leq 3.14)$.
d) Bestäm $P(2.1 \leq X < 2.3)$.
e) Bestäm $\lambda_X(2.7)$.

ÖVNING 3.12

Låt Y vara en kontinuerlig slumpvariabel med täthet $f_Y(y) = cy, 0 \leq y \leq 1$.
a) Bestäm c.
b) Rita upp $f_Y(y)$.
c) Beräkna $F_y(y)$ för alla y och rita upp $F_Y(y)$.
d) Beräkna $\lambda_Y(y)$.

ÖVNING 3.13

Vid en produktionsprocess vill man tillverka kolvar med en viss diameter. Man har dock inte perfekt precision utan absolutfelet Y i kolvens diameter kan beskrivas av kontinuerlig slumpvariabel som antar värden mellan 0 och 5 mm och vars täthetsfunktion är omvänt proportionell mot absolutfelet.
a) Bestäm täthetsfunktionen $f_Y(y)$ för *alla* värden på y.
b) Bestäm fördelningsfunktionen $F_Y(y)$ för *alla* värden på y.
c) Beräkna sannolikheten att absolutfelet är högst 2 mm.
d) Beräkna sannolikheten att absolutfelet är mellan 3 och 4 mm.

ÖVNING 3.14

I Övning 3.8 på sidan 49 definierades fördelningsfunktionen $F_Y(y)$ som $F_Y(y) = 0$ för $y < 0$, $F_Y(y) = y^2$ för $0 \leq y \leq 1$ och $F_Y(y) = 1$ för $y > 1$. Bestäm täthetsfunktionen $f_Y(y)$ för alla y.

3.5 Lägesmått och spridningsmått

3.5.1 Lägesmått

En slumpvariabels värde är ju slumpmässigt så man kan därför inte entydigt säga att en den är stor eller liten. Likväl vill man ofta kunna uttala sig om en slumpvariabel är stor eller inte – inte minst vid jämförelse av olika flera slumpvariabler. Eftersom en slumpvariabel kan anta olika värden måste man på något sätt väga in sannolikheten för olika värden för att beskriva om den är stor eller inte. Detta kan göras på olika sätt och man brukar kalla dessa storheter för lägesmått. Det vanligaste lägesmåttet är *väntevärde*.

DEFINITION 3.7 (VÄNTEVÄRDE)

Väntevärdet för en slumpvariabel X betecknas med $E(X)$, μ_X, eller bara μ om det inte kan förväxlas med andra väntevärden, och är ett reellt tal. För en diskret slumpvariabel definieras det som

$$E(X) = \sum_k k \cdot p_X(k), \tag{3.2}$$

och för en kontinuerlig som

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx. \tag{3.3}$$

Väntevärdet för en funktion $g(\cdot)$ av en slumpvariabel X betecknas $E(g(X))$ och definieras analogt som $E(g(X)) = \sum_k g(k)p_X(k)$, respektive $E(g(X)) = \int g(x)f_X(x)dx$.

ANMÄRKNING 3.10

Definitionen gäller endast om summan/integralen är absolut-konvergent, dvs. om $\sum_k |k|p_X(k) < \infty$ *respektive* $\int_{-\infty}^{\infty} |x|\, f_X(x)\, dx < \infty$. *Om summan/integralen är oändlig sägs* X *sakna väntevärde. Fallet att* X *saknar väntevärde kan i själva verket delas upp i olika fall. Om man betraktar de negativa och positiva delarna av utfallen var för sig och tittar på*

$$E(X^-) = \sum_{k<0} k p_X(k) \quad \textit{och}$$
$$E(X^+) = \sum_{k>0} k p_X(k),$$

(eller motsvarande integraler i kontinuerliga fallet) så sägs X sakna väntevärde om ingen av summorna är konvergenta, medan väntevärdet är $+\infty$ om $E(X^-)$ är konvergent och $E(x^+)$ ej är konvergent, respektive att väntevärdet är $-\infty$ om det omvända gäller.

ANMÄRKNING 3.11

Observera att väntevärdet $E(X)$ är ett reellt tal – ett siffervärde (förutsatt att det existerar). Detta till skillnad från slumpvariabeln själv som ju är en funktion på utfallsrummet och som kan anta olika värden.

Den fysikaliska tolkningen av väntevärde är *tyngdpunkt*. Om man ritar upp tätheten/sannolikhetsfunktionen för en slumpvariabel X är $E(X)$ positionen som gör att figuren balanserar (se Figur 3.10 nedan).

[Täthetsfunktion med tyngdpunkt inritad]

Figur 3.10. *Täthetsfunktionen $f_X(x)$ och väntevärdet $E(X)$. Observera att tyngdpunkten är det värde som gör att figuren "väger jämt".*

EXEMPEL 3.14 (*Tändstickor, väntevärde*)

I Exempel 3.6 på sidan 45 definierades en slumpvariabel Z som angav antal tändstickor i en slumpvis vald tändsticksask Sannolikhetsdefinitionen definierades av $p_Z(47) = 0.02$, $p_Z(48) = 0.08$, $p_Z(49) = 0.20$, $p_Z(50) = 0.40$, $p_Z(51) = 0.20$, $p_Z(52) = 0.08$, $p_Z(53) = 0.02$ (se Figur 3.4 på sidan 45). Väntevärdet blir för denna slumpvariabel $E(Z) = \sum_k k p_Z(k) = 47 \cdot 0.02 + 48 \cdot 0.08 + 49 \cdot 0.20 + 50 \cdot 0.40 + 51 \cdot 0.20 + 52 \cdot 0.08 + 53 \cdot 0.02 = 50$. Om vi tänker i termer av tyngdpunkt bör det inte förvåna att tyngdpunkten är just 50 (se Figur 3.4). Man kan i själva verket visa att varje slumpvariabel Y vars fördelning (täthetsfunktion eller sannolikhetsfunktion) är symmetrisk kring ett tal c har $E(Y) = c$ förutsatt att väntevärdet existerar.

Förutom att ett väntevärde kan tolkas som en fördelnings tyngdpunkt finns det också ett nära samband mellan väntevärde och medelvärde som vi nu

antyder. Antag att vi gör ett större antal observationer $y_1, \ldots y_n$ av en diskret slumpvariabel Y med sannolikhetsfunktion $p_Y(k)$. Eftersom vi gör många observationer borde, enligt frekvenstolkningen av sannolikhet (Avsnitt 2.3), den relativa frekvensen $\tilde{p}_k$ av dessa som antar värdet k vara ungefär lika med motsvarande sannolikhet p_k. Medelvärdet $\bar{y} = (y_1 + \ldots y_n)/n$ kan beräknas på den alternativa formen $\bar{y} = \sum_k k\tilde{p}_k$, se Övning 3.19. Från detta får vi alltså $E(Y) = \sum_k k p_Y(k) \approx \sum_k k\tilde{p}_k = \bar{y}$. En tolkning av väntevärde är således att det bör ligga nära medelvärdet av många oberoende observationer.

I senare avsnitt kommer vi att studera funktioner av slumpvariabler och då även intressera oss för väntevärden av dessa. Redan i innevarande avsnitt behöver vi väntevärdet för kvadratfunktionen varför vi redan nu formulerar ett viktigt resultat om väntevärdet av funktioner av slumpvariabler.

SATS 3.4 (VÄNTEVÄRDET AV EN FUNKTION AV EN SLUMPVARIABEL)

Låt X vara en slumpvariabel, $g(\cdot)$ en reell funktion och låt slumpvariabeln Y vara definierad av $Y = g(X)$. Då gäller att

$$E(Y) = E(g(X)) = \begin{cases} \sum_k g(k)\, p_X(k) & \text{om } X \text{ är diskret och} \\ \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx & \text{om } X \text{ är kontinuerlig.} \end{cases}$$

BEVIS

Vi bevisar satsen i det diskreta fallet. Beviset i det kontinuerliga är liknande om än inte identiskt. Kärnan i beviset är att $P(g(X) = j) = \sum_{\{k;g(k)=j\}} P(X = k)$. Vi får därför

$$\begin{aligned} E(Y) &= \sum_j jP(Y = j) = \sum_j jP(g(X) = j) \\ &= \sum_j \sum_{\{k;g(k)=j\}} jP(X = k) = \sum_j \sum_{\{k;g(k)=j\}} g(k)P(X = k) \\ &= \sum_k g(k)P(X = k). \end{aligned}$$

Den sista likheten beror på att dubbelsumman till vänster bara är ett annat sätt att summera över alla möjliga k.

Väntevärdet är det vanligaste lägesmåttet, men även *medianen* är vanligt förekommande.

DEFINITION 3.8 (MEDIAN)

Talet $x_{0.5}$ är *Median* för en slumpvariabel X om den satisfierar $P(X \leq x_{0.5}) \geq 0.5$ och $P(X \geq x_{0.5}) \geq 0.5$.

ANMÄRKNING 3.12

Om X är kontinuerlig är en ekvivalent definition att $x_{0.5}$ löser $F_X(x) = 0.5$.

[Bild saknas]

Figur 3.11. *Vänstra figuren visar en fördelningsfunktion och dess median, högra figuren visar motsvarande täthetsfunktion och dess median.*

EXEMPEL 3.15 (*Median av diskret slumpvariabel*)

Betrakta slumpvariabeln Y som antar värdena 1, 2, 3, 4 och 5 med samma sannolikhet 0.2. Det betyder att $p_X(k) = 0.2$ för $k = 1, \ldots 5$. Medianen blir i detta fall $x_{0.5} = 3$.

EXEMPEL 3.16 (*Median av kontinuerlig slumpvariabel*)

Antag att en kontinuerlig slumpvariabel Y har täthetsfunktionen $f_Y(t) = t/2$, $0 \leq t \leq 2$. Det medför att fördelningsfunktionen blir $F_Y(t) = \int_0^t f_Y(s)ds = \int_0^t t/2ds = t^2/4$ $(0 \leq t \leq 2)$. Om vi löser ekvationen $t^2/4 = 0.5$ erhåller man lösningen $y_{0.5} = \sqrt{2} \approx 1.414$ som alltså är medianen.

Det är värt att notera att medianen inte behöver vara unik om $F_X(t)$ är konstant lika med 0.5 i något intervall vilket följande exempel visar.

EXEMPEL 3.17

Låt Z anta värdet 0 med sannolikhet 0.5 och 1 med sannolikhet 0.5 ($p_Z(0) = p_Z(1) = 0.5$). Då är alla värden i intervallet $(0,\ 1)$ median.

När ska man använda väntevärde respektive median?
På denna fråga finns inget entydigt svar. Inte så sällan sammanfaller de (och då spelar det ju ingen roll), t.ex. för symmetriska täthets- eller sannolikhetsfunktioner – median och väntevärde är då bägge lika med symmetripunkten. Annars är det vanligast att man använder väntevärdet, en anledning till detta är nog dess koppling till medelvärdet som är centralt inom statistiken som vi ska se senare i boken. En annan anledning är att väntevärdet har flera beräkningsmässiga fördelar gentemot median (mer om dessa i senare avsnitt). Ibland är det dock lämpligare att använda medianen som lägesmått. T.ex. påpekades det tidigare att väntevärdet kan vara oändligt eller inte väldefinierat, och då är medianen det enda valet. Men även i fall där slumpvariabler har väntevärden brukar man använda medianen om fördelningen har s.k. *tunga svansar*. Med detta menas att fördelningen/slumpvariabeln kan anta mycket stora (och/eller väldigt stora negativa) värden. Dessa värden påverkar väntevärdet i mycket hög grad medan värdena mer i mitten (som har lejonparten av sannolikhetsmassan) får föga inflytande på väntevärdet vilket kanske inte är önskvärt. Liknande sker inom statistiken som vi ska se senare. När man t.ex. studerar löner är det vanligare att man redovisar medianlön än medellön. Orsaken till detta är att några få människors höga löner drar upp medellönen till ett belopp som ganska få tjänar, så medellönen avspeglar inte en typisk lön – därför redovisar man oftare medianlönen som bättre avspeglar vad man vill (referens till inferenskapitel??).

Medianen är för övrigt ett specialfall av de mer allmänna begreppen kvantil och percentil som vi passar på att definiera.

DEFINITION 3.9 (KVANTIL, KVARTIL OCH PERCENTIL)

För $0 < \alpha < 1$ definieras α-*kvantilen* x_α till en slumpvariabel X som lösningen $x = x_\alpha$ till ekvationen $F_X(x) = 1-\alpha$. Valen av $\alpha = 0.75$ respektive $\alpha = 0.25$ har liksom medianen ($\alpha = 0.5$) fått egna namn. Lösningen $x_{0.75}$ till $F_X(x) = 1 - 0.75$ kallas *första kvartilen* (eller *nedre kvartilen*) och betecknas med Q_1 och lösningen $x_{0.25}$ till $F_X(x) = 1 - 0.25$ kallas för *tredje kvartilen* (eller *övre kvartilen*) och betecknas Q_3. *Andra kvartilen* är detsamma som medianen.
Percentil definieras i termer av procent, men här syftar man på chansen

att vara mindre än värdet ifråga. En slumpvariabels $100 \cdot r$-*percentil* definieras som lösningen till ekvationen $F_X(x) = r$. Det betyder t.ex. att en slumpvariabels 90%-percentil är densamma som dess 0.1-kvantil.

[Bild saknas]

Figur 3.12. *Figuren visar en täthetsfunktion och dess första och tredje kvartil samt median.*

ANMÄRKNING 3.13

Det bör påpekas att t.ex. 0.1-kvantilen är det tal som har sannolikhetsmassan 0.9 till vänster om sig och 0.1 till höger. Mer naturligt hade kanske varit att använda de omvända förhållandena. Orsaken beror på att man inom statistikteorin ofta använder kvantiler för att det skall vara små sannolikheter att få större *värden. Mer om detta i kommande kapitel. Percentil däremot, används oftast inte på detta sätt. När begreppen kvantil och percentil dyker upp bör man vara uppmärksam på vad som menas – tyvärr är det så att begreppen definieras olika av olika författare.*

EXEMPEL 3.18 (*Kvartiler*)

Betrakta en kontinuerlig slumpvariabel Y med fördelningsfunktion

$$F_Y(y) = \begin{cases} 1 - e^{-y/6} & \text{för } 0 < y < \infty, \\ 0 & \text{för } y \leq 0. \end{cases}$$

Detta är en fördelningsfunktion eftersom den är icke-avtagande, den går mot 0 då $y \to -\infty$ (i själva verket är den lika med 0 för negativa y) och den går mot 1 då $y \to \infty$.

För att beräkna medianen löser vi alltså $F_Y(y) = 0.5$, dvs $1 - e^{-y/6} = 0.5$. Detta betyder att $e^{-y/6} = 0.5$, dvs. $-y/6 = \ln(0.5)$. Om vi förenklar detta får vi alltså att medianen ges av $y_{0.5} = 6\ln 2 \approx 4.16$.
På liknande sätt får man att första kvartilen löser $F_Y(y) = 0.25$ med lösning $y_{0.75} = 6\ln(1/0.75) \approx 1.72$ samt att tredje kvartilen ges av lösningen till $F_Y(y) = 0.75$ med lösning $y_{0.25} = 6\ln(1/0.25) \approx 8.32$.
Slutligen blir 0.1-kvantilen (tillika 90%-percentilen) lösningen till $F_Y(y) = 0.9$ vilken ges av $y_{0.1} = 6\ln(1/0.1) \approx 13.8$.

3.5.2 Spridningsmått

I förra delavsnittet beskrevs en slumpvariabels kanske viktigaste enskilda sammanfattande storhet, nämligen lägesmåttet som på ett eller annat sätt anger ett enskild tal som beskriver hur stor en slumpvariabel ”typiskt” är. Eftersom det är en slumpvariabel antar ju inte variabeln detta värde alltid. En relevant fråga är därför hur mycket slump variabeln är behäftad med. Ligger den t.ex. alltid ganska nära sitt väntevärde eller avviker den ofta mycket från väntevärdet? När man vill beskriva detta använder man något *spridningsmått*. Vi börjar med att definiera *varians*.

DEFINITION 3.10 (VARIANS)

Variansen $V(X)$ för en slumpvariabel X med väntevärde μ definieras som

$$V(X) = E((X-\mu)^2) = \begin{cases} \sum_k (k-\mu)^2 p_X(k), \\ \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx, \end{cases}$$

om detta väntevärde är ändligt. Den senare likheten gäller om X är diskret respektive om X är kontinuerlig och följer av Sats 3.4, sidan 59.

EXEMPEL 3.19 (*Tändstickor, varians*)

I Exempel 3.6 på sidan 45 beräknades väntevärdet för antal tändstickor i en tändsticksask till $\mu = 50$. Variansen för antalet tändstickor blir $V(X) = \sum_{k=47}^{53}(k-50)^2 p_X(k) = (-3)^2 p_X(47) + (-2)^2 p_X(48) + (-1)^2 p_X(49) + (0)^2 p_X(50) + 1^2 p_X(51) + 2^2 p_X(52) + 3^2 p_X(53) = 1.4$.

EXEMPEL 3.20 (*Variansexempel*)

Antag att Y är diskret med sannolikhetsfunktion $p_Y(0) = 0.1$, $p_Y(1) = 0.25$, $p_Y(2) = 0.63$, samt $p_Y(3) = 0.02$. För att räkna ut variansen av Y måste vi först beräkna väntevärdet $E(Y) = \sum_k k p_Y(k) = 0p_Y(0) + 1p_Y(1) + 2p_Y(2) + 3p_Y(3) = 1.57$. Variansen blir därför $V(Y) = \sum_k (k - \mu)^2 p_Y(k) = (-1.57)^2 p_Y(0) + (-0.57)^2 p_Y(1) + 0.43^2 p_Y(2) + 1.43^2 p_Y(3) = 0.4851$.

Det är betydligt lättare att räkna ut variansen i föregående exempel om man använder följande matnyttiga resultat.

SATS 3.5 (VARIANSBERÄKNING)

För en slumpvariabel X med väntevärde μ_X gäller

$$\begin{aligned} V(X) &= E(X^2) - \mu_X^2 \\ &= E(X^2) - (E(X))^2 \\ &= \begin{cases} \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - \mu_X^2, & \text{om } X \text{ är kontinuerlig} \\ \sum_k k^2 p_X(k) - \mu_X^2, & \text{om } X \text{ är diskret.} \end{cases} \end{aligned}$$

BEVIS

Vi visar satsen i det kontinuerliga fallet; det diskreta visas analogt. Enligt definitionen av varians gäller

$$\begin{aligned} V(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx = \int_{-\infty}^{\infty} (x^2 - 2\mu x - \mu^2) f_X(x)\,dx \\ &= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - 2\mu \int_{-\infty}^{\infty} x f_X(x)\,dx + \mu^2 \int_{-\infty}^{\infty} f_X(x)\,dx. \end{aligned}$$

Den andra integralen blir μ per definition, så hela termen blir $-2\mu^2$, och den sista integralen blir 1 eftersom $f_X(x)$ är en täthetsfunktion. Detta tillsammans bevisar satsens påstående.

EXEMPEL 3.21 (*Variansexempel, alternativ beräkning*)

I föregående exempel (Exempel 3.20) beräknades variansen enligt definitionen. Vi gör nu detsamma med hjälp av räkneregeln från Sats 3.5. Det

gäller att $\sum_k k^2 p_Y(k) = 0^2 \cdot 0.1 + 1^2 \cdot 0.25 + 2^2 \cdot 0.63 + 3^2 \cdot 0.02 = 2.95$. Väntevärdet beräknades tidigare till $E(X) = 1.57$. Räkneregeln ger därför att $V(X) = 2.95 - 1.57^2 = 0.4851$.

Ett bättre och väsentligt mer använt spridningsmått än varians är *standardavvikelse*. Skälet att vi trots detta definierat varians först är att standardavvikelsen definieras med hjälp av variansen; man måste alltså beräkna variansen för att få fram standardavvikelsen.

DEFINITION 3.11 (STANDARDAVVIKELSE)

Standardavvikelsen $D(X)$ för en slumpvariabel X definieras som

$$D(X) := \sqrt{V(X)}.$$

Standardavvikelsen betecknas ofta med σ_X eller σ.

Observera att standardavvikelsen, liksom väntevärde och varians är reella tal till skillnad från slumpvariabeln som antar olika värden vid skilda observationer (den är ju en funktion på utfallsrummet). Värt att påpeka är även att standardavvikelsen, liksom väntevärdet men till skillnad från variansen, har samma enhet som variabeln själv. Om t.ex. slumpvariabeln är en storhet i m är väntevärde och standardavvikelse också det medan variansens enhet blir m^2. Detta kan sägas vara huvudanledningen till att standardavvikelsen är ett bättre spridningsmått än variansen.

EXEMPEL 3.22 (*Variansexempel, standardavvikelse*)

I Exempel 3.20 och exempel 3.21 beräknades variansen till $V(X) = 0.4851$ vilket gör att standardavvikelsen blir $D(X) = \sqrt{0.4851} \approx 0.6965$.

Det finns inget entydigt enkelt sätt att relatera standardavvikelsen till hur mycket en slumpvariabel typiskt kan avvika från sitt väntevärde. Vill man ändå ha en sådan vägledning kan man inte så sällan använda principen att slumpvariabeln ofta ligger inom en standardavvikelse från sitt väntevärde men att större avvikelser förekommer, medan avvikelser på mer än två eller tre standardavvikelser från väntevärdet är sällsynta respektive mycket sällsynta.

Ibland är det av intresse att mäta spridning i relation till slumpvariabelns storhet. Man får på så sätt ett dimensionslöst mått. Ett dylikt mått är emeller-

tid bara meningsfullt för positiva slumpvariabler varför definitionen endast gäller för dessa.

DEFINITION 3.12 (VARIATIONSKOEFFICIENT)

Variationskoefficienten $R(X)$ för en positiv slumpvariabel X definieras som

$$R(X) := \frac{D(X)}{E(X)}.$$

En nackdel med variationskoefficienten är förstås att den bara definieras för positiva slumpvariabler. En fördel är dock att den som sagt är dimensionslös. Med detta menas att variationskoefficienten inte påverkas av vilken enhet slumpvariabeln mäts i utan den ”mäter” standardavvikelsen i termer av väntevärdet. Om man t.ex. betraktar en slumpvariabel som anger en vikt i kg blir variationskoefficienten densamma om man övergår till enheten g, medan standardavvikelsen multipliceras med 1000 när man övergår till g.

Som redan nämnts är standardavvikelsen det vanligaste spridningsmåttet.

EXEMPEL 3.23 (*Variansexempel, variationskoefficient*)

I Exempel 3.20 och 3.22 beräknades väntevärde respektive standardavvikelse för en slumpvariabel X till $E(X) = 1.57$ respektive $D(X) = 0.6965$. Vi får därmed variationskoefficienten till $R(X) = 0.6965/1.57 \approx 0.444$.

3.5.3 Olikheter

Vi avslutar detta avsnitt med två olikheter som knyter samman sannolikhetsuttryck med väntevärde respektive standardavvikelse. Poängen med dessa är att man kan uttala sig konservativt om sannolikheter, dvs. ge övre gränser för sannolikheter, när man bara känner till väntevärde/standardavvikelsen för slumpvariabeln men inte hela dess slumpstruktur.

SATS 3.6 (MARKOVS OLIKHET)

Låt X vara en positiv slumpvariabel ($X \geq 0$) med ändligt väntevärde $E(X)$. Då gäller, för varje $a > 0$, att

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

ANMÄRKNING 3.14

Markovs olikhet är uppkallad efter den ryske matematikern Andrei Markov (1856-1922).

BEVIS

Vi visar olikheten i det kontinuerliga fallet (det diskreta fallet visas analogt). Från definitionen av väntevärde följer

$$E(X) = \int_0^\infty x\, f_X(x)\, dx = \int_0^a x\, f_X(x)\, dx + \int_a^\infty x\, f_X(x)\, dx$$
$$\geq 0 + a \int_a^\infty f_X(x)\, dx = a\, P(X \geq a).$$

Detta bevisar satsen.

SATS 3.7 (CHEBYSHEVS OLIKHET)

Låt Z vara en slumpvariabel med ändligt väntevärde μ och standardavvikelse σ. Då gäller, för varje $a > 0$, att

$$P(|Z - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

ANMÄRKNING 3.15

Chebyshevs olikhet är uppkallad efter den ryske matematikern Pafnutij Tjebysjov (1821-1894). Olikheten använder den engelska stavningen vilken ibland även används för personen, även Tjebyshov och andra varianter förekommer.

BEVIS

Vi antar även i detta bevis att slumpvariabeln är kontinuerlig. Från definitionen av varians har vi att

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty} (z-\mu)^2 f_Z(z)\,dz \\ &= \int_{|z-\mu|<a} (z-\mu)^2 f_Z(z)\,dz + \int_{|z-\mu|\geq a} (z-\mu)^2 f_Z(z)\,dz \\ &\geq 0 + a^2 \int_{|z-\mu|\geq a} f_Z(z)\,dz \\ &= a^2\, P(|Z-\mu| \geq a).\end{aligned}$$

Genom att dividera höger- och vänsterled med a^2 visas satsens påstående.

Dessa lättbevisade och till synes enkla satser är i själva verket mycket kraftfulla redskap i många olika sammanhang. Vi kommer senare att använda Chebyshevs olikhet för att visa stora talens lag (Avsnitt 3.12) men visar redan nu ett exempel på hur Markovs olikhet kan användas.

EXEMPEL 3.24 (*Markovs olikhet*)

Antag att man endast känner till väntevärdet för en positiv slumpvariabel och att detta är $\mu = 7$. Om vi vill få en uppskattning av 0.1-kvantilen kan vi använda Markovs olikhet. Det gäller nämligen att $P(X \geq 70) \leq 0.1$. Således gäller att $x_{0.1} \leq 70$. Man kan alltså ge övre begränsningar på kvantiler trots att man bara känner väntevärdet för en stokastisk variabel.

ÖVNING 3.15

Betrakta den diskreta slumpvariabeln X med sannolikhetsfunktion $p_X(1) = 0.2$, $p_X(2) = 0.1$, $p_X(3) = 0.3$, $p_X(4) = 0.1$,$p_X(5) = 0.3$. Beräkna väntevärdet $E(X)$ och standardavvikelsen $D(X)$.

ÖVNING 3.16

Betrakta den kontinuerliga slumpvariabeln X med täthetsfunktion $f_X(x) = 2$ för $2.5 \leq x \leq 3$ (och $f_X(x) = 0$ för övriga x).
a) Beräkna väntevärdet $E(X)$.

b) Beräkna medianen $x_{0.5}$.
c) Beräkna variansen $V(X)$.
d) Beräkna standardavvikelsen $D(X)$.
e) Beräkna variationskoefficienten $R(X)$

ÖVNING 3.17

Betrakta den kontinuerliga slumpvariabeln Y med täthetsfunktion $f_Y(y) = 2y$, $0 \leq y \leq 1$ (som definierades i Övning 3.12, sidan 56).
a) Beräkna väntevärdet $E(Y)$.
b) Beräkna medianen $y_{0.5}$.
c) Beräkna variansen $V(Y)$.
d) Beräkna standardavvikelsen $D(Y)$.

ÖVNING 3.18

Man kan visa att för positiva kontinuerliga slumpvariabler gäller även $E(X) = \int_0^\infty (1 - F_X(x))\,dx$ (observera att integralen går från 0 - även om slumpvariabeln t.ex. bara kan anta värden större än 5). Visa att detta gäller för Y från föregående uppgift (som definierades i Övning 3.12).

ÖVNING 3.19

Visa att $(y_1 + \ldots y_n)/n = \sum_k k\tilde{p}_k$ där $\tilde{p}_k = \#\{y_i; y_i = k,\ k = 1, \ldots n\}/n$, dvs. antalet observationer som antar värdet k dividerat med antal observationer.

ÖVNING 3.20

Låt X vara en kontinuerlig slumpvariabel med symmetrisk täthetsfunktion, dvs. antag att det finns ett tal a sådant att $f_X(a - y) = f_X(a + y)$ för alla y. Visa att i så fall är a såväl median som väntevärde (om detta existerar).

ÖVNING 3.21

Ge en övre begränsning på 0.05-kvantilen av en positiv slumpvariabel X med väntevärde 10.

ÖVNING 3.22

En slumpvariabel Y har väntevärde 0 och varians $\sigma^2 = 4$. Ge en uppskattning av sannolikheten absolutbeloppet $|Y|$ ej överstiger 4.

3.6 Blandade och singulära fördelningar

Vi har studerat diskreta slumpvariabler i Avsnitt 3.2, och kontinuerliga slumpvariabler i Avsnitt 3.4, men dessa två huvudtyper täcker inte in alla förekommande typer av slumpvariabler. Till exempel kan man tänka sig slumpvariabler som är blandningar av diskreta och kontinuerliga variabler.

3.6.1 Blandning av diskret och kontinuerlig fördelning

EXEMPEL 3.25 (*Väntetid vid trafikljus*)

I en gatukorsning med trafikljus är de röda och gröna perioderna lika långa, båda a sekunder. Vi bortser från den tid det tar för trafikljusen att slå om.

Betrakta en bil som anländer till korsningen vid en slumpmässig tidpunkt och låt $X =$ ”väntetiden till grönt ljus”.

Då gäller uppenbarligen att $P(X = 0) = 1/2$ eftersom trafikljuset visar grönt hälften av tiden. Om bilen anländer under en röd period kommer däremot X att vara en kontinuerlig slumpvariabel med värde i intervallet $(0, a)$.

Låt $Y := X \mid (X > 0)$, dvs. väntetiden givet att bilen får vänta. Om bilen anländer vid en slumpmässigt vald tidpunkt i ett rött intervall gäller uppenbarligen att $P(Y \leq t) = t/a$, så att $Y \sim \mathrm{Re}(0, a)$.

Vi kan beskriva X som

$$X = \begin{cases} 0 & \text{med sannolikhet } \frac{1}{2}, \\ Y & \text{med sannolikhet } \frac{1}{2}, \end{cases} \tag{3.4}$$

där $Y \sim \mathrm{Re}(0, a)$, med $E(Y) = a/2$ och $E(Y^2) = a^2/3$.

Med hjälp av framställningen (3.4) och betingning kan vi beräkna

$$E(X) = 0 \cdot \frac{1}{2} + E(Y) \cdot \frac{1}{2} = \frac{a}{2} \cdot \frac{1}{2} = \frac{a}{4}$$

och

$$E(X^2) = 0^2 \cdot \frac{1}{2} + E(Y^2) \cdot \frac{1}{2} = \frac{a^2}{3} \cdot \frac{1}{2} = \frac{a^2}{6},$$

så att

$$V(X) = \frac{a^2}{6} - \left(\frac{a}{4}\right)^2 = \frac{a^2}{6} - \frac{a^2}{16} = \frac{5a^2}{48}.$$

Detta är ett exempel på en blandad fördelning, som varken är diskret eller kontinuerlig. Allmänt kan vi skriva en blandning som

$$X = \begin{cases} Y & \text{med sannolikhet } p, \\ Z & \text{med sannolikhet } 1 - p, \end{cases}$$

där Y och Z har fördelningsfunktioner $F_Y(t)$ respektive $F_Z(t)$. Då gäller

$$\begin{aligned} F_X(t) &= P(X \leq t) = p \cdot P(Y \leq t) + (1 - p) \cdot P(Z \leq t) \\ &= p \cdot F_Y(t) + (1 - p) \cdot F_Z(t). \end{aligned}$$

Om både Y och Z är diskreta blir även X diskret och om både Y och Z är kontinuerliga så blir även X kontinuerlig.

Det intressanta fallet är, som i trafikljusexemplet 3.25, då Y är diskret och Z kontinuerlig (eller tvärtom). Då blir X varken diskret eller kontinuerlig, utan en blandning av dessa.

Väntevärde och varians för X kan beräknas med hjälp av betingning, som i exemplet.

$$\begin{aligned} E(X) &= p \cdot E(Y) + (1 - p) \cdot E(Z), \\ E(X^2) &= p \cdot E(Y^2) + (1 - p) \cdot E(Z^2), \\ V(X) &= E(X^2) - (E(X))^2. \end{aligned}$$

3.6.2 * Singulära fördelningar

De fördelningar som vi kallat kontinuerliga i Avsnitt 3.4 bör egentligen kallas *absolutkontinuerliga*. Det finns nämligen fördelningsfunktioner $F(t)$ som är kontinuerliga, men inte kan skrivas på formen

$$F(t) = \int_{-\infty}^{t} f(x)\,dx \tag{3.5}$$

för någon täthetsfunktion $f(x)$.
Motsvarande fördelningar är alltså inte (absolut)kontinuerliga, men inte heller diskreta eftersom F(t) är kontinuerlig och alltså saknar hopp. De kan inte heller vara blandningar av diskreta och kontinuerliga fördelningar.

Vi påminner om att det för absolutkontinuerliga fördelningar gäller att täthetsfunktionen f kan skrivas som derivatan av fördelningsfunktionen F, $f(t) = F'(t)$ i alla punkter där $f(t)$ är kontinuerlig.

DEFINITION 3.13 (SINGULÄR FÖRDELNING)

En slumpvariabel sägs ha en *singulär* fördelning om fördelningsfunktionen $F(t)$ är kontinuerlig men inte kan uttryckas på formen (3.5) för någon täthetsfunktion f.

Lyckligtvis är singulära fördelningar sällan förekommande och läsaren kommer knappast att stöta på någon i något praktiskt sammanhang. För fullständighets skull visar vi ändå med ett exempel att de faktiskt existerar.

EXEMPEL 3.26 (*Cantorfördelning*)

Vi konstruerar den singulära Cantorfördelningen på intervallet $(0, 1)$ som ett gränsvärde.
Låt $U_1, U_2, \ldots$ vara en oändlig följd oberoende likafördelade slumpvariabler med $P(U_i = 0) = P(U_i = 2) = 1/2$.
Vi genomför konstruktionen stegvis.

Steg 1: Låt $X_1 := U_1/3$.
Då gäller att

$$X_1 = \begin{cases} 0 & \text{med sannolikhet } \frac{1}{2}, \\ \frac{2}{3} & \text{med sannolikhet } \frac{1}{2} \end{cases}$$

och

$$F_{X_1}(t) = \begin{cases} 0 & \text{då } t < 0, \\ \frac{1}{2} & \text{då } 0 \leq t < \frac{2}{3}, \\ 1 & \text{då } t \geq \frac{2}{3}, \end{cases}$$

med $F'_{X_1}(t) = 0$ utom i $t = 0$ och $t = 2/3$ där den är odefinierad, så att $F'_{X_1}(t) = 0$ på hela intervallet $(0, 1)$ utom i två punkter, dvs. på delintervall av sammanlagd längd 1.
Det största hoppet hos $F_{X_1}(t)$ är 1/2, för $t = 0$ och $t = 2/3$.

Steg 2: Låt $X_2 := U_1/3 + U_2/3^2 = X_1 + U_2/3^2$.
Då gäller att X_2 antar värdena 0, 2/9, 2/3 och 8/9 med sannolikhet 1/4 vardera. Vidare är $F'_{X_2}(t) = 0$ utom i dessa punkter, dvs. på delintervall av sammanlagd längd 1.
Det största hoppet hos $F_{X_2}(t)$ är $1/4 = (1/2)^2$, för $t = 0$, $t = 2/9$, $t = 2/3$ och $t = 8/9$.

Steg n: Låt $X_n = U_1/3 + U_2/3^2 + \cdots + U_n/3^n = X_{n-1} + U_n/3^n$.
Då gäller att X_n kan anta 2^n olika värden, alla med sannolikhet $(1/2)^n$. Detta innebär att $F'_{X_n}(t) = 0$ utom i dessa hoppunkter, dvs. på intervall av sammanlagd längd 1 och att samtliga hopp är $(1/2)^n$.
Vi får alltså under hela konstruktionen en fördelning där fördelningsfunktionen är konstant utom i hoppunkterna, dvs. derivatan är 0 i delintervall av sammanlagd längd 1 och det största hoppet halveras för varje steg och går mot 0 då $n \to \infty$.

Låt X_∞ beteckna gränsvariabeln då $n \to \infty$. Dess fördelningsfunktion, $F_{X_\infty}(t)$, är uppenbarligen kontinuerlig för alla $0 \leq t \leq 1$, eftersom hoppens storlek går mot 0. Vidare har den $F'_{X_\infty}(t) = 0$ på intervall av sammanlagd längd 1, så att det inte kan finnas någon täthetsfunktion.
En Cantorfördelad slumpvariabel kan alltså skrivas

$$X = \sum_{n=1}^{\infty} \frac{U_n}{3^n},$$

där U_n är oberoende och 0 eller 2 med samma sannolikhet.
Observera att alla tal $0 \leq x \leq 1$ kan skrivas decimalt i bas 3 som

$$x = \sum_{n=1}^{\infty} \frac{c_n}{3^n},$$

där c_n är 0, 1 eller 2.

Eftersom $E(U_n) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 2 = 1$ kan vi beräkna

$$E(X) = E\left(\sum_{n=1}^{\infty} \frac{U_n}{3^n}\right) = \sum_{n=1}^{\infty} \frac{E(U_n)}{3^n} = \sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1/3}{1-1/3} = \frac{1}{2},$$

vilket är rimligt av symmetriskäl.
Vidare är $V(U_n) = E((U_n - 1)^2) = \frac{1}{2} \cdot (-1)^2 + \frac{1}{2} \cdot 1^2 = 1$, så att

$$V(X) = V\left(\sum_{n=1}^{\infty} \frac{U_n}{3^n}\right) = \sum_{n=1}^{\infty} \frac{V(U_n)}{(3^n)^2} = \sum_{n=1}^{\infty} \frac{1}{9^n} = \frac{1/9}{1-1/9} = \frac{1}{8}.$$

Det finns alltså tre huvudtyper av slumpvariabler: diskreta, (absolut)-kontinuerliga och singulära. Dessutom kan man tänka sig blandningar av dessa, så att den mest allmänna varianten är

$$X = \begin{cases} Y & \text{med sannolikhet } p_1, \\ Z & \text{med sannolikhet } p_2, \\ W & \text{med sannolikhet } 1 - p_1 - p_2, \end{cases}$$

där Y är diskret, Z är (absolut)kontinuerlig och W är singulär.

ÖVNING 3.23

I ett visst betjäningssystem finns en betjänare och antalet kunder i systemet, X, har fördelning

$$P(X = k) = \left(\frac{2}{3}\right)^k \cdot \frac{1}{3}, \quad \text{för } k = 0, 1, 2, \ldots.$$

Kundernas betjäningstider är oberoende av varandra och exponentialfördelade med väntevärde 2 minuter.
Betrakta en kund som anländer vid en slumpmässig tidpunkt och låt W beteckna hennes kötid.

a) Uttryck denna storhet som en blandning av en diskret och en kontinuerlig variabel.
b) Beräkna väntevärdet för variabeln.(L)

ÖVNING 3.24

Låt X vara en blandning av Y, med sannolikhet p, och Z, med sannolikhet $1 - p$.
Visa att $V(X) = p \cdot V(Y) + (1 - p) \cdot V(Z) + p(1 - p) \cdot (E(Y) - E(Z))^2$.
Anm. Notera att om $E(Y) = E(Z)$ så gäller
$V(X) = p \cdot V(Y) + (1 - p) \cdot V(Z)$.

3.7 Några vanliga diskreta fördelningar

Somliga typer av slumpvariabler dyker upp ofta varför det kan vara bra att studera dem lite mer i detalj en gång för alla. En diskret slumpvariabels slumpstruktur bestäms ju av dess fördelningsfunktion (eller dess sannolikhetsfunktion). Man brukar därför oftast använda ordet ”fördelning” när man beskriver en viss slumpstruktur. En diskret fördelning är således en fördelning svarande mot en diskret slumpvariabel och motsvarande för kontinuerlig fördelning. I detta avsnitt kommer vi att gå igenom ett antal diskreta fördelningar, beskriva i vilka situationer de kan uppstå, samt härleda egenskaper för dem.

3.7.1 Enpunktsfördelning

Den enklaste fördelningen får nog sägas vara den som svarar mot en slumpvariabel ”utan slump”, dvs. en slumpvariabel som bara kan anta ett enda värde, a säg. En sådan fördelning kallas enpunktsfördelning och säges ibland vara trivial eller urartad.

DEFINITION 3.14 (ENPUNKTSFÖRDELNING)

En slumpvariabel X säges vara *enpunktsfördelad* med värde a om $p_X(a) = P(X = a) = 1$.

Det är inte svårt att beräkna läges- och spridningsmått för denna variabel.

SATS 3.8

För en enpunktsfördelad slumpvariabel med värde a gäller $E(X) = a$, $D(X) = V(X) = 0$.

BEVIS

Vi har $E(X) = \sum_k p_X(k) = ap_X(a) = a$, och $V(X) = \sum_k (k - a)^2 p_X(k) = (a - a)^2 p_X(a) = 0$. Av det senare följer direkt att $D(X) = R(X) = 0$.

3.7.2 Tvåpunktsfördelning

Den enklaste icke-triviala diskreta fördelningen är den där två värden kan antas, *tvåpunktsfördelningen*.

DEFINITION 3.15 (TVÅPUNKTSFÖRDELNING)

En diskret slumpvariabel Y säges vara *tvåpunktsfördelad* med värden a och b ($a \neq b$) om $p_Y(a) = p$ och $p_Y(b) = 1 - p =: q$, för något $0 \leq p \leq 1$.

[Bild saknas]

Figur 3.13. *Sannolikhetsfunktion för en tvåpunktsfördelning med a =, b = och p =.*

ANMÄRKNING 3.16

Ofta, men inte alltid, är inom sannolikhetsteorin q reserverad till att betyda $1 - p$*, dvs. komplementärsannolikheten till p. Detta kan man dock inte alltid utgå ifrån utan det bör framgå av sammanhanget.*

SATS 3.9

För en tvåpunktsfördelad slumpvariabel gäller

$$E(Y) = pa + qb, \quad V(Y) = pq(a-b)^2 \quad \text{och} \quad D(Y) = \sqrt{pq}|a-b|.$$

BEVIS

Satsen följer direkt från definitionerna. Variansuttrycket får man med formeln $V(X) = \sum_k k^2 p_Y(k) - (E(Y))^2 = a^2 p + b^2 q - (ap + bq)^2$ vilket efter lite algebra blir formeln ovan när man utnyttjar att $q = 1 - p$.

ÖVNING 3.25

Ett lotteri kostar en krona att delta i. Med sannolikhet 0.005 tjänar man 100 kr (dvs. man får 101 kr tillbaka) och med resterande sannolikhet förlorar man insatsen. Låt Z ange nettoresultatet för ett spel. Härled sannolikhetsfunktionen och beräkna väntevärde och varians.

3.7.3 Bernoullifördelning

Ett viktigt specialfall av tvåpunktsfördelningen är när $a = 1$ och $b = 0$. Man brukar ibland beskriva de två möjliga utfallen som ”lyckat” respektive ”misslyckat”, och lyckat ger värdet 1 och misslyckat värdet 0.

DEFINITION 3.16 (BERNOULLIFÖRDELNING)

En slumpvariabel Z är *Bernoullifördelad* om $P(Z = 0) = p_Z(0) = 1 - p$ och $P(Z = 1) = p_Z(1) = p$, för något $0 \leq p \leq 1$. Man skriver detta med beteckningen $Z \sim Be(p)$.

ANMÄRKNING 3.17

Bernoullifördelningen är uppkallad efter den schweiziske matematikern Jakob Bernoulli (1654-1705) som för övrigt hade flera nära släktingar som också var framstående matematiker.

Som en direkt följd av läges- och spridningsmåtten för tvåpunktsfördelningen får vi motsvarande resultat för Bernoullifördelningen.

FÖLJDSATS 3.1

Om $Z \sim Be(p)$ gäller $E(Z) = p$, $V(Z) = p(1 - p)$ och $D(Z) = \sqrt{p(1 - p)}$.

EXEMPEL 3.27 (*Lotteri*)

Ett lotteri ger vinst med sannolikhet $p = 0.05$ och ingen vinst med resterande sannolikhet. Låt $Z = 1$ indikera att en inköpt lott ger vinst och $Z = 0$ att det var en nitlott. Det gäller i så fall att $Z \sim Be(0.05)$ och $E(Z) = 0.05$ samt $D(Z) = \sqrt{pq} = \sqrt{0.05 \cdot 0.95} \approx 0.218$.

3.7.4 Diskret likformig fördelning

En annan vanligt förekommande slumpvariabel är när alla heltal mellan 1 och något positivt heltal n kan antas, och alla utfall har samma sannolikhet. Vi säger då att slumpvariabeln är diskret likformig.

DEFINITION 3.17 (DISKRET LIKFORMIG FÖRDELNING)

En slumpvariabel X säges vara *diskret likformigt fördelad* på heltalen $1, \ldots, n$ om $p_X(k) = 1/n$ för $k = 1, \ldots, n$.

ANMÄRKNING 3.18

Notera att denna fördelning är ett specialfall av likformig sannolikhetsfördelning som definierades i Avsnitt 2.3 på sidan 11.

[Bild saknas]

Figur 3.14. *Diskret likformig fördelning.*

SATS 3.10

För en slumpvariabel X med diskret likformig fördelning på $1, \ldots, n$ gäller

$$E(X) = \frac{n+1}{2},$$
$$V(X) = \frac{(n+1)(n-1)}{12},$$
$$D(X) = \sqrt{\frac{(n+1)(n-1)}{12}},$$
$$R(X) = \sqrt{\frac{n-1}{3(n+1)}}.$$

BEVIS

Från definitionen av väntevärde får vi

$$E(X) = \sum_{k=1}^{n} k p_X(k) = (1 + \ldots + n)/n = (n+1)/2.$$

Den sista likheten erhölls från relationen $1 + \ldots + n = n(n+1)/2$. För variansen beräknar vi

$$\sum_{k=1}^{n} k^2 p_X(k) = \frac{1^2 + \ldots + n^2}{n} = \frac{n^2}{3} + \frac{n}{2} + \frac{1}{6}.$$

Den sista likheten följer av relationen $1^2 + \ldots n^2 = n^3/3 + n^2/2 + n/6$. Vi använder nu räkneregeln från Sats 3.5 på sidan 64 och får

$$V(X) = \frac{n^2}{3} + \frac{n}{2} + \frac{1}{6} - \left(\frac{n+1}{2}\right)^2,$$

vilket efter lite förenklingar ger just $V(X) = (n+1)(n-1)/12$.

EXEMPEL 3.28 (*Examinationsordning*)

I en klass bestående av 20 studenter skall turordningen för muntlig examination bestämmes med lottens hjälp. 20 lappar numreras från 1 till 20 och läggs i en skål. Du drar en av lapparna utan att titta ner i skålen. Låt

X ange din turordning. Då gäller att X är diskret likformigt fördelad på $1, \ldots, 20$. Din förväntade turordning blir $(20+1)/2 = 10.5$ med standardavvikelse $\sqrt{(20+1)(20-1)/12} \approx 5.76$.

ÖVNING 3.26

Låt X vara ett slumpmässigt heltal mellan 1 och 100. Ange sannolikhetsfunktionen för X och beräkna väntevärde och standardavvikelse.

3.7.5 Binomialfördelning

En situation som ofta dyker upp i olika tillämpningar är att ett försök som kan sluta på två möjliga sätt, lyckat eller misslyckat, upprepas ett bestämt antal gånger. Antag att försöket upprepas $n \geq 1$ oberoende gånger och att ett enskilt försöket blir lyckat med sannolikhet p. Låt Y vara slumpvariabeln som anger hur många försök som lyckas (bland de n försöken). Vi ska nu härleda sannolikhetsfunktionen $p_Y(k)$ för $k = 0, \ldots, n$; för övriga heltal är sannolikhetsfunktionen förstås 0. Att inget av försöken lyckas har sannolikhet $p_Y(0) = (1-p)^n$ eftersom försöken är oberoende och misslyckas med sannolikhet $1-p$. Att exakt ett försök lyckas kan ske på olika sätt. Om vi låter de n försökens resultat betecknas med L eller M beroende på om det är lyckat eller misslyckat kan vi fått 1 lyckat på följande sätt: $(LMM\ldots M)$, $(MLMM\ldots M)$, $(MMLM\ldots M)$, …, $(MM\ldots ML)$, dvs. n olika sätt beroende på vilken gång det misslyckade inträffade. Vart och ett av dessa scenarier har sannolikhet $p(1-p)^{n-1}$ eftersom ett försök lyckas och resterande $n-1$ misslyckas Totalt får vi således $p_Y(1) = np(1-p)$. Mer allmänt kan vi få k av n lyckade på många olika sätt. Hur många sätt detta kan ske på har vi faktiskt redan gått igenom i Avsnitt 2.4 på sidan 16. Där visades nämligen att man kan välja ut k bland n på $\binom{n}{k}$ antal sätt. Varje sådant val har sannolikheten $p^k(1-p)^{n-k}$ eftersom k försök ska lyckas och $n-k$ misslyckas. Totalt får vi således den allmänna formeln $p_Y(k) = \binom{n}{k}p^k(1-p)^{n-k}$, $k = 0, 1, \ldots n$. Denna fördelning dyker upp i många sammanhang och kallas binomialfördelningen.

DEFINITION 3.18 (BINOMIALFÖRDELNING)

En diskret slumpvariabel Y sägs vara *binomialfördelad* med parametrar n och p ($0 \leq p \leq 1$) om sannolikhetsfunktionen ges av

$$p_Y(k) = P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \ldots, n.$$

Man skriver $Y \sim \text{Bin}(n, p)$.

Att sannolikhetsfunktionen summerar sig till 1 följer av binomialteoremet. Det gäller ju nämligen att

$$1 = 1^n = (p + (1-p))^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}.$$

[Bild saknas]

Figur 3.15. *Sannolikhetsfunktionen för binomialfördelningen.*

SATS 3.11

Om $Y \sim \text{Bin}(n, p)$ gäller

$$E(Y) = np,$$
$$V(Y) = np(1-p),$$
$$D(Y) = \sqrt{np(1-p)},$$

BEVIS

$E(Y) = \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k}$ (termen för $k = 0$ kan tas bort eftersom dess bidrag blir 0). Men,

$$k\binom{n}{k} = k\frac{n!}{k!(n-k)!} = \frac{n!}{(k-1)!(n-k)!} = n\frac{(n-1)!}{(k-1)!(n-k)!}$$
$$= n\binom{n-1}{k-1}.$$

Om vi dessutom bryter ut ett p får vi således

$$E(Y) = np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}.$$

Om vi skiftar summationsindex ett steg nedåt ser vi att summan utgör alla termer för en $\text{Bin}(n-1, p)$ variabel som alltså summerar sig till 1. Av detta får vi att $E(Y) = np$. För att härleda variansuttrycket behöver vi beräkna $E(Y^2)$ som kan skrivas som

$$E(Y^2) = \sum_{k=1}^{n} k^2\binom{n}{k}p^k(1-p)^{n-k}$$
$$= \sum_{k=1}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} + \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k}.$$

Den andra termen blir np enligt ovan och den första kan man visa att den blir $n(n-1)p^2$ på liknande sätt och överlåts till läsaren (Övning 3.30).

Det kan ta lite tid att räkna ut sannolikheterna ovan med en miniräknare. Om n (och k) är lite större finns även viss fara för avrundningsfel eftersom vi då multiplicerar mycket stora tal med mycket små. Än mer tidsödande blir det att beräkna fördelningsfunktionen. Av denna anledning finns fördelningsfunktionen för binomialfördelningen tabulerad längst bak i boken (Tabell 2 på sidan ??). Tabellen finns för $n = 2$ upp till $n = 20$ och för $p = 0.05,\ 0.1,\ 0.15,\ 0.2,\ 0.25,\ 0.3,\ 0.4$ och $p = 0.5$. T.ex. får man ur Tabell 2 att om X är $\text{Bin}(n = 10, p = 0.3)$ så gäller att $F_X(5) = 0.9527$. Om man har ett p som är större än 0.5 kan tabellen ändå kanske användas. Genom att räkna antalet ”misslyckade” i stället för antalet ”lyckade” får man nämligen också en binomialfördelning med samma n, men med nytt p som är 1 minus

det ursprungliga p:et. Detta resultat är ganska användbart varför vi formulerar det som en sats.

SATS 3.12

Om X är Bin(n, p) så gäller att $Y = n - X$ är Bin$(n, 1 - p)$.

Hur gör man då om varken p eller $1 - p$ finns i Tabell 2? Om n är litet (mindre än t.ex. 5) bör man räkna exakt med miniräknare eller dator. Har man tillgång till dator med lämplig mjukvara behöver man för den delen aldrig använda Tabell 2. För lite större n kan man ju beräkna eftersökt sannolikhet med närmast mindre och närmast större p vilket ger begränsningar på sannolikheten, möjligen kan man även våga sig på en linjär interpolation för att få en approximation av eftersökt sannolikhet. Om i stället n är stort (Tabell 2 går upp till $n = 20$) kan binomialfördelningen approximeras med någon annan fördelning (normalfördelningen eller Poissonfördelningen). Mer om detta i Avsnitt 3.14 på sidan 165.

Om man först gör n_1 oberoende upprepningar av ett försök med sannolikhet p för lyckat, och därefter gör n_2 nya upprepningar av samma försök (med samma sannolikhet p för lyckat) så har vi tillsammans $n_1 + n_2$ upprepningar av ett försök, så antalet lyckade totalt bör ju vara Bin$(n_1 + n_2, p)$. Att så verkligen är fallet följer av Sats 3.13 nedan. Vi har ännu inte alla verktyg för att bevisa satsen utan sparar detta till Exempel 3.45 på sidan 142.

SATS 3.13

Låt $X \sim$ Bin(n_1, p) och $Y \sim$ Bin(n_2, p) vara oberoende slumpvariabler. Då gäller att $Z = X + Y \sim$ Bin$(n_1 + n_2, p)$.

Vi avslutar delavsnittet med ett exempel där vi använder resultaten ovan samt även redogör för hur Tabell 2 används.

EXEMPEL 3.29

En DNA sekvenator lyckas sekvensera ett hårstrå med sannolikhet 0.7 vid varje försök, och olika försök lyckas oberoende av varandra. Dag 1 sekvenseras 7 hårstrån och dag 2 sekvenseras 4 hårstrån. Antag att vi vill beräkna chansen att a) exakt 4 hårstrån lyckas dag 1, b) att minst 3 lyckas dag 2, och c) att högst 6 lyckas totalt. Låt X vara antal som lyckas dag 1 och Y antal som lyckas dag 2. Då gäller att $X \sim$ Bin$(n = 7, p = 0.7)$ och

$Y \sim \text{Bin}(n = 4, p = 0.7)$ och dessa är oberoende. Uppgift a består alltså av att beräkna $P(X = 4)$ vilket vi gör på två sätt. Dels kan man göra det direkt från sannolikhetsfunktionen: $P(X = 4) = \binom{7}{4}0.7^4 0.3^3 = 0.227$. Man kan i stället använda Tabell 2 även om det just i detta fall inte blir lättare. Eftersom $p = 0.7$ som är större än 0.5 och således inte finns med i tabellen så betraktar vi i stället antalet misslyckade $7 - X$ som är $\text{Bin}(n = 7, p = 0.3)$. Vi får $P(X = 4) = P(7 - X = 7 - 4) = P(7 - X = 3)$. Det är dock inte sannolikhetsfunktionen utan fördelningsfunktionen som finns tabulerad, så vi utnyttjar följande likhet: $P(7-X = 3) = P(7-X \leq 3) - P(7 - X \leq 2)$. Från Tabell 2 kan vi läsa av dessa värden och ser att vi får $P(X = 4) = P(7 - X \leq 3) - P(7 - X \leq 2) = 0.8740 - 0.6471 = 0.2269$.

I b) vill vi beräkna $P(Y \geq 3)$. ”Minst tre lyckade” är detsamma som ”högst en misslyckad”, dvs. $P(Y \geq 3) = P(4-Y \leq 4-1) = P(4-Y \leq 1)$. Antalet misslyckade dag 2, dvs. $4 - Y$, är $\text{Bin}(n = 4, p = 0.3)$, så från Tabell 2 får vi 0.6517. Vi hade även kunnat *räkna* ut denna sannolikhet: $P(Y \geq 3) = P(Y = 3) + P(Y = 4) = \binom{4}{3}0.7^3 0.3^1 + \binom{4}{4}0.7^4 0.3^0 = 0.4116 + 0.2401 = 0.6517$.

För c) använder vi Sats 3.13 ovan. Denna medför att $Z = X + Y$, antalet lyckade totalt på de två dagarna, är $\text{Bin}(n = 11, p = 0.7)$ (eftersom vi totalt har 11 hårtstrån och varje sekvensering lyckas med sannolikhet 0.7 är ju detta inget förvånande). Vi får att $P(Z \leq 6) = P(11-Z \geq 11-6) = P(11-Z \geq 5) = 1-P(11-Z \leq 4)$, och denna är enligt Tabell 2 lika med 0.7897 ($11 - Z \sim \text{Bin}(n = 11, p = 0.3)$). Även denna sannolikhet hade vi kunnat *räkna* ut med hjälpa av sannolikhetsfunktionen: $P(11 - Z \leq 4) = \sum_{k=0}^{4} P(11 - Z = k)$, men eftersom detta tar någon minut med en miniräknare kommer tabellen väl till pass.

ÖVNING 3.27

Låt $X \sim \text{Bin}(n = 12, p = 0.15)$. Bestäm

a) $E(X)$ och $D(X)$

b) $P(X \leq 1)$

c) $P(X \geq 3)$

d) $P(X = 4)$

ÖVNING 3.28

Låt $Y \sim \text{Bin}(n = 16, p = 0.75)$. Bestäm

a) $E(Y)$ och $D(Y)$.
b) $P(Y \leq 12)$
c) $P(Y > 10)$

ÖVNING 3.29

Antag att i genomsnitt var 10:e bil som passerar Rialaavfarten på E18 kör för fort och att olika bilar håller oberoende hastigheter (således antar vi t.ex. att det inte finns köbildningar). En polis mäter hastigheten på 15 bilar. Vad är sannolikheten att exakt 3 bilar kör för fort respektive sannolikheten att minst 3 av dessa kör för fort?

ÖVNING 3.30

Visa att $\sum_{k=1}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} = n(n-1)p^2$ som användes i beviset av Sats 3.11.

ÖVNING 3.31

Ulrika har ett nytt spam-filter för att rensa bort massförsändelser m.m. bland sina e-brev. Hennes uppfattning är att ca 20% av alla e-brev är spamförsändelser. Antag att hon senaste helgen fick 7 e-brev på lördagen och 5 e-brev på söndagen. Antag att spamfiltret är perfekt, dvs. fångar upp alla spam och inga riktiga försändelser. Vad är sannolikheten att
a) minst 3 e-brev spamklassades på lördagen?
b) inget e-brev spamklassades på söndagen?
c) högst ett e-brev spamklassades under helgen?

3.7.6 Hypergeometrisk fördelning

I föregående avsnitt beskrevs hur ett försök som kunde sluta på två olika sätt (”lyckat” och ”misslyckat”) upprepades ett givet antal gånger n, och varje försök lyckades oberoende av varandra och med samma sannolikhet p. I innevarande avsnitt betraktar vi en liknande situation, men med skillnaden att chansen för ”lyckat” respektive ”misslyckat” ändras efter hand pga att det finns ett ändligt antal lyckade respektive misslyckade ”försök”. Vi har

således en ändlig mängd element och varje enhet är en av två sorter (lyckad-misslyckad, svart-vit, sjuk-frisk, ...). När man tar ett element ändras de kvarvarande proportionerna lyckade och misslyckade vilket gör att motsvarande sannolikheter ändras.

Ett konkret exempel är om vi har en vanlig kortlek med 52 kort. Antag att vi samlar på hjärter och skall dra fem kort. Vad är chansen att det blir t.ex. tre hjärter? Eftersom sannolikheten för hjärter i respektive dragning beror på vad vi fått i tidigare dragningar skulle vi kunna erhålla sannolikheten genom att betinga med avseende om vi får hjärter eller ej i respektive dragning, lite på samma sätt som i Figur 2.5 på sidan 21 när vi drog två kort och beräknade chansen att få ett ess i respektive dragning. Detta är dock en mödosam väg eftersom vi i detta fall måste göra fyra betingningar innan vi kan beräkna sannolikheten att få hjärter sista kortet vi drar. Ett betydligt enklare sätt att beräkna sannolikheten att få tre hjärter är genom att konstatera att varje uppsättning av fem kort har samma sannolikhet. Vår kortdragning är således likformig bland alla uppsättningar med 5 kort. Svaret på frågan är således antal sätt som man kan dra 5 kort ur en kortlek så att 3 blir hjärter dividerat med antal sätt totalt som man kan dra 5 kort (se Klassiska sannolikhetsdefinitionen, Sats 2.2 på sidan 11). Antalet sätt att välja ut 5 kort bland 52 vet vi sedan tidigare att det är $\binom{52}{5}$. Men på hur många sätt kan vi göra det så att 3 blir hjärter (och två icke-hjärter)? Jo, då måste vi välja de tre hjärterna bland de 13 som finns ($\binom{13}{3}$ sätt) och så övriga två bland icke-hjärter ($\binom{39}{2}$ sätt). Dessa kan ju kombineras hur man vill, så svaret är att vi kan välja tre hjärter på $\binom{13}{3}\binom{39}{2} = 211926$ sätt. Vi har således kommit fram till att sannolikheten att få 3 hjärter när vi drar 5 kort ur en kortlek är $\binom{13}{3}\binom{39}{2}/\binom{52}{5} = 0.0815$.

Om vi ersätter kortlekens storlek 52 med N, ursprungligt antal lyckade med m (13 i kortexemplet) och antalet dragna element med n får vi en allmän fördelning som vi nu definierar.

DEFINITION 3.19 (HYPERGEOMETRISK FÖRDELNING)

En diskret slumpvariabel X sägs vara *hypergeometriskt fördelad* med parametrar N, n och m (heltal där $0 \leq n \leq N$ och $0 \leq m \leq N$) om sannolikhetsfunktionen ges av

$$p_X(k) = P(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, \quad k = 0, \ldots, n.$$

Man skriver $X \sim \text{Hyp}(N, n, m)$.

ANMÄRKNING 3.19

Om $n > m$ eller $n > N - m$, dvs. att vi drar fler element än det finns lyckade eller misslyckade, kan X inte anta alla värden från 0 till n, varför vi i sådana fall kan inskränka antalet möjliga utfall. T.ex. kan ju om $m = 1$ X aldrig blir större än 1 även om vi drar fler element. Som tur är behöver vi inte oroa oss för det eftersom sannolikheten ovan i dylika fall blir 0 (t.ex. är $\binom{1}{2} = 0$).

ANMÄRKNING 3.20

Binomialkoefficienterna ovan kan vara numeriskt instabila om N är stor (t.ex. ≥ 15). Man får då förenkla uttrycken eftersom flertalet faktorer tar ut varandra i täljare och nämnare.

För att försäkra oss om att det verkligen är en sannolikhetsfunktion bör vi checka att sannolikheterna för de olika utfallen summerar sig till 1, dvs. att $\sum_k \binom{m}{k}\binom{N-m}{n-k}/\binom{N}{n} = 1$, vilket är detsamma som att visa att $\sum_k \binom{m}{k}\binom{N-m}{n-k} = \binom{N}{n}$. Det senare är emellertid en väljkänd kombinatorisk likhet. Vi hoppar därför över dess bevis. I ord säger den att, för att välja n bland N måste vi välja k bland de m första och resterande $n - k$ bland de $N - m$ övriga, för något k.

I inledningen av avsnittet förklarades en situation då den hypergeometriska fördelningen uppstår. Detta kan lätt generaliseras som ett allmänt resultat vilket beskrivs i följande sats.

SATS 3.14

Antag att en mängd innehåller N element, varav m $(0 \leq m \leq N)$ är lyckade och $N - m$ är misslyckade. Antag vidare att vi drar n element helt slumpmässigt utan återläggning, dvs. utan att lägga tillbaka elementen mellan dragningarna. Om vi låter X ange antalet lyckade element vi drar så är $X \sim \text{Hyp}(N, n, m)$.

Vi härleder nu momenten för hypergeometrisk fördelning.

SATS 3.15

Om $X \sim \text{Hyp}(N, n, m)$ gäller

$$E(Y) = np,$$
$$V(Y) = np(1-p)\frac{N-n}{N-1},$$
$$D(Y) = \sqrt{np(1-p)\frac{N-n}{N-1}},$$

där $p = m/N$ är andelen lyckade.

ANMÄRKNING 3.21

Faktorn $(N-n)/(N-1)$ *brukar kallas ändlighetskorrektion. Anledningen är att den går mot 1 ju större populationen* N *är, och den spelar således bara någon väsentlig roll om andelen element som dras* (n/N) *inte är försumbar.*

BEVIS

Det gäller att

$$E(X) = \sum_k k\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} = m\sum_k \frac{\binom{m-1}{k-1}\binom{N-m}{n-k}}{\binom{N}{n}}$$
$$= n\frac{m}{N}\sum_k \frac{\binom{m-1}{k-1}\binom{N-m}{n-k}}{\binom{N-1}{n-1}} = n\frac{m}{N}.$$

Den sista likheten följer av att vi summerar sannolikheterna för en $\text{Hyp}(N-1, n-1, m-1)$ vilket blir 1. Beviset för variansen sparas till Övning 3.34.

EXEMPEL 3.30

En pokerspelare får fem kort och hoppas på sin favoritfärg hjärter. Om vi låter Y beteckna antalet hjärter bland de fem korten betyder det att $X \sim \text{Hyp}(N = 52, n = 5, m = 13)$ eftersom det finns 52 kort, spelaren

”drar” 5 kort, och totalt finns 13 hjärter. Chansen att t.ex. få tre hjärter blir då

$$P(X = 3) = \frac{\binom{13}{3}\binom{39}{2}}{\binom{52}{5}} = \frac{\frac{13\cdot12\cdot11}{3\cdot2\cdot1}\frac{39\cdot38}{2\cdot1}}{\frac{52\cdot51\cdot50\cdot49\cdot48}{5\cdot4\cdot3\cdot2\cdot1}} = 0.0815,$$

vilket visats tidigare. Förväntat antal hjärter blir $E(X) = 5\frac{13}{52} = 1.25$, variansen blir $V(X) = 5 \cdot 0.25 \cdot 0.75\frac{47}{51} = 0.0864$ och standardavvikelsen blir $D(X) = \sqrt{V(X)} = 0.930$.

Den hypergeometriska fördelningen förekommer således när vi studerar egenskaper hos element eller individer i *ändliga* populationer. Detta gäller således de flesta biologiska populationer, medan mer abstrakta populationer som t.ex. populationen bestående av alla potentiella mätningar av någonting är oändlig och då är binomialfördelningen den relevanta. Det som skiljer binomialfördelningen från hypergeometrisk fördelning är att för binomialfallet är chansen för ”lycket” densamma hela tiden, oberoende av tidigare dragning, medan chansen för ”lyckat” efter hand ändras i det hypergeometriska fallet, och hur det ändras beror på tidigare dragningar.

Om vi har en ändlig population men lägger tillbaka det dragna elementet innan vi drar ett nytt har vi samma sannolikhet för lyckat i varje dragning, och denna sannolikhet beror inte på tidigare dragningar. Vi har således även då att göra med binomialfördelningen. Av denna anledning brukar man man säga att binomialfördelningen handlar om dragning *med återläggning* medan hypergeometrisk handlar om dragning *utan återläggning*.

Om vi har en stor population och drar utan återläggning så ändras sannolikheten för lyckat beroende på tidigare dragningar, men dock inte så mycket. Om vi t.ex. ska dra två element bland hundra varav 50 är ”lyckade” så är chansen att få lyckat i andra dragningen $49/99 \approx 0.495$ respektive $50/99 \approx 0.505$ beroende på om vi fick lyckat i första dragningen eller inte, vilket inte är någon större skillnad. Dessutom är de två fördelningarnas väntevärden desamma, och om populationen (N) är stor är varianserna också nästa lika. Man kan därför kanske ana att hypergeometrisk fördelning liknar binomialfördelningen i detta fall. stämmer och kommer att tas upp senare i Avsnitt 3.14 på sidan 165 då vi diskuterar approximationer av olika fördelningar.

ÖVNING 3.32

En urna innehåller 10 bollar varav 4 är röda och 6 vita och Vilma ska dra tre bollar utan återläggning. Låt X beteckna antalet röda bollar som Vilma får.

a) Beräkna $E(X)$ och $D(X)$.
b) Beräkna $p_X(2)$

ÖVNING 3.33

Ett lager består av 100 exemplar av en viss elektronisk komponent. Man vet att 6 av dessa har smärre defekter men inte vilka (defekten upptäcks genom tidskrävande mätningar). Till en kund vill man leverera 10 komponenter utan defekt och chauffören chansar på att bara ta med sig elva (slumpvis) valda komponenter utan att undersöka dem. Vad är chansen att kunden blir nöjd (dvs. att hon får minst 10 enheter utan defekt)?

ÖVNING 3.34

Visa att uttrycket för variansen i Sats 3.15 stämmer. (L)

ÖVNING 3.35

Antag att $Y \sim \text{Bin}(n = 5, p = 0.3)$ och $X \sim \text{Hyp}(N = 100, n = 5, m = 30)$, dvs. andelen lyckade är $m/N = 0.3$ precis som sannolikheten för lyckat i binomialfallet. Illustrera att dessa fördelningar då är ganska lika genom att beräkna $P(Y = 2)$ och $P(X = 2)$.

3.7.7 Poissonfördelning

Vi ska nu definiera Poissonfördelningen. I motsats till de tidigare definierade fördelningarna kan man inte lika lätt beskriva ett slumpexperiment då Poissonfördelningen uppstår. Likväl är det en fördelning som ofta dyker upp. Situationerna den dyker i upp kan i huvudsak delas upp i två fall: 1. där man har en binomialfördelning eller hypergeometrisk fördelning med litet p, och 2. i fallet där händelser sker slumpmässigt i tiden och vi betraktar antalet händelser i ett tidsintervall. Den första situationen beskrivs mer i detalj i Avsnitt 3.14 och den senare i Avsnitt **??** på sidan **??**. Vi definierar nu fördelningen.

DEFINITION 3.20 (POISSONFÖRDELNING)

En diskret slumpvariabel X sägs vara *Poissonfördelad* med parameter λ ($\lambda > 0$) om sannolikhetsfunktionen ges av

$$p_X(k) = P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots.$$

Man skriver $X \sim \text{Po}(\lambda)$.

ANMÄRKNING 3.22

Fördelningen är uppkallad efter den franske matematikern Siméon Denis Poisson (1781-1840).

Vi bör ju förstås verifiera att det verkligen är en sannolikhetsfunktion, dvs. att sannolikheterna summerar sig till 1. Men, det är förhoppningsvis bekant att Taylorutvecklingen av e^x ges av $1 + x^1/1! + x^2/2! + \ldots$. Av detta följer att $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$, vilket i sin tur implicerar att sannolikheterna summerar sig till 1.

I Figur 3.16 illustreras Poissonfördelningen för ett par olika λ. Som synes

Bild saknas

Figur 3.16. *Sannolikhetsfunktionen för Poissonfördelning för några olika val av λ.*

tenderar fördelningen ge större värden ju större λ är. Ett faktum som även illustreras av väntevärdets form i följande sats.

SATS 3.16

Om $Z \sim \text{Po}(\lambda)$ gäller

$$E(Z) = \lambda,$$
$$V(Z) = \lambda,$$
$$D(Z) = \sqrt{\lambda}.$$

BEVIS

Vi har $E(Z) = \sum_{k=0}^{\infty} k\lambda^k e^{-\lambda}/k! = \lambda \sum_{k=1}^{\infty} \lambda^{k-1} e^{-\lambda}/(k-1)! = \lambda$, där den sista likheten följer av att man genom att byta summationsindex ser att vi summerar alla sannolikheter för en $\text{Po}(\lambda)$-fördelningen vilket således blir 1. På liknande sätt visar man att $E(Z^2)$ blir $\lambda^2 + \lambda$ genom att i summauttrycket skriva k^2 som $k(k-1)+k$ och summera termerna var för sig. Variansen blir således $V(Z) = E(Z^2) - (E(Z))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$ vilket visar vad satsen påstår. Standardavvikelsen fås som vanligt genom att beräkna roten ur variansen.

Fördelningsfunktionen $F_Z(k)$ för Poissonfördelningen finns tabulerad för ett antal värden på λ längst bak i boken (Tabell 3 på sidan ??). T.ex. ser man där att om $Z \sim \text{Po}(\lambda = 6)$ gäller $F_Z(7) = 0.7440$. Sannolikhetsfunktionen $p_Z(k)$ får man med hjälp av relationen $p_Z(k) = F_Z(k) - F_Z(k-1)$, eller förstås genom formeln för sannolikhetsfunktionen direkt. Tabellen är gjord för att underlätta beräkningar av fördelningsfunktionen då man ju måste summera sannolikhetsfunktionen för ett utfall upp till efterfrågat värde. Om man har ett λ som ligger mellan två λ-värden i Tabell 3 bör man utföra beräkningarna med miniräknare eller dator. Om ett mindre fel inte spelar någon roll kan man förstås alternativt titta i tabellen för närmaste värde på λ. Om man har ett λ som är större än det största som finns i Tabell 3 (λ=15) kan Poissonfördelningen approximeras med normalfördelningen. Mer om hur denna approximation sker, och lite om dess bakomliggande orsak, finner du i Avsnitt 3.14 på sidan 165.

EXEMPEL 3.31

Antalet gånger en stor produktionsmaskin går sönder under ett år Z kan med god approximation beskrivas av en Poissonfördelning med väntevär-

de $\lambda = 3$. Chansen att maskinen går sönder exakt 4 gånger under ett år är således $p_Z(4) = 3^4 e^{-3}/4! = 0.168$. Sannolikheten att maskinen går sönder högst 3 gånger blir från Tabell 3 0.6472. Vill man t.ex. räkna ut $P(Z > 1)$ använder man likheten $P(Z > 1) = 1 - P(Z \leq 1)$ och $P(Z \leq 1)$ kan hittas i Tabell 3, alternativt beräknas från sannolikhetsfunktionen via $P(Z \leq 1) = p_Z(0) + p_Z(1)$. I bägge fallen får man att $P(Z > 1) = 0.8809$.

ÖVNING 3.36

Låt $Y \sim \text{Po}(\lambda = 1.32)$. Bestäm
a) $P(Y = 1)$,
b) $P(Y > 1$,
c) $E(Y)$ och $D(Y)$.

ÖVNING 3.37

Låt $Y \sim \text{Po}(\lambda = 12)$. Bestäm
a) $P(Y = 10)$,
b) $P(Y \geq 13)$

ÖVNING 3.38

Beräkna variationskoefficienten för en $X \sim \text{Po}(\lambda)$.

ÖVNING 3.39

I Sverige sker i genomsnitt 15 stormar per år (påhittad uppgift). Antalet stormar enskilda år kan med god approximation beskrivas av Poissonfördelningen. Vad är under dessa förutsättningar sannolikheten att det blir fler än 20 stormar?

ÖVNING 3.40

Lina uppskattar att hon i genomsnitt får ca 1.6 e-brev per timme under arbetsdagen, och att flödet är ganska jämnt spritt över dagen. I Avsnitt **??**

kommer vi se att en rimlig modell för hur e-brev kommer till hennes e-brevlåda är Poissonprocessen. Denna modell medför att antalet e-brev hon för en tidsperiod av längd t är Poissonfördelad med parameter $\lambda = 1.8 \cdot t$.

a) Vad är chansen att hon får mer än 3 e-brev under en timme?
b) Vad är sannolikheten att hon får högst 10 e-brev under en arbetsdag på 8 timmar? (Använd det närmsta λ i Tabell 3.)

3.7.8 Geometrisk fördelning och besläktade fördelningar

Både för binomialfördelningen och hypergeometrisk fördelning var *antalet* försök bestämt (till n). I andra situationer är så inte fallet utan man upprepar försöket ända tills någon händelse inträffar. Vi har då att göra med geometrisk fördelning, eller för-första-gångenfördelningen, beroende på vad man räknar.

Antag att ett försök kan utfalla på två sätt, ”lyckat” eller ”misslyckat”. Antag vidare att chansen för lyckat försök är p och att resultaten av försöken är oberoende. Detta försök upprepas tills dess att det för första gången blir ett lyckat försök. Låt X ange antalet försök som erfordras. För att X skall vara lika med k måste vi således ha $k-1$ misslyckade i rad följt av 1 lyckat, och enligt multiplikationsprincipen blir sannolikheten för detta således $q^{k-1}p$ där $q = 1-p$. Vi säger att X har för-första-gångenfördelning. Om vi i stället bara räknar de misslyckade innan det lyckade och kaller detta för Y kommer ju denna variabel vara 1 mindre än X. Vi får således att $P(Y = k) = P(X = k+1)$. Denna variabel brukar kallas för en geometriskt fördelad slumpvariabel. Vi inför nu dessa definitioner.

DEFINITION 3.21 (FÖR-FÖRSTA-GÅNGEN OCH GEOMETRISK FÖRDELNING)

En diskret slumpvariabel X sägs vara *för-första-gångenfördelad* med parameter p $(0 < p < 1)$ om sannolikhetsfunktionen ges av

$$p_X(k) = P(X = k) = pq^{k-1}, \quad k = 1, 2, \ldots.$$

Man skriver $X \sim \text{ffg}(p)$.

En diskret slumpvariabel Y sägs vara *geometriskt fördelad* med parameter p $(0 < p < 1)$ om sannolikhetsfunktionen ges av

$$p_Y(k) = P(Y = k) = pq^{k}, \quad k = 0, 1, 2, \ldots.$$

Man skriver $Y \sim \text{Geo}(p)$.

ANMÄRKNING 3.23

Det är förstås lite onödigt att införa två slumpvariabler för i princip samma slumpförsök. I viss litteratur skiljer man heller inte på de två utan använder de två som synonymer, då oftast för fallet att man räknar antal försök totalt (inklusive det lyckade). I engelskspråkig litteratur används endast benämning "Geometric distribution" och syftar då oftast på vad vi ovan kallar ffg. Man bör alltså vara observant på vad som menas när dessa fördelningar nämns.

I Figur 3.17 illustreras den geometriska fördelningen för ett par olika p.

Bild saknas

Figur 3.17. *Sannolikhetsfunktionen för geometrisk fördelning för några olika val av p.*

EXEMPEL 3.32

Erik lägger en patiens som går ut med sannolikheten 0.01. Han har gett sig sjutton på att den skall gå ut och fotsätter tills att den gör det. Antalet patienser han måste lägga X är då ffg($p = 0.01$) och chansen att det skall ske efter exakt 10 försök ges av $p_X(10) = 0.01 \cdot 0.99^9 = 0.00913$. Om vi i ställer räknar Y som antalet misslyckade försök innan det lyckade är $Y \sim \text{Geo}(p = 0.01)$ och $P(Y = 10) = 0.01 \cdot 0.99^10 = 0.00904$. Att de inte blir samma beror ju på att den första sannolikheten är för händelsen att vi gör 10 försök totalt och den andra att vi gör 10 misslyckade följt av ett lyckat (att patiensen går ut).

Till skillnad från flera av de tidigare fördelningarna har fördelningsfunktionen $F_X(k)$ (respektive $F_Y(k)$) enkla uttryck. Det gäller nämligen att $F_X(k) = P(X \leq k) = 1 - P(X > k)$, och $P(X > k)$ är detsamma som

sannolikheten att vi inte har fått något lyckat bland de k första försöken, dvs. att vi bara fått misslyckade bland de k första försöken. Vi har således att $P(X > k) = q^k$ och

$$F_X(k) = 1 - q^k, \qquad F_Y(k) = 1 - q^{k+1}.$$

Av denna anledning behövs ingen tabell för ffg eller geometrisk fördelning. Vi räknar nu ut väntevärde och varians för fördelningen.

SATS 3.17

Om $X \sim \text{ffg}(p)$ och $Y \sim \text{Geo}(p)$ gäller

$$E(X) = \frac{1}{p}, \qquad E(Y) = \frac{q}{p},$$
$$V(X) = \frac{q}{p^2}, \qquad V(Y) = \frac{q}{p^2},$$
$$D(X) = \frac{\sqrt{q}}{p}, \qquad D(Y) = \frac{\sqrt{q}}{p}.$$

BEVIS

Vi visar satsen för ffg och lämnar motsvarande för geometrisk fördelning till läsaren (Övning 3.47). Eftersom den geometriska fördelningen alltid är ett mindre än ffg bör det ju dock inte överraska att väntevärdet är ett mindre ($1/p - 1 = q/p$), och kanske inte heller att spridningen, t.ex. mätt med varians eller standardavvikelse, är oförändrad. För X gäller

$$E(X) = \sum_{k=1}^{\infty} kq^{k-1}p = p\sum_{k=1}^{\infty} \frac{d}{dq}q^k,$$

eftersom $\frac{d}{dq}q^k = kq^{k-1}$. Eftersom $q = 1 - p$ så gäller att $0 < q < 1$ och för sådana q är summan ifråga absolutkonvergent. Vi får därför byta ordning på summation och derivering trots att summan innehåller oändligt många termer, och vi får $E(X) = p\frac{d}{dq}\sum_{k=1}^{\infty} q^k$. Men $\sum_{k=1}^{\infty} q^k = -1 + \sum_{k=0}^{\infty} q^k = -1 + (1-q)^{-1}$. Genom att derivera detta m.a.p. q får vi att $E(X) = p(1-q)^{-2} = 1/p$ vilket skulle bevisas. För variansen beräknar vi

först $E(X^2)$. Vi får

$$E(X^2) = \sum_{k=1}^{\infty} k^2 q^{k-1} p = \sum_{k=1}^{\infty} (k(k-1)+k) q^{k-1} p$$
$$= qp \sum_{k=1}^{\infty} \frac{d^2}{dq^2} q^k + \frac{1}{p},$$

där vi för första termen brutit ut qp, sedan skrivit $k(k-1)q^{k-2}$ som andraderivatan av q^k, och använt att den andra termen är väntevärdet som vi just räknat ut till $1/p$. Eftersom summan är absolutkonvergent kan vi byta ordning på summation och derivering och vi får $\frac{d^2}{dq^2} \sum_{k=1}^{\infty} q^k = \frac{d^2}{dq^2} - 1 + (1-q)^{-1} = 2(1-q)^{-3}$. Detta medför att $E(X^2) = (2-p)/p^2$ och således att variansen ges av $V(X) = E(X^2) - (E(X))^2 = (2-p)/p^2 - 1/p^2 = q/p^2$.

Från väntevärdets uttryck ser vi således att vi i genomsnitt behöver fler försök ju mindre p är. Något som inte är förvånande – ju mindre chans vi har att lyckas i ett enskilt försök ju fler försök lär behövas innan vi lyckas. Även spridningen (variansen eller standardavvikelsen) ökar då p minskar. I patiensexemplet ovan (Exempel 3.32) har vi att $E(X) = 100$ och $D(X) = 99.5$.

För-första-gångenfördelningen är ett specialfall av en mer allmän fördelningen som kallas negativ binomialfördelning. Denna fördelning uppstår för samma typ av experiment, dvs. att ett försök som resulterar i lyckat utfall med sannolikhet p upprepas oberoende gånger, men nu upprepas försöket tills att man gjort r ($r \geq 1$) lyckade försök. ffg-fördelningen är således specialfallet att $r = 1$. Vad är då chansen att vårt r:te lyckade inträffar just efter k försök? Jo, bland de $k-1$ första försöken skall vi ha fått $r-1$ lyckade och $k-1-(r-1) = k-r$ misslyckade, samt att vi i det k:te försöket ska få ett lyckat. Chansen för det senare är förstås p. Chansen att få $r-1$ lyckade på $k-1$ försök är $\binom{k-1}{r-1} p^{r-1} q^{k-r}$, eftersom det kan ske på $\binom{k-1}{r-1}$ sätt, och varje sätt har sannolikheten $p^{r-1} q^{k-r}$. Den eftersökta sannolikheten blir således $p \cdot \binom{k-1}{r-1} p^{r-1} q^{k-r}$. Vi definierar nu denna fördelning.

DEFINITION 3.22 (NEGATIV BINOMIALFÖRDELNING)

En diskret slumpvariabel X sägs vara *negativ binomialfördelad* med parametrar $r \geq 1$ (heltal) och p ($0 < p < 1$) om sannolikhetsfunktionen ges

av

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r q^{k-r}, \quad k = r, r+1, r+2, \ldots.$$

Man skriver $X \sim \text{NegBin}(r, p)$.

EXEMPEL 3.33

Vid en flygplats landar plan i tid (dvs. utsatt tid plus/minus 10 minuter) oberoende av varandra med sannolikhet 0.6. Då är antalet plan X som behöver landar för att 5 skall ha landat i tid $\text{NegBin}(r = 5, p = 0.6)$. T.ex. är chansen att det behövs 10 landningar $p_X(10) = \binom{9}{4} 0.6^5 0.4^5 = 0.1003$.

Det går att härleda väntevärde och varians utifrån definitionen och sannolikhetsfunktionen, men ett betydligt lättare sätt är att använda momenten för ffg-fördelningen tillsammans med insikten att en negativ binomialfördelad slumpvariabel är detsamma som summan av r ffg-variabler, alla med samma p. För att använda detta måste vi dock lära oss om egenskaper hos summor av slumpvariabler vilket behandlas i Avsnitt 3.11.3 på sidan 144. Beviset av satsen nedan sparas därför till det avsnittet, närmare bestämt till Exempel 3.48 på sidan 147.

SATS 3.18

Om $X \sim \text{NegBin}(r, p)$ så gäller

$$E(X) = \frac{r}{p},$$
$$V(X) = \frac{rq}{p^2},$$
$$D(X) = \frac{\sqrt{rq}}{p}.$$

ÖVNING 3.41

Låt $X \sim \text{ffg}(p = 0.3)$ Bestäm
a) $P(X = 4)$,

b) $P(X \geq 4)$,
c) $P(X \leq 2)$,
d) $E(X)$ och $D(X)$.

ÖVNING 3.42

Låt $Y \sim \text{Geo}(p = 0.8)$ Bestäm
a) $P(Y = 1)$,
b) $P(Y \geq 1)$,
c) $P(Y \leq 2)$,
d) $E(Y)$ och $D(Y)$.

ÖVNING 3.43

Antag att kanalarna i Nederländerna fryser till och blir skridskodugliga någon gång under vintern med sannolikheten 0.2 och att detta sker oberoende mellan vintrar (det senare är ett högst rimligt antagande). Vad är chansen att det kommer att ges möjlighet att åka skridskor längs kanalerna de närmsta 3 åren?
a) Vad är chansen att det kommer att ges möjlighet att åka skridskor längs kanalerna de närmsta 3 åren?
b) Vad är förväntat antal isfria vintrar innan första isen ligger?

ÖVNING 3.44

Beräkna variationskoefficienten för $X \sim \text{ffg}(p)$, $Y \sim \text{Geo}(p)$ och $Z \sim \text{NegBin}(r, p)$.

ÖVNING 3.45

Låt $X \sim \text{NegBin}(r = 3, p = 0.75)$. Bestäm
a) $P(X = 3)$,
b) $P(X = 6)$,
c) $P(X \leq 6)$.
d) $E(X)$ och $D(X)$.

ÖVNING 3.46

Tom sätter (dvs. kastar bollen i korgen) varje enskilt straffkast i basketboll med sannolikhet 0.8. En träning avslutas med att man skall skjuta straffar tills man satt 10 kast.

a) Vad är chansen att Tom bara behöver 10 kast?
b) Vad är chansen att han klarar det på exakt 13 försök?
c) Hur många försök är det störst sannolikhet att Tom exakt behöver?

ÖVNING 3.47

Visa att väntevärde och varians för en geometriskt fördelad slumpvariabel stämmer enligt Sats 3.17 på sidan 96. (L)

3.8 Några vanliga kontinuerliga fördelningar

På samma sätt som att vissa typer av diskreta fördelningar/slumpvariabler dyker upp ofta gäller även för kontinuerliga fördelningar/slumpvariabler. Vi studerar nu några av de vanligast förekommande kontinuerliga fördelningarna och härleder egenskaper för desamma.

3.8.1 Kontinuerlig likformig fördelning

Vi har tidigare stött på likformig sannolikhetsfördelning både gällande på godtyckliga diskreta utfallsrum (Definition 2.3 på sidan 11) och för diskreta slumpvariabler (Avsnitt 3.7.4). Även för kontinuerliga slumpvariabler är likformig fördelning vanligt förekommande. Den kanske vanligaste varianten är att slumpvariabeln är likformigt fördelat mellan 0 och 1, men vi behandlar det allmänna fallet då slumpvariabeln är likformigt fördelad mellan a och b ($a < b$). Eftersom vi betraktar en kontinuerlig slumpvariabel kan vi inte definiera den som att alla värden mellan a och b är lika sannolika - enskilda värden har alltid sannolikhet 0 för kontinuerliga slumpvariabler. Det man i stället menar är att sannolikheten att slumpvariabeln ligger i något givet intervall inom $(a,\ b)$ beror bara på intervallets bredd och inte på *var* intervallet ligger. Vi definierar nu denna fördelningen.

DEFINITION 3.23 (LIKFORMIG KONTINUERLIG FÖRDELNING)

En kontinuerlig slumpvariabel U sägs vara *likformigt fördelad* mellan a och b ($a < b$) om täthetsfunktionen ges av

$$f_U(x) = \frac{1}{b-a}, \quad a \leq x \leq b \; (f_U(x) = 0 \text{ annars}).$$

Man skriver $U \sim \text{Re}(a, b)$.

ANMÄRKNING 3.24

Likformig fördelning heter "uniform distribution" på engelska varför förkortningen $U(a, b)$ ofta förekommer.

Som synes är täthetsfunktionen konstant, för de värden som den är positiv. (Detta svarar mot att sannolikhetsfunktionen är konstant för diskret likformig sannolikhetsfördelning.) Fördelningsfunktionen $F_U(u)$ ges således av

$$F_U(u) = \begin{cases} 0 & \text{om } u < a, \\ \frac{u-a}{b-a} & \text{om } a \leq u \leq b, \\ 1 & \text{om } u > b. \end{cases}$$

I Figur 3.18 nedan visas täthetsfunktionen och fördelningsfunktionen. Från

Bild saknas

Figur 3.18. *Täthetsfunktion och fördelningsfunktion för likformig fördelning*

figuren är det uppenbart att väntevärdet, dvs. tyngdpunkten, ligger mitt emellan a och b. I satsen nedan visas detta, liksom uttrycket för standardavvikelsen.

SATS 3.19

Om $U \sim \text{Re}(a, b)$ så gäller

$$E(U) = \frac{a+b}{2},$$
$$V(U) = \frac{(b-a)^2}{12},$$
$$D(U) = \frac{b-a}{\sqrt{12}}.$$

BEVIS

Vi visar påståendet om väntevärdet och överlåter beviset av variansen/standardavvikelsen åt läsaren (Övning 3.51).

$$E(U) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}.$$

EXEMPEL 3.34

En person som inte kan tidtabellen anländer till en busshållplats där bussarna går var tionde minut. Då är väntetiden T för personen likformigt fördelad mellan 0 och 10 ($T \sim \text{Re}[0, 10]$, enhet minuter). Förväntad väntetid blir $E(T) = (0 + 10)/2 = 5$ minuter och standardavvikelsen för väntetiden blir $D(T) = 10/\sqrt{12} \approx 2.89$. Chansen att personen får vänta mer än 7 minuter blir $1 - F_T(7) = 1 - (7-0)/(10-0) = 0.3$ och chansen att väntetiden blir högst 5 minuter blir $F_T(5) = (10-5)/(10-0) = 0.5$.

ÖVNING 3.48

Låt $U \sim \text{Re}[10, 20]$. Bestäm

a) $P(10 \leq U \leq 13$.

b) $P(U > 12)$.

c) $E(U)$ och $D(U)$.

ÖVNING 3.49

Låt $U \sim \text{Re}[-5, 5]$. Bestäm
a) $f_U(u)$
b) $F_U(u)$
c) Intensiteten $\lambda_U(u)$.

ÖVNING 3.50

Ett s.k. slumptal brukar oftast innebära ett tal u från $\text{Re}[0, 1]$ (se mer om detta i Kapitel **??** på sidan **??**). Antag att $U \sim \text{Re}[0, 1]$.
a) Vad är chansen att första decimalen är 5?
b) Vad är chansen att andra decimalen är 5?
c) Vad är variationskoefficienten $R(U)$?

ÖVNING 3.51

Visa att för $U \sim \text{Re}[a, b]$ så gäller att $V(U) = (b-a)^2/12$. (L)

3.8.2 Exponentialfördelning

Vi ska nu gå igenom en positiv kontinuerlig fördelning som förekommer ofta i tillämpningar, nämligen exponentialfördelningen som vi tidigare stött på. Antag att Y är tiden tills en händelse inträffar och att intensiteten med vilket händelsen inträffar är konstant, dvs. att $\lambda_Y(y) = \beta$ för något $\beta > 0$. Innebörden av detta är att chansen att händelsen inträffar i ett kort intervall $(t, t+h)$, betingat av att den inte inträffat innan t, är lika med λh. Det speciella är att intensiteten är *oberoende* av hur långt tid t som har passerat, så intensiteten är konstant. Man brukar därför prata om *minneslöshet* vilket nämndes i Exempel 3.12 på sidan 54.

Hur ser då fördelningsfunktionen $F_Y(y)$ ut för denna intensitet? Vi har ju att

$$\lambda_Y(y) = \frac{f_Y(y)}{1 - F_Y(y)} = \beta,$$

och det gäller alltid att täthetsfunktionen satisfierar $f_Y(y) = F_Y'(y)$. Detta är ju en bekant matematisk funktion. Uttrycket till vänster i ekvationen kan nämligen skrivas som $-\frac{d}{dy}\ln(1 - F_Y(y))$ vilket alltså skall vara lika med β.

Lösningen till detta är ju exponentialfunktionen. Således får vi att $1-F_Y(y) = e^{-\beta y}$ vilket medför att $F_Y(y) = 1 - e^{-\beta y}$ och $f_y(y) = \beta e^{-\beta y}$. Vi inför nu detta som definition.

DEFINITION 3.24 (EXPONENTIALFÖRDELNING)

En kontinuerlig slumpvariabel Y sägs vara *exponentialfördelad* med intensitetsparameter β om täthetsfunktionen ges av

$$f_Y(y) = \beta e^{-\beta y}, \ y > 0, \quad (f_Y(y) = 0, \ y \leq 0).$$

Man skriver $Y \sim \text{Exp}(\beta)$.

ANMÄRKNING 3.25

I vissa böcker använder man en annan parametrisering, nämligen inversen $1/\beta$. Anledningen till detta är, som vi skall se, att väntevärdet är lika med $1/\beta$. När man ser exponentialfördelningen bör man därför dubbelkolla om parametern som anges är intensitetsparametern (som i denna bok) eller väntevärdet.

I Figur 3.19 visas täthetsfunktionen och fördelningsfunktionen för exponentialfördelningen. Vi har redan härlett fördelningsfunktionen och intensiteten

Bild saknas

Figur 3.19. *Täthetsfunktion och fördelningsfunktion för exponentialfördelningen.*

för exponentialfördelningen och sett att dessa satisfierar

$$F_Y(y) = 1 - e^{-\beta y} \qquad \text{och} \qquad \lambda_Y(y) = \beta.$$

EXEMPEL 3.35

Radioaktiva atomer sönderfaller slumpmässigt i tiden med intensitet $\beta = 7$ per sekund. Det betyder att tiden T till nästa sönderfall är Exp(7). Vi kan då beräkna chansen att denna tid överstiger 0.5 sekunder som $P(T > 0.5) = 1 - F_T(0.5) = e^{-7\cdot 0.5} \approx 0.030$. Chansen att det dröjer mellan 0.1 och 0.5 ges av $P(0.1 \leq T \leq 0.5) = F_T(0.5) - F_T(0.1) = (1 - e^{-7\cdot 0.5}) - (1 - e^{-7\cdot 0.1}) \approx 0.466$.

Vi härleder nu momenten för exponentialfördelningen.

SATS 3.20

Om $Y \sim \text{Exp}[\beta]$ så gäller

$$E(Y) = \frac{1}{\beta},$$
$$V(Y) = \frac{1}{\beta 2},$$
$$D(Y) = \frac{1}{\beta}.$$

BEVIS

Väntevärdet ges av $E(Y) = \int_0^\infty y\beta e^{-\beta y} dy$. Om vi partialintegrerar detta får man

$$\begin{aligned} E(Y) &= \left[-ye^{-\beta y}\right]_0^\infty + \int_0^\infty e^{-\beta y} dy \\ &= 0 + \left[-\frac{e^{-\beta y}}{\beta}\right]_0^\infty = \frac{1}{\beta}. \end{aligned}$$

Att $\left[-ye^{-\beta y}\right]_0^\infty = 0$ följer av att $ye^{-\beta y}$ går mot 0 då y går mot oändligheten, och för $y = 0$ blir uttrycket också 0. För att beräkna $E(Y^2)$ använder vi också partialintegration.

$$E(Y^2) = \int_0^\infty y^2 \beta e^{-\beta y} dy = \left[-y^2 e^{-\beta y}\right]_0^\infty + \int_0^\infty 2ye^{-\beta y} dy.$$

Den första termen blir 0 av liknande anledning som ovan, och den andra termen är $\frac{2}{\beta}E(Y)$ vilket gör att $E(Y^2) = \frac{2}{\beta^2}$. Variansen blir därför $V(Y) = E(Y^2) - (E(Y))^2 = 1/\beta^2$.

ÖVNING 3.52

Antag att $Y \sim \text{Exp}(\beta = 4)$. Beräkna
a) $P(Y \leq 0.5)$,
b) $P(0.2 \leq Y \leq 0.4)$
c) $E(Y)$ och $V(Y)$.

ÖVNING 3.53

Låt $T \sim \text{Exp}(\beta = 1/5)$. Beräkna
a) $P(2 \leq T \leq 4)$,
b) $P(4 \leq T \leq 6)$,
c) Variationskoefficienten $R(T)$.

ÖVNING 3.54

Livslängden på ett lysrör som är tänt hela tiden kan beskrivas av en exponentialfördelning (anledningen är att det inte åldras, och när det inte tänds och släcks utsätts det för konstant ”söndringsintensitet”). Antag att det går sönder med intensiteten 3 per år. Beräkna
a) sannolikheten att ett lysrör håller mer än ett år,
b) sannolikheten att det går sönder första månaden,
c) förväntad tid tills det går sönder.

ÖVNING 3.55

Visa att täthetsfunktionen för exponentialfördelningen verkligen är en täthetsfunktion, dvs. att dess integral blir 1.

3.8.3 Normalfördelning

Det är nu dags att gå igenom den kanske viktigaste av alla fördelningar: normalfördelningen, även kallad Gaussfördelningen och ”the bell-shaped curve”. Att den är så viktig beror på att om man summerar många slumpvariabler (tänk: något slumpmässigt som beror på många olika faktorer) så är resultatet nästan alltid normalfördelat. Mer om detta i Avsnitt 3.13 längre fram.

DEFINITION 3.25 (NORMALFÖRDELNING)

En kontinuerlig slumpvariabel X sägs vara *normalfördelad* med parametrar μ och $\sigma > 0$ om täthetsfunktionen ges av

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Man skriver $X \sim \mathrm{N}(\mu, \sigma)$.

ANMÄRKNING 3.26

Som vi senare ska se är första parametern fördelningens väntevärde och den senare parameterns standardavvikelse. I somlig litteratur står variansen (dvs. σ^2) i stället för standardavvikelsen i andra positionen. Om det alltså står 4 kan vissa mena att variansen är 4 medan andra att standardavvikelsen är fyra (och variansen därmed 16!) – således anledning att vara uppmärksam på vad som menas.

I Figur 3.20 visas täthetsfunktionen för $\mathrm{N}(\mu, \sigma)$. Att visa att $f_X(x)$ verkligen

Bild saknas

Figur 3.20. *Täthetsfunktion för en $N(\mu, \sigma)$-fördelning.*

är en täthetsfunktion, dvs. att $\int_{-\infty}^{\infty} f_X(x)dx = 1$, är lite svårare än vad vi

tidigare behövt visa. Om vi gör substitutionen $z = (x - \mu)/\sigma$ får vi

$$\int_{-\infty}^{\infty} f_X(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}dz,$$

så det gäller således att visa att $\int_{-\infty}^{\infty} e^{-z^2/2}dz = \sqrt{2\pi}$. Kvadraten på denna integral, dvs. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(z^2+y^2)/2}dzdy$, måste således visas bli 2π. Detta visar man genom att övergå till polära koordinater ($z = r\cos u$ och $y = r\sin v$) vilket ger integralen

$$\int_0^{2\pi} \left(\int_0^{\infty} re^{-r^2/2}dr \right) du.$$

Den inre integralen kan vi lösa explicit eftersom $-re^{-r^2/2}$ är derivatan av $e^{-r^2/2}$, så $\int_0^{\infty} re^{-r^2/2}dr = 1$ och hela uttrycket blir således 2π vilket var det vi skulle visa.

Fördelningsfunktionen $F_X(t)$ ges således av

$$F_X(t) = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Denna integral har inget slutet uttryck. För givet μ, σ och t kan den däremot beräknas numeriskt.

Från Figur 3.20 är det, via tolkningen av väntevärde som jämviktspunkt, uppenbart att fördelningen har väntevärde μ. Vi visar dock detta, liksom att standardavvikelsen är σ i följande sats.

SATS 3.21

Om $X \sim \mathrm{N}(\mu, \sigma)$ så gäller

$$E(Y) = \mu,$$
$$V(Y) = \sigma^2,$$
$$D(Y) = \sigma.$$

BEVIS

Från definitionen har vi $E(X) = \int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$. Om vi liksom tidigare gör substitutionen $z = (x - \mu)/\sigma$ får vi

$$E(X) = \int_{-\infty}^{\infty} (\sigma z + \mu)\frac{1}{\sqrt{2\pi}} e^{-z^2/2}dz.$$

Den första termen ger bidrag 0 eftersom det är en udda funktion ($\int_{-\infty}^{\infty} ze^{-z^2/2}dz = 0$) medan den andra termen ger bidraget μ eftersom vi tidigare visat att

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}dz = 1.$$

Detta visar således att $E(X) = \mu$. För variansen beräknar vi först $E(X^2)$. Efter variabelsubstitutionen $z = (x - \mu)/\sigma$ får vi

$$E(X^2) = \int_{-\infty}^{\infty} (\sigma z + \mu)^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2}dz.$$

Om vi utvecklar kvadraten $(\sigma z + \mu)^2 = \sigma^2 z^2 + 2\sigma\mu z + \mu^2$ får vi med hjälp av tidigare beräkningar att den andra termens bidrag blir 0 och den tredje termen ger bidraget μ^2. Den första termens bidrag får vi genom partiell integration av $z^2 e^{-z^2/2} = z \cdot ze^{-z^2/2}$. Vi får

$$\begin{aligned}\sigma^2 \int_{-\infty}^{\infty} z \cdot \frac{ze^{-z^2/2}}{\sqrt{2\pi}}dz &= \sigma^2 \left[z \cdot \frac{-e^{-z^2/2}}{\sqrt{2\pi}} \right]_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}dz \\ &= 0 + \sigma^2.\end{aligned}$$

Tillsammans får vi således $E(X^2) = \sigma^2 + \mu^2$, och variansen blir $V(X) = E(X^2) - (E(X))^2 = \sigma^2$.

Att väntevärdet är μ är som sagt uppenbart genom tolkningen av väntevärde som jämviktspunkten i fördelningen. Att standardavvikelsen är just σ kan man däremot inte ”se” från en figur av täthetsfunktionen eller täthetsfunktionens matematiska form. Att spridning, och därmed standardavvikelsen, ökar med σ är dock uppenbart om man betraktar Figur 3.21 där täthetsfunktionen för $\mathrm{N}(\mu, \sigma)$ finns inritad för ett antal olika värden på σ.

Vi har redan vid flera tillfällen gjort substitutionen $z = (x - \mu)/\sigma$ och då fått en täthetsfunktion som svarar mot att $\mu = 0$ och $\sigma = 1$. En normalfördelad slumpvariabel med dessa μ och σ sägs ha den standardiserade normalfördelningen

DEFINITION 3.26 (STANDARDISERAD NORMALFÖRDELNING)

En normalfördelad slumpvariabel Z med parametrar $\mu = 0$ och $\sigma = 1$ sägs vara *standardiserad normalfördelad*, $Z \sim \mathrm{N}(0, 1)$. Täthetsfunktionen $f_Z(z)$ och fördelningsfunktionen $F_Z(z)$ för en standardiserad normal-

Bild saknas

Figur 3.21. *Täthetsfunktionen för en N(μ, σ)-fördelning för ett antal olika σ värden.*

fördelning har egna beteckningar $\varphi(z)$ respektive $\Phi(z)$. Dessa definieras således av

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < x < \infty,$$
$$\Phi(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

I Figur 3.22 visas täthetsfunktionen $\varphi(x)$ för N(0, 1) och i Figur 3.23 dess fördelningsfunktion $\Phi(x)$. Att beräkna $\Phi(x)$ för ett givet värde x måste göras genom numerisk integration och kräver därför matematisk programvara. Av den anledningen finns $\Phi(x)$ tabulerad längst bak i boken (Tabell 4) vilket vi kommer att ha användning av om vi ska beräkna sannolikheter för utfall av normalfördelade slumpvariabler. Att dessa funktioner har fått egna beteckningar beror på att de används ofta. I själva verket är det så att det räcker att kunna räkna ut dessa för att kunna beräkna $F_X(x)$ för en godtycklig normalfördelning vilket vi nu skall se.

Låt $X \sim \mathrm{N}(\mu, \sigma)$ och $Z \sim \mathrm{N}(0, 1)$. Vi har tidigare visat att

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Genom variabelsubstitutionen $z = (x - \mu)/\sigma$ får vi att detta kan skrivas som

$$F_X(x) = \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Bild saknas

Figur 3.22. *Täthetsfunktionen $\varphi(x)$ för en $N(0,1)$-fördelning (standardiserad normalfördelning).*

Bild saknas

Figur 3.23. *Fördelningsfunktionen $\Phi(x)$ för en $N(0,1)$-fördelning (standardiserad normalfördelning).*

Eftersom täthetsfunktionen av en slumpvariabel är derivatan av fördelningsfunktionen får vi även att

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right).$$

Vi formulerar detta viktiga och användbara resultat i en sats.

SATS 3.22

Om $X \sim \mathrm{N}(\mu, \sigma)$ så gäller

$$f_X(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right),$$
$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right),$$
$$P(a < X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

ANMÄRKNING 3.27

Notera att sannolikheterna ovan kan skrivas med eller utan strikta olikheter. För alla kontinuerliga slumpvariabler gäller nämligen t.ex. att $P(a < X < b) = P(a \leq X \leq b)$.

Vi visar nu genom ett par exempel användbarheten av resultaten ovan. Vi använder oss då av Tabell 4 på sidan ?? som ger $\Phi(z)$ för positiva z i hundradelsformat upp till 4.09.

EXEMPEL 3.36

Antag att vikten på förvildade svenska minkar är normalfördelad med väntevärde μ=4.7 kg och standardavvikelse σ=1.2 kg. Täthetsfunktionen blir således

$$f_X(x) = \frac{1}{1.2\sqrt{2\pi}} e^{-(x-4.7)^2/(2\cdot 1.2^2)}.$$

T.ex. blir täthetsfunktionen i punkten x=4.0 kg $f_X(4.0) = (1.2\sqrt{2\pi})^{-1} e^{-(4-4.7)^2/(2\cdot 1.2^2)} = 0.2804$. Väntevärdet är alltså μ=4.7 kg, och standardavvikelsen är σ =1.2 kg. Fördelningsfunktionen $F_X(x)$ ges av $\Phi\left(\frac{x-4.7}{1.2}\right)$. Om vi t.ex. vill beräkna sannolikheten att en slumpvis vald mink väger mindre är 5 kg ges den av $F_X(5) = \Phi\left(\frac{5-4.7}{1.2}\right) = \Phi(0.25)$. Eftersom $\Phi(\cdot)$ finns tabulerad i Tabell 4 kan vi avläsa att eftersökt sannolikhet blir 0.5987. Detta får man genom att gå ner i tabellen tills man stöter på rad 0.2. Sedan går man åt höger till hundradelen 0.05 där man avläser $\Phi(0.25) = 0.5987$. Om vi vill beräkna sannolikheten att vikten understiger 4 kg blir denna $F_X(4) = \Phi\left(\frac{4-4.7}{1.2}\right) = \Phi(-0.58)$. Vi kan nu inte använda Tabell 4 direkt eftersom den endast finns för positiva x. Emellertid är φ symmetrisk kring 0 av vilket det följer att $\Phi(-x) = 1 - \Phi(x)$

vilket även illustreras i Figur 3.24. Låt oss avslutningsvis beräkna sannolikheten att en slumpvis vald mink väger mer än 3.5 kg. Denna ges av

$$\begin{aligned} P(X > 3.5) &= 1 - P(X \leq 3.5) = 1 - F_X(3.5) \\ &= 1 - \Phi\left(\frac{3.5 - 4.7}{1.2}\right) = 1 - \Phi(-1) = \Phi(1) = 0.8413. \end{aligned}$$

Bild saknas

Figur 3.24. *Illustration av att* $\Phi(-x) = 1 - \Phi(x)$.

EXEMPEL 3.37

Vid en industri produceras järnbalkar som väger 2000 kg. Den exakta vikten är förstås inte 2000.00000 kg. I stället kan man för dessa balkar anta att vikten för en enskild balk med god approximation kan beskrivas som $X \sim \mathrm{N}(\mu = 2000, \sigma = 2.3\ kg)$. Om balkar som avviker med mer än 5kg från avsedd vikt efterbehandlas för att få en vikt närmre 2000 kg betyder det att andelen balkar som behöver efterbehandlas ges av

$$\begin{aligned} P(X < 1995) + P(X > 2005) &= 1 - P(1995 < X < 2005) \\ &= 1 - \left(\Phi\left(\frac{2005 - 2000}{2.3}\right) - \Phi\left(\frac{1995 - 2000}{2.3}\right)\right) \\ &= 1 - (\Phi(2.17) - \Phi(-2.17)) = 2(1 - \Phi(2.17)) = 0.030. \end{aligned}$$

Sannolikheten att en balk väger mindre än 1997 kg ges av

$$\begin{aligned} P(X \leq 1997) = F_X(1997) &= \Phi(\frac{1997 - 2000}{2.3}) = \Phi(-1.30) \\ &= 1 - \Phi(1.30) = 0.0968. \end{aligned}$$

På motsvarande sätt får man att sannolikheten att balken väger mer än 2007 kg ges av $P(X > 2007) = 1 - F_X(2007) = 1 - \Phi(\frac{2007-2000}{2.3}) = 0.0012$.

I Sats 3.22 preciserades hur man uttrycker sannolikheter av typen $P(X < a)$, $P(X > b$ och $P(a < X < b)$ i termer av $\Phi(x)$, fördelningsfunktionen för N(0, 1) (standardiserad normalfördelning). Ibland vill man ta reda på sannolikheten för händelser där en- eller tvåsidiga intervallet anges i termer av μ och σ. T.ex. kanske man vill veta vad sannolikheten är att den normalfördelade slumpvariabeln avviker mer än en standardavvikelse från sitt väntevärde. Man vill alltså beräkna $P(X < \mu-\sigma)+P(X > \mu+\sigma)$. Enligt formlerna från Sats 3.22 blir detta

$$\Phi\left(\frac{\mu-\sigma-\mu}{\sigma}\right) + 1 - \Phi\left(\frac{\mu+\sigma-\mu}{\sigma}\right) = \Phi(-1) + 1 - \Phi(1)$$
$$= 2(1 - \Phi(1)) = 0.3174.$$

Som synes beror inte svaret på μ eller σ. Sannolikheten för avvikelser från väntevärdet μ, uttryckt i enheter av standardavvikelsen σ, är således oberoende av μ och σ. Man kan alltså kan tala om sådana händelser generellt. Det gäller rent allmänt för $a < b$ att

$$P(\mu + a\sigma \leq X \leq \mu + b\sigma) = \Phi(b) - \Phi(a).$$

Speciellt gäller att chansen att slumpvariabeln håller sig inom 1, 2 respektive 3 standaravvikelser från väntevärdet blir

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.6826,$$
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9544,$$
$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.9974.$$

Chansen att ligga utanför dessa intervall blir komplementsannolikheten, dvs. 1 minus sannolikheterna ovan. Det är alltså inte särskilt ovanligt (0.3174) att hamna mer än en standardavvikelse från väntevärdet. Att hamna mer än två standardavvikelser ifrån väntevärdet är däremot ganska ovanligt (0.0456). Sannolikheten att avvika med mer än tre standardavvikelser från väntevärdet är mycket ovanligt (0.0026; mindre än 3 promille). Detta förklarar det gamla uttrycket att någon är ”mer än tre sigma” om udda personer.

Tidigare i detta kapitel definierades α-kvantilen x_α som det värde x som gör att $P(X > x_\alpha) = \alpha$. Kvantiler för standard normalfördelningen förekommer så ofta så att dessa getts en egen beteckning λ_α. Den som läser

inferenskapitlen i denna eller annan bok kommer med största sannolikhet lära sig $\lambda_{0.025}$, och troligen ytterligare någon kvantil, utantill. Om vi låter $Z \sim \mathrm{N}(0,1)$ så definieras således λ_α som lösningen till $P(Z > \lambda_\alpha) = \alpha$. Men $P(Z > \lambda_\alpha) = 1 - \Phi(\lambda_\alpha)$ så λ_α löser tydligen $\Phi(\lambda_\alpha) = 1 - \alpha$. Tag t.ex. $\alpha = 0.05$. Vi skall då söka i normalfördelningstabellen efter det tal x som gör att $\Phi(x) = 1 - \alpha = 0.95$. Efter en stunds tittande ser man att svaret $\lambda_{0.05}$ bör ligga mellan 1.64 och 1.65, troligen ganska mitt emellan dessa två värden. För att få bättre precision finns emellertid en speciell tabell, Tabell 4, som ger kvantilvärden för de vanligast förekommande α-värdena. I denna tabell ser man att $\lambda_{0.05} = 1.6449$.

I statistiska sammanhang vill man ofta ange symmetriska intervall så att standardnormalfördelningen ligger inom intervallet med en fördefinierad sannolikhet. Om vi t.ex. vill veta vilket tal $z > 0$ som gör att $P(-z \leq Z \leq z) = 0.95$ kan även detta uttryckas med hjälp av kvantiler. Om vi vill att 95% av sannolikhetsmassan ska ligga innanför det symmetriska intervallet $[-z,\ z]$ inser man att 2.5% måste ligga till vänster därom och 2.5% till höger (se Figur 3.25 för en illustration av det allmänna fallet). Således måste z ges av

Bild saknas

Figur 3.25. *I figuren finns det symmetriska intervallet innehållande sannolikhetsmassa* $1 - \alpha$ *skuggat, samt kvantilen* $\lambda_{\alpha/2}$ *inritad.*

$\lambda_{0.025} = 1.9600$ och $-z$ blir då $-\lambda_{0.025} = -1.9600$.

Uttryckt annorlunda är alltså sannolikheten 0.95 att en normalfördelad slumpvariabel inte avviker mer än 1.96 standardavvikelser från sitt väntevärde. Detta stämmer väl överens med vad vi tidigare visat, nämligen att sannolikheten att avvika mer än 2 standardavvikelser var 0.9544.

Vi har alltså kommit fram till följande allmänna formel

$$P(-\lambda_{\alpha/2} \leq Z \leq \lambda_{\alpha/2}) = 1 - \alpha.$$

ÖVNING 3.56

Låt $Z \sim \mathrm{N}(0, 1)$. Beräkna
a) $P(Z < 1.5)$,
b) $P(1.0 < Z < 1.5)$,
c) $P(Z > 0.5)$.

ÖVNING 3.57

Låt $Z \sim \mathrm{N}(0, 1)$. Beräkna
a) $P(Z < -0.3)$,
b) $P(Z > -0.59)$,
c) $P(-1.2 < Z < 1.5)$.

ÖVNING 3.58

Låt $X \sim \mathrm{N}(\mu = 10, \sigma = 2)$. Beräkna
a) $P(X < 11)$,
b) $P(10 < X < 15)$,
c) $P(X > 13.5)$.

ÖVNING 3.59

Låt $X \sim \mathrm{N}(\mu = 10, \sigma = 2)$. Beräkna
a) $P(X < 7)$,
b) $P(X > 6.5)$,
c) $P(8.5 < X < 11)$,

ÖVNING 3.60

Låt $X \sim \mathrm{N}(\mu = 10, \sigma = 2)$. Beräkna
a) $P(10 + 1.2 \cdot 2 \leq X)$,
b) $P(10 - 0.68 \cdot 2 > X)$,
c) $P(10 - 1.28 \cdot 2 \leq X \leq 10 - 0.34 \cdot 2)$,

ÖVNING 3.61

Låt $X \sim \mathrm{N}(\mu, \sigma)$ och låt $a > 0$ och $b > 0$ vara givna. Uttryck följande sannolikheter i termer av $\Phi(x)$ för *positiva* x.

a) $P(X > -a)$,

b) $P(-a < X < a)$,

c) $P(-a < X < b)$,

ÖVNING 3.62

Låt $X \sim \mathrm{N}(\mu, \sigma)$. Uttryck följande sannolikheter i termer av $\Phi(x)$. Låt $0 < a < b$ vara givna.

a) $P(\mu - a\sigma \leq X \leq \mu a\sigma)$,

b) $P(\mu \leq X \leq \mu b\sigma)$,

c) $P(\mu - b\sigma < X < \mu - a\sigma)$,

ÖVNING 3.63

Beräkna följande standardiserade normalfördelningskvartiler approximativt. (L)

a) $\lambda_{0.32}$,

b) $\lambda_{0.07}$,

c) $\lambda_{0.02}$

ÖVNING 3.64

Låt $X \sim \mathrm{N}(\mu = 100,\ \sigma = 10)$. Beräkna följande kvantiler x_α. (L)

a) $x_{0.05}$,

b) $x_{0.025}$,

c) $x_{0.005}$.

ÖVNING 3.65

Låt $X \sim \mathrm{N}(\mu,\ \sigma)$. Uttryck kvantilen x_α med hjälp av kvantilen λ_α för den standardiserade normalfördelningen.

ÖVNING 3.66

Antag att svenska kvinnors längd är normalfördelad med väntevärde 167 cm och standardavvikelse 5 cm. En kvinna väljs på måfå. Vad är sannolikheten att

a) hon är längre än 175 cm,
b) hon är kortare än 165 cm,
c) hennes längd är mellan 160 cm och 170 cm.

ÖVNING 3.67

Antalet DNA-substitutioner som skett hos en familj av växtarter under de senaste 50 miljoner åren beskrivs väl med normalfördelningen med väntevärde 50 och samma varians, dvs. med standardavvikelse $\sqrt{50} \approx 7.07$. (Exakt gäller dock inte detta eftersom antal substitutioner är måste vara ett heltal vilket inte gäller för normalfördelningen – vi ”glömmer” dock denna skillnad.). Låt Y vara antalet substitutioner för en given väst. Vad är sannolikheten att

a) högst 60 substitutioner skett,
b) högst 70 substitutioner skett,
c) högst 80 substitutioner skett.

ÖVNING 3.68

Antalet malariaparasiter per ml blod hos en malariasmittad individ är normalfördelad med väntevärde 3200 och standardavvikelse 1000, barn har samma standardavvikelse men väntevärde 4000 (påhittade uppgifter). Kraftig feber brukar inträffa när parasitnivån överstiger 5000. Hur stor del av barnen bör få kraftig feber respektive hur stor andel bland vuxna?

ÖVNING 3.69

Senare i boken (sidan 149) kommer vi definiera vad som kallas momentgenererande funktion för en slumpvariabel och som definieras som $\phi_X(t) = E(e^{tX})$. Beräkna momentgenererande funktionen för $X \sim \mathrm{N}(\mu, \sigma)$. (L)

3.8.4 *Fler kontinuerliga fördelningar

Lognormal fördelning
Gammafördelning
Betafördelning
t-, F- och χ^2fördelningarna
Weibullfördelning
Paretofördelning

3.9 Flerdimensionella slumpvariabler

I det här avsnittet kommer vi behandla slumpvariabler med mer än en dimension. Vi kommer dock nästan uteslutande exemplifiera med 2-dimensionella slumpvariabler, men att generalisera till högre dimensioner är inte svårt.

3.9.1 Definition av flerdimensionella slumpvariabler

Följande definitioner för en tvådimensionell slumpvariabel och dess fördelningsfunktion är helt analoga med deras endimensionella motsvarigheter.

DEFINITION 3.27 (TVÅDIMENSIONELL SLUMPVARIABEL)

En *tvådimensionell slumpvariabel* $(X, Y) = (X(u), Y(u))$ är en tvådimensionell funktion definierad på ett utfallsrum Ω, $(X, Y) : \Omega \longmapsto \mathcal{R} \times \mathcal{R}$.

ANMÄRKNING 3.28

Flerdimensionell slumpvariabel definieras helt analogt, $(X_1, \ldots, X_n) : \Omega \longmapsto \mathcal{R}^n$

DEFINITION 3.28 (FÖRDELNINGSFUNKTION FÖR TVÅDIMENSIONELL SLUMPVARIABEL)

Fördelningsfunktionen $F_{X,Y}(x, y)$ för en tvådimensionell slumpvariabel definieras som $F_{X,Y}(x, y) = P(X \leq x,\ Y \leq y)$.

ANMÄRKNING 3.29

Fördelningsfunktionen för en flerdimensionell slumpvariabel definieras helt analogt som $F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots X_n \leq x_n)$

På liknande sätt som för endimensionella slumpvariabler finns diskreta och kontinuerliga tvådimensionella slumpvariabler och till dessa hör en sannolikhetsfunktion respektive täthetsfunktion.

DEFINITION 3.29 (DISKRET TVÅDIMENSIONELL SLUMPVARIABEL OCH DESS SANNOLIKHETSFUNKTION)

En tvådimensionell slumpvariabel (X, Y) sägs vara *diskret* om den endast kan anta ändligt eller uppräkneligt oändligt antal värden. *Sannolikhetsfunktionen* $p_{X,Y}(x, y)$ för en sådan slumpvariabel definieras av

$$p_{X,Y}(j, k) = P(X = j, Y = k) \qquad j = 0, 1, \ldots, k = 0, 1, \ldots.$$

ANMÄRKNING 3.30

Som tidigare antar vi att slumpvariabelns komponenter antar icke-negativa heltal, i annat fall skall summationen ske över de värden komponenterna kan anta.

Sannolikhetsfunktionen hänger ihop med fördelningsfunktionen på samma sätt som i det endimensionella fallet, nämligen att

$$F_{X,Y}(x, y) = \sum_{j=0}^{x} \sum_{k=0}^{y} p_{X,Y}(j, k).$$

Vi illustrerar med en tvådimensionell diskret likformig sannolikhetsfunktion från Exempel 3.38 nedan.

[Bild saknas]

Figur 3.26. *Illustration av den tvådimensionella sannolikhetsfunktionen* $p_{X,Y}(j.k) = 1/36,\ j = 1, \ldots, 6,\ k = 1, \ldots, 6.$

EXEMPEL 3.38

Antag att man skall kasta två tärningar, en röd och en svart, och låt (X, Y) beteckna antalet prickar på den röda respektive svarta tärningen. Vi har tidigare påpekat att av symmetriskäl är alla 36 utfall lika sannolika vilket alltså betyder att $p_{X,Y}(j,k) = 1/36$ för $j = 1, \ldots, 6$ och $k = 1, \ldots, 6$ (och $p_{X,Y}(j,k) = 0$ för övriga j och k). Av detta följer att $\sum_{j=1}^{6} \sum_{k=1}^{6} p_{X,Y}(j,k) = 1$ vilket ju måste gälla. Man kan även efter lite räknande konstatera att motsvarande fördelningsfunktion satisfierar $F_{X,Y}(j,k) = jk/36$ (för $j = 1, \ldots, 6$ och $k = 1, \ldots, 6$).

DEFINITION 3.30 (KONTINUERLIG SLUMPVARIABEL OCH TÄTHETSFUNKTION)

En tvådimensionell slumpvariabel (X, Y) sägs vara *kontinuerlig* om det finns en funktion $f_{X,Y}(x, y)$ så att för ”alla” mängder A gäller

$$P((X, Y) \in A) = \iint_A f_{X,Y}(x, y)\, dxdy.$$

Funktionen $f_{X,Y}(x, y)$ kallas för slumpvariabelns *täthetsfunktion*.

ANMÄRKNING 3.31

Ibland används begreppen simultan *fördelningsfunktion, sannolikhetsfunktion eller täthetsfunktion när man vill betona att funktionen gäller för bägge slumpvariablerna.*

Täthetsfunktionen hänger ihop med fördelningsfunktionen på samma sätt som i det endimensionella fallet, nämligen att

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s, t) dsdt.$$

Vi illustrerar täthetsfunktionen för den tvådimensionella tätheten som ges i Exempel 3.39, se Figur 3.9.1.

EXEMPEL 3.39

Betrakta den tvådimensionella slumpvariabeln (X, Y) med täthetsfunktion $f_{X,Y}(x, y) = x + y$, $0 \leq x \leq 1$, $0 \leq y \leq 1$ (således är $f_{X,Y}(x, y) = 0$ för alla övriga (x, y), se Figur 3.9.1). Även här ser vi

[Bild saknas]

Figur 3.27. *Illustration av den tvådimensionella täthetsfunktionen* $f_{X,Y}(x, y) = x + y$, $0 \leq x \leq 1,\ 0 \leq y \leq 1$.

att $\int\int f_{X,Y}(x, y)dxdy = 1$ som det ska. För $A = [0,\ 0.5] \times [0,\ 0.5]$ får vi $P((X, Y) \in A) = \int_0^{0.5}\int_0^{0.5}(x + y)dxdy = 1/8$ samt för $B = [0.5,\ 1] \times [0.5,\ 1]$ får vi $P((X, Y) \in B) = \int_{0.5}^{1}\int_{0.5}^{1}(x+y)dxdy = 3/8$. Område B har således tre gånger så hög sannolikhet som A trots att områdena är lika stora.

För en tvådimensionell slumpvariabel kan man ibland primärt vara intresserad av den ena komponenten. Man kan då studera den marginella fördelningen. Vi har följande samband mellan en tvådimensionell slumpvariabel och dess ena komponent.

SATS 3.23 (MARGINELL FÖRDELNING)

Betrakta en tvådimensionell slumpvariabel (X, Y) med fördelningsfunktion $F_{X,Y}(x, y)$. Den *marginella fördelningsfunktionen* för slumpvariabeln X ges då av $F_X(x) = F_{X,Y}(x, \infty)$. Om (X, Y) är diskret ges den marginella sannolikhetsfunktionen för X av $p_X(j) = \sum_k p_{X,Y}(j, k)$. Om (X, Y) är kontinuerlig ges den marginella täthetsfunktionen för X av $f_X(x) = \int f_{X,Y}(x, y)dy$.

ANMÄRKNING 3.32

Den marginella fördelningen för Y fås på motsvarande sätt genom att byta plats på variablerna ovan.

BEVIS

Vi visar de två första påståendena, det tredje överlåts till läsaren (Övning 3.73). Det gäller att $F_X(x) = P(X \leq x) = P(X \leq x, Y < \infty) = F_{X,Y}(x, \infty)$. På motsvarande sätt gäller

$$\begin{aligned} p_X(j) &= P(X = j) = P(X = j, -\infty < Y < \infty) \\ &= \sum_k P(X = j, Y = k) = \sum_k p_{X,Y}(j, k), \end{aligned}$$

där den tredje likheten följer från Axiom 3 i Kolmogorovs axiomsystem 2.2 på sidan 6.

3.9.2 Kovarians och korrelation

När man behandlar tvådimensionella slumpvariabler är det viktigt hur stora de två komponenterna är (vilket kan kvantifieras med hjälp av de marginella fördelningarna för X respektive Y). En annan viktig aspekt är huruvida de två variablerna påverkar varandra, man brukar tala om beroenden mellan variabler. Hur beroende två variabler är kan kvantifieras på olika sätt, de vanligaste endimensionella måtten är kovarians och korrelationskoefficient vilka vi nu definierar.

DEFINITION 3.31 (KOVARIANS OCH KORRELATIONSKOEFFICIENT)

Betrakta en tvådimensionell slumpvariabel (X, Y) med ändliga väntevärden (μ_X och μ_Y) och standardavvikelser (σ_X och σ_Y). *Kovariansen* mellan X och Y, betecknat $C(X, Y)$, definieras som $C(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$. För en diskret respektive kontinuerlig slumpvariabel innebär dessa definitioner således

$$C(X, Y) = \sum_{j,k}((j - \mu_X)(k - \mu_Y)p_{X,Y}(j, k), \text{ respektive}$$

$$C(X, Y) = \int\int((x - \mu_X)(y - \mu_Y)f_{X,Y}(x, y)dxdy.$$

Korrelationskoefficienten $\rho = \rho(X, Y)$ definieras som

$$\rho(X, Y) = \frac{C(X, Y)}{\sigma_X \sigma_Y}.$$

ANMÄRKNING 3.33

Kovarians förkortas med C eftersom det på engelska heter covariance.

ANMÄRKNING 3.34

Såväl kovarians som korrelationskoefficient är reella tal. Enheten för kovarians är produkten av enheterna som X och Y anges i. Korrelationskoefficienten är däremot dimensionslös vilket gör den mer användbar: om vi byter måttenhet för variablerna ändras kovariansen men inte korrelationskoefficienten.

Ofta kan följande räkneregel, som liknar räkneregeln för beräkning av varians från Sats 3.5 på sidan 64, vara nyttig att använda vid beräkning av kovarians.

SATS 3.24

För en tvådimensionell slumpvariabel med kovarians $C(X, Y)$ och marginella väntevärden μ_X respektive μ_Y gäller

$$C(X, Y) = E(XY) - \mu_X \mu_Y.$$

BEVIS

Vi visar satsen för det kontinuerliga fallet. Det gäller nämligen att

$$\begin{aligned}
C(X, Y) &= \int\int (x - \mu_X)(y - \mu_Y) f_{X,Y}(x, y) dxdy \\
&= \int\int (xy - x\mu_Y - \mu_X y - \mu_X \mu_Y)\, f_{X,Y}(x, y) dxdy \\
&= \int\int xy f_{X,Y}(x, y) dxdy - \mu_Y \int\int x f_{X,Y}(x, y) dxdy \\
&\quad - \mu_X \int\int y f_{X,Y}(x, y) dxdy + \mu_X \mu_Y \int\int f_{X,Y}(x, y) dxdy \\
&= \int\int xy f_{X,Y}(x, y) dxdy - \mu_Y \int x f_X(x) dx \\
&\quad - \mu_X \int y f_Y(y) dy + \mu_X \mu_Y \\
&= \int\int xy f_{X,Y}(x, y) dxdy - \mu_Y E(X) - \mu_X E(Y) + \mu_X \mu_Y \\
&= E(XY) - \mu_X \mu_Y.
\end{aligned}$$

Vi nämnde som en anmärkning att korrelationskoefficienten är dimensionslös. Man kan visa följande starkare resultat:

SATS 3.25

Korrelationskoefficienten $\rho = \rho(X, Y)$ för en tvådimensionell slumpvariabel satisfierar

$$-1 \leq \rho \leq 1.$$

BEVIS

Eftersom $((x - \mu_X)/\sigma_X + (y - \mu_Y)/\sigma_Y)^2$ är en positiv funktion gäller att

$$E[\left((X - \mu_X)/\sigma_X^2 + (Y - \mu_Y)/\sigma_Y^2\right)^2] \geq 0.$$

Om vi utvecklar kvadraten blir vänster sida lika med

$$V(X)/\sigma_X^2 + V(Y)/\sigma_Y^2 + 2\rho(X, Y) = 2 + 2\rho(X, Y).$$

Detta är positivt bara om $\rho \geq -1$. Att $\rho \leq 1$ visas analogt men för funktionen $((x - \mu_X)/\sigma_X - (y - \mu_Y)/\sigma_Y)^2$.

I beviset ovan använder vi att $E((X - \mu_X)^2/a)) = E((X - \mu_X)^2)/a$. Detta är en konsekvens av att väntevärdet är en integral eller summa, och konstanter kan brytas ut ur dessa. Ett formellt bevis görs längre fram i Sats 3.33.

Om $\rho(X, Y) > 0$ säger man att X och Y är *positivt* korrelerade medan X och Y sägs vara *negativt* korrelerade om $\rho(X, Y) < 0$. Att två variabler är positivt korrelerade betyder att om den ena slumpvariabeln är stor så tenderar även den andra att vara det, liksom att om med ena är liten tenderar den andra också att vara det. Vid negativ korrelation tenderar den ena att vara liten när den andra är stor. Om $|\rho|$ är nära 1 säger man att korrelationen är stark medan den sägs vara svag om $|\rho|$ ligger nära 0. Man kan visa att $|\rho| = 1$ om och endast om Y kan skrivas som $Y = aX + b$ för några konstanter a och b. Att verkligen $|\rho| = 1$ för detta val visas i Övning 3.85 på sidan 156.

3.9.3 Oberoende slumpvariabler

Vi definierar nu de två viktiga begreppen okorrelerade och oberoende. Dessa är viktiga för att alla räkningar blir mycket enklare när dessa kriterier är uppfyllda, och inte så sällan är de uppfyllda i tillämpningar.

DEFINITION 3.32 (OBEROENDE OCH OKORRELERADE)

Två slumpvariabler X och Y sägs vara *oberoende* om deras simultana fördelning satisfierar $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ för det diskreta fallet, respektive att $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ för det kontinuerliga fallet. Slumvariablerna sägs vara *okorrelerade* om $\rho(X, Y) = 0$.

ANMÄRKNING 3.35

En alternativ mer allmängiltig definition av oberoende är att X och Y är oberoende om deras fördelningsfunktion satisfierar $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.

Oberoende är ett starkare kriterium än okorrelerad vilket följande sats visar.

SATS 3.26

Om X och Y är oberoende så är de även okorrelerade.

BEVIS

Vi visar satsen för det kontinuerliga fallet. Från antagandet har vi

$$E(XY) = \int\int xyf_{X,Y}(x, y)dxdy = \int\int xyf_X(x)f_Y(y)dxdy$$
$$= \int xf_X(x)dx \int yf_Y(y)dy = E(X)E(Y).$$

Från Sats 3.24 gäller därför att $C(X, Y) = 0$ vilket i sin tur medför att $\rho(X, Y) = 0$ vilket skulle visas.

3.9.4 Betingade fördelningar

Ibland har man observerat den ena av de två slumpvariablerna och man är intresserad av den andra slumpvariabeln givet (eller betingat av) vetskapen

om den första variabelns värde. Vi har då att göra med betingade fördelningar som är nära besläktade med betingade sannolikheter som definierades i Avsnitt 2.5 på sidan 19.

DEFINITION 3.33 (BETINGAD FÖRDELNING)

Låt (X, Y) vara en tvådimensionell slumpvariabel. Om (X, Y) är diskret så definieras den *betingade sannolikhetsfunktion* $p_{X|Y}(j \mid k)$ för X, betingat av att $Y = k$, av

$$p_{X|Y}(j \mid k) = P(X = j \mid Y = k) = \frac{p_{X,Y}(j, k)}{p_Y(k)}, \ j = 0, 1, \ldots.$$

Om (X, Y) är kontinuerlig definieras den *betingade täthetsfunktion* $f_{X|Y}(x \mid y)$ för X, betingat av att $Y = y$, av

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad -\infty < x < \infty.$$

Den *betingade fördelningsfunktionen.* $F_{X|Y}(x \mid y)$ för X, betingat av att $Y = y$, fås på vanligt sätt genom summering/integrering: $F_{X|Y}(x \mid y) = \sum_{j \le x} p_{X|Y}(j \mid y)$ respektive $F_{X|Y}(x \mid y) = \int_{-\infty}^{x} f_{X|Y}(t \mid y)dt$.

ANMÄRKNING 3.36

Betingade fördelningar uppfyller kraven för en fördelning. Att så verkligen är fallet kan lätt visas i det diskreta fallet genom att visa att fördelnings- och sannolikhetsfunktionerna ovan uppfyller definitionerna för fördelnings- och sannolikhetsfunktion (Definition 3.28 och 3.29). Beviset är däremot krångligare i det kontinuerliga fallet beroende på att betingningshändelsen då har sannolikhet 0 och hoppas därför över.

I Avsnitt 2.5 infördes två ekvivalenta definitioner för oberoende *händelser*. Den ena var att $P(A \cap B) = P(A)P(B)$ vilket påminner om Definition 3.32 ovan för oberoende slumpvariabler. Den andra var att $P(A \mid B) = P(A \cap B)/P(B)$. Även denna har en motsvarighet för slumpvariabler.

DEFINITION 3.34 (ALTERNATIV DEFINITION AV OBEROENDE SLUMPVARIABLER)

Två diskreta slumpvariabler X och Y är *oberoende* om $p_{X|Y}(j \mid k) = p_X(j)$ för alla j och sådana k för vilket $p_Y(k) > 0$. Två kontinuerliga

slumpvariabler X och Y är *oberoende* om $f_{X|Y}(x \mid y) = f_X(x)$ för alla x och sådana y för vilket $f_Y(y) > 0$.

Att de två definitionerna är ekvivalenta visas lätt. T.ex. gäller ju att om $p_{X|Y}(j \mid k) = p_X(j)$ får man att från definitionen av betingad sannolikhetsfunktion (Definition 3.29) $p_{X,Y}(j,k) = p_{X|Y}(j \mid k)p_Y(k) = p_X(j)p_Y(k)$ vilket medför oberoende enligt den ursprungliga definitionen.

Det påpekades ovan att betingade fördelningar verkligen är fördelningar i sig. Speciellt kan man beräkna väntevärden, varianser och standardavvikelser på samma sätt som för ”vanliga” fördelningar. Man brukar använda beteckningarna $E(X|Y = y)$, $V(X \mid Y = y)$ och $D(X \mid Y = y)$ vilka alltså beräknas enligt följande formler:

$$
\begin{aligned}
E(X|Y = k) &= \sum_j j p_{X|Y}(j \mid k) \\
V(X \mid Y = k) &= \sum_j j^2 p_{X|Y}(j \mid k) - (E(X|Y = k))^2 \\
D(X \mid Y = k) &= \sqrt{V(X \mid Y = k)},
\end{aligned}
$$

för det diskreta fallet och motsvarande integraler för det kontinuerliga fallet.

EXEMPEL 3.40

Antag att vi ska kasta två tärningar och är intresserad av summan S av antalet prickar. Låt X ange antalet prickar på den svarta och Y antal prickar på den vita tärningen. Vi ska nu studera fördelningen för summan av antalet prickar betingat av att vi vet antalet prickar $X = k$ på den svarta tärningen. Man kan räkna fram att $p_{S|X}(s \mid k) = 1/6$, $s = k+1, \ldots, k+6$, eftersom den vita tärningen blir 1 till 6 med sannolikhet 1/6 vardera. För väntevärdet får vi

$$
E(S \mid X = k) = \sum_{s=k+1}^{k+6} s p_{S|X}(s \mid k) = ((k+1)+\ldots+(k+6))/6 = k+\frac{7}{2}.
$$

Det gäller vidare att $E(S^2 \mid X = k) = ((k+1)^2 + \ldots + (k+6)^2)/6$ och efter lite räknande får man att $V(S \mid X = k) = 35/12$.

ÖVNING 3.70

Betrakta den diskreta tvådimensionella slumpvariabeln $p_{X,Y}(j,k) = c(j+k)$, $j = 1,2,3$ och $k = 1,2,3$ för något c. Bestäm c.

a) Bestäm c.

b) Beräkna $E(X)$, $V(X)$, $C(X,Y)$.

c) Beräkna $E(X|Y=3)$.

ÖVNING 3.71

I Exempel 3.39 på sidan 121 beskrevs en kontinuerlig tvådimensionell fördelning med täthetsfunktion $f_{X,Y}(x,y) = x+y$, $0 \leq x \leq 1,\ 0 \leq y \leq 1$. Beräkna $E(X)$, $E(Y)$, $V(X)$, $V(Y)$, $C(X,Y)$ och $\rho(X,Y)$.

ÖVNING 3.72

Betrakta igen en kontinuerlig tvådimensionell slumpvariabel med täthetsfunktion $f_{X,Y}(x,y) = x+y$, $0 \leq x \leq 1,\ 0 \leq y \leq 1$.

a) Beräkna $f_{X|Y}(x \mid y)$.

b) Beräkna $E(X \mid Y = y)$.

b) För vilket y är denna störst/minst?

ÖVNING 3.73

Bevisa det tredje påståendet i Sats 3.23, dvs. att den marginella täthetsfunktionen för X ges av $f_X(x) = \int_0^\infty f_{X,Y}(x,y)dy$. (L)

ÖVNING 3.74

Låt (X,Y) vara en tvådimensionell diskret slumpvariabel. Visa att om deras fördelningsfunktion $F_{X,Y}(x,y)$ satisfierar $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ för alla x och y så är X och Y oberoende. (L)

3.10 Några vanliga flerdimensionella fördelningar

Vi ska i detta avsnitt gå igenom två vanligt förekommande flerdimensionella fördelningar som är generaliseringar av deras endimensionella motsvarigheter. Vi börjar med den diskreta multinomialfördelningen och tar därefter upp den tvådimensionella normalfördelningen.

3.10.1 Multinomialfördelning

Binomialfördelningen uppstod i situationer där ett slumpförsök som kunde utfalla på två sätt upprepades n gånger (se sidan 81). Det förekommer också ofta situationer där slumpförsök kan utfalla på fler än två sätt, och i denna situation uppstår multinomialfördelningen.

EXEMPEL 3.41

En tärning kastas 10 gånger. Antag att vi vill beräkna sannolikheten att det resulterade i 2 st 1:or, 1 2:a, 1 3:a, 2 st 4:or, 1 5:a och 3 st 6:or. I detta fall har alla sex utfall samma sannolikhet (vilket inte alltid behöver vara fallet). Chansen för en given sekvens av 10 kast är således $(1/6)^10$ för varje sekvens. För att få det utfall vi är intresserade av kan vi dock välja flera olika sekvenser. Man kan nämligen visa en generalisering av situationen då binomialkoefficienterna uppstod, som säger att antal sätt att välja 2 1:or, 1 2:a osv, är lika med $\frac{10!}{2!1!1!2!1!3!} = 151200$. Detta produktbråk brukar ofta skrivas som $\binom{10}{2\ 1\ 1\ 2\ 1\ 3}$.Vi får således att den efterfrågade sannolikheten blir $\binom{10}{2\ 1\ 1\ 2\ 1\ 3}(1/6)10 \approx 0.0025$.

Den allmänna multinomialfördelningen är en generalisering av experimentet ovan.

DEFINITION 3.35 (MULTINOMIALFÖRDELNINGEN)

Den k-dimensionella slumpvariabeln $(X_1, X_2, \ldots, X_k)$ sägs vara multinomialfördelad med parametrar n (positivt heltal) och $p_1, \ldots, p_r$, sådana att $p_i > 0$ och $\sum_{i=1}^{r} p_i = 1$, om sannolikhetsfunktionen ges av

$$p_{X_1,\ldots,X_r}(k_1, \ldots, k_r) = \binom{n}{k_1 \ \ldots \ k_r} p_1^{k_1} \cdot \ldots \cdot p_r^{k_r},$$

för sådana $k_1, \ldots, k_r$ så att $k_i \geq 0$ och $\sum_{i=1}^{r} k_i = n$.

ANMÄRKNING 3.37

Egentligen är slumpvariabeln (k-1)-dimensionell eftersom $\sum_{i=1}^{r} X_i = n$ *måste gälla för att ha positiv sannolikhet, så t.ex. den sista koordinaten är överflödig. Specialfallet* $r = 2$ *har ju tidigare studerats, binomialfördelningen, där utfallen kallades "lyckat" och "misslyckat" och då utelämnades antalet misslyckade eftersom summan var given.*

Att detta verkligen är en sannolikhetsfunktion beror på den kombinatoriska likheten

$$(a_1 + \ldots + a_r)^n = \sum_{k_1,\ldots,k_r k_1+\ldots+k_r=n} \binom{n}{k_1 \ \ldots \ k_r} a_1^{k_1} \cdot \ldots \cdot a_r^{k_r},$$

för godtyckliga reella tal $a_1, \ldots, a_r$, och applicerat på $a_i = p_i$ gäller $\sum_i a_i = 1$ varför vänster sida blir $1^n = 1$.

Låt oss nu räkna ut väntevärde, varians och kovarianser för multinomialfördelningen. Betrakta specifikt händelsen i $(1 \leq i \leq r)$ som i varje enskilt försök har sannolikhet p_i att inträffa. Vi kan då klumpa ihop de övriga utfallen till en händelse som vi kan kalla "icke-i" (eller "misslyckade") som således har sannolikhet $\sum_{j \neq i} p_j = 1 - p_i$. Således blir X_i binomialfördelad: $X_i \sim \text{Bin}(n, p_i)$, av vilket följer att $E(X_i) = np_i$, $V(X_i) = np_i(1 - p_i)$ och $D(X_i) = \sqrt{np_i(1 - p_i)}$. Vad gäller kovariansen $C(X_i, X_j)$ kan man inte lika lätt resonera sig fram till vad den borde bli. Däremot inser man att den borde vara negativ eftersom att om X_i är ovanligt stor, dvs. att ovanligt många försök resulterat i utfall i, så borde det bli ovanligt få som resulterar i utfall j eftersom antalet försök är givet. Vi visar vad det blir i följande sats.

SATS 3.27

Låt $(X_1, X_2, \ldots, X_k)$ vara multinomialfördelad med parametrar n och $p_1, \ldots, p_r$. Då gäller för $i \neq j$

$$\begin{aligned} E(X_i) &= np_i, \\ V(X_i) &= np_i(1 - p_i), \\ C(X_i, X_j) &= -np_i p_j. \end{aligned}$$

ANMÄRKNING 3.38

Som vanligt får man standardavvikelsen och korrelationen via dessa uttryck, man får då $D(X_i) = \sqrt{np_i(1 - p_i)}$ *respektive* $\rho(X_i, X_j) = -\sqrt{\frac{p_i p_j}{(1-p_i)(1-p_j)}}$.

BEVIS

Vi har ovan bevisat uttrycken för väntevärde och varians genom att konstatera att $X_i \sim \text{Bin}(n, p_i)$. Återstår att visa kovariansen för vilket vi måste beräkna $E(X_iX_j)$. Vi gör endast detta för specialfallet att $r = 3$, det allmänna fallet är mycket likt men mer notationstungt eftersom vi måste hålla reda på alla komponenter. Vi beräknar $E(X_1X_2)$ och skall då summera över alla värden på k_1, k_2 och k_3 sådana att de är icke-negativa och så att $k_1 + k_2 + k_3 = n$. Man kan inse att detta blir exakt de termer som anges nedan:

$$\begin{aligned} E(X_1X_2) &= \sum_{k_1=0}^{n}\sum_{k_2=0}^{n-k_1} k_1k_2\frac{n!}{k_1!k_2!(n-k_1-k_2)!}p_1^{k_1}p_2^{k_2}p_3^{n-k_1-k_2} \\ &= n(n-1)p_1p_2\sum_{k_1=1}^{n}\sum_{k_2=1}^{n-k_1}\frac{(n-2)!}{(k_1-1)!(k_2-1)!(n-k_1-k_2)!} \\ &\quad \times p_1^{k_1-1}p_2^{k_2-1}p_3^{n-k_1-k_2}. \end{aligned}$$

Att nedre summationsgränserna kan ändras till 1 i stället för 0 beror på faktorn k_1k_2. Det som står innanför summationen är nu multinomialsannolikheterna för en multinomialfördelning med parametrar $n-2$ och samma p_1, p_2 och p_3. Detta ser man om man byter index till $j_1 = k_1 - 1$ och $j_2 = k_2 - 1$. Multinomialsannolikheterna summarar sig till 1 varför vi får $E(X_1X_2) = n(n-1)p_1p_2$ så kovariansen blir

$$C(X_1, X_2) = n(n-1)p_1p_2 - np_1np_2 = -np_1p_2.$$

För multinomialfördelning kan man även räkna ut betingade fördelningar. Vi vet att $X_i \sim \text{Bin}(n, p_i)$, men vilken fördelning har X_i betingat på att $X_j = k$ $(0 \le k \le n)$ för någon annan komponent $j \neq i$? Detta kan man härleda rent logiskt, samt även genom formelmanipulation. Vi börjar med det förra eftersom det bidrar mer till förståelsen. Frågan är alltså vilken fördelning X_i har om vi vet att $X_j = k$. Dvs, bland de n försöken vet vi att k resulterade i utfall j och därmed vet vi även resterande $n-k$ *inte* resulterade i utfall j. De kvarvarande försöken har nu fått en förhöjd sannolikhet att resultera i utfall i (liksom för alla övriga "icke-j" utfall. Chansen att ett försök blir utfall i betingat av att det inte blev utfall j är $P(\text{utfall } i|\text{ ej utfall } j) = p_i/(1-p_j)$. De kvarvarande försöken sker fortfarande oberoende. Därför är fördelningen för X_i betingat av att $X_j = k$ binomialfördelad med parametrar $n-k$ och $p_i/(1-p_j)$. Man brukar skriva detta som att $X_i \mid X_j = k \sim \text{Bin}(n-k, p_i/(1-p_j))$.

Sannolikhetsfunktionen blir således

$$p_{X_i|X_j}(j \mid k) = \binom{n-k}{j} \left(\frac{p_i}{1-p_j}\right)^j \left(1 - \frac{p_i}{1-p_j}\right)^{n-k-j},$$

$j = 0, \ldots, n-k$..

Mer allmänt kan man på analogt sätt visa att hela resterande vektorn $X_1, \ldots X_{j-1}, X_{j+1}, \ldots X_r$, betingat av att $X_j = k$ är multinomialfördelad med parametrar $n-k$ och $p_1/(1-p_j), \ldots, p_{j-1}/(1-p_j), p_{j+1}/(1-p_j), \ldots, p_r/(1-p_j)$.

EXEMPEL 3.42

En stryktipsrad består av 13 matcher där man skall tippa 1, x eller 2, beroende på om man tror att hemmalaget vinner, det blir oavgjort, eller bortalaget vinner matchen. Statistiskt sett brukar hemmalagen vinna ca 50% av matcherna, oavgjort blir det ca 30% av matcherna, medan bortalaget vinner ca 20% av matcherna (gissade siffror). Under dessa förutsättningar, och förutsatt att man inte har ytterligare information om enskilda matcher, är antalet hemmavinster, oavgjorda och bortavinster på en stryktipsrad, (Y_1, Y_x, Y_2) då multinomialfördelade med parametrar $n = 13$, $p_1 = 0.5$, $p_x = 0.3$ och $p_2 = 0.2$. T.ex. gäller då att $P(Y_1 = 7, Y_x = 4, Y_2 = 2) = \binom{13}{7\,4\,2} 0.5^7 0.3^4 0.2^2 \approx 0.065$. Det gäller även att $E(Y_1) = 13 \cdot 0.5 = 6.5$, $D(Y_x) = \sqrt{13 \cdot 0.3 \cdot 0.7} \approx 1,65$, och att $C(Y_x, Y_2) = -13 \cdot 0.3 \cdot 0.2 = -0.78$.

3.10.2 Tvådimensionell normalfördelning

Vi definierar nu den tvådimensionella normalfördelningen som ofta kallas bivariat normalfördelning.

DEFINITION 3.36 (BIVARIAT NORMALFÖRDELNING)

Den kontinuerliga två-dimensionella slumpvariabeln (X, Y) sägs vara *bivariat normalfördelad* med parametrar $\mu_x, \mu_y, \sigma_x, \sigma_y$ och ρ ($\sigma_x > 0, \sigma_y > 0$ och $-1 < \rho < 1$) när dess simultana täthetsfunktion $f_{X,Y}(x, y)$ för alla x och y ges av

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp \frac{1}{2(1-\rho^2)} Q_\rho\left(\frac{x-\mu_x}{\sigma_x}, \frac{y-\mu_y}{\sigma_y}\right),$$

där den kvadratiska formen $Q_\rho(u, v)$ definieras av

$$Q_\rho(u, v) = u^2 + v^2 - 2\rho uv.$$

Man skriver att $(X, Y) \sim \mathrm{N}(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$.

ANMÄRKNING 3.39

Att använda sig av täthetsfunktionen för att beräkna sannolikheten för att $(X, Y) \in A$ numeriskt är relativt komplicerat, såvida inte A är en "enkel" mängd. Specialfallet att $\rho = 0$ är dock väsentligt lättare.

Täthetsfunktionen illustreras i Figur 3.28 nedan.

Bild saknas

Figur 3.28. *Bild på 2-dimensionell normalfördelad täthetsfunktion.*

Ett specialfall av denna komplicerade täthetsfunktion är när $\rho = 0$. Produkttermen i den kvadratiska formen försvinner då och man kan skriva tätheten som $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ där $f_X(x)$ och $f_Y(y)$ bägge är täthetsfunktioner för normalfördelning med väntevärde och standardavvikelse μ_x och σ_x respektive μ_y och μ_y. Eftersom den simultana täthetsfunktionen kan skrivas som en produkt av två marginella täthetsfunktioner gäller det alltså att X och Y är oberoende i detta fall (Definition 3.32 på sidan 126).

EXEMPEL 3.43

Antag att $(X, Y) \sim \mathrm{N}(\mu_x = 5, \mu_y = 10, \sigma_x = 2, \sigma_y = 3, \rho = 0)$. Vi kan då t.ex. beräkna $P(3 \leq X \leq 6, 8 \leq Y \leq 14)$. Eftersom X och Y är oberoende gäller att $P(3 \leq X \leq 6, 8 \leq Y \leq 14) = P(3 \leq X \leq 6)P(8 \leq Y \leq 14) = (F_X(6) - F_X(3))(F_Y(14) - F_Y(8))$. Från den endimensionella

normalfördelningen vet vi dessutom att $F_X(x) = \Phi(\frac{x-\mu_x}{\sigma_x})$, där $\Phi(\cdot)$ är fördelningsfunktionen för den standardiserade normalfördelningen. Vi får därför att att eftersökt sannolikhet blir

$$\left(\Phi(\frac{6-5}{2}) - \Phi(\frac{3-5}{2})\right)\left(\Phi(\frac{14-10}{3}) - \Phi(\frac{8-10}{3})\right).$$

Genom att använda Tabell 4 får att eftersökt sannolikhet blir 0.350.

Vad som inte är lika lätt att visa, men som gäller även för fallet att $\rho \neq 0$ är att de marginella fördelningarna för X respektive Y är ovan nämnda normalfördelningar.

SATS 3.28

Antag att $(X, Y) \sim \mathrm{N}(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$. Då gäller att de marginella fördelningarna ges av $X \sim \mathrm{N}(\mu_x, \sigma_x)$ och $Y \sim N(\mu_y, \sigma_y)$.

BEVIS

Det gäller att $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$. Om vi gör substitutionen $z = (y - \mu_y)/\sigma_y$, låter $C = 1/2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}$, och bryter ut faktorer som inte beror av z får vi

$$f_X(x) = C\sigma_y \exp\left(-\frac{1}{2(1-\rho^2)}(\frac{x-\mu_x}{\sigma_x})^2\right) \times \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2(1-\rho^2)}\left[z^2 - 2\rho\frac{x-\mu_x}{\sigma_x}z\right]\right)dz.$$

Om man kvadratkompletterar exponenten innanför integralen får vi

$$z^2 - 2\rho\frac{x-\mu_x}{\sigma_x}z = (z - \rho\frac{x-\mu_x}{\sigma_x})^2 - \rho^2(\frac{x-\mu_x}{\sigma_x})^2.$$

Om man bryter ut den sista termen (med faktorn framför) kan man genom att lägga till lämplig konstant faktor få en endimensionell normaltäthet innanför integralen som således integrerar sig till 1. Utförandet av dessa manipulationer lämnas till läsaren (Övning 3.80). Kvar får man

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x}e^{-(x-\mu_x)^2/2\sigma_x^2},$$

vilket just betyder att $X \sim \mathrm{N}(\mu_x, \sigma_x)$. Beviset för Y är av symmetriskäl helt analogt.

Satsen ovan ger alltså förklaring av innebörden av fyra utav de fem parametrarna, de är väntevärden och standardavvikelser för respektive slumpvariabel. Eftersom den kvarvarande parametern har getts beteckningen σ kan man kanske gissa att den just betecknar korrelationen $\rho(X, Y)$. Detta stämmer vilket följande sats styrker. Beviset för satsen liknar beviset ovan men är mer tekniskt och utelämnas därför.

SATS 3.29

Antag att $(X, Y) \sim \mathrm{N}(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$. Då är $\rho(X, Y) = \rho$, dvs. parametern ρ är korrelationen mellan X och Y.

Vi går igenom ytterligare ett resultat för bivariat normalfördelning, nämligen vad som händer med fördelningen för ena variabeln om vi observerar den andra. Vi vill alltså härleda den betingade fördelningen för X betingat av att $Y = y$. Vi söker därför $f_{X|Y}(x \mid y)$ som är en täthetsfunktion för X. Om vi bakar ihop sådant som inte beror på x får vi från definitionen av betingad täthetsfunktion (Definition 3.33 på sidan 127)

$$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{f_{X,Y}x, y)}{f_Y(y)} \\ &= C_1 f_{X,Y}(x, y) \\ &= C_2 \exp\left(-\frac{1}{2(1-\rho^2)}[(\frac{x-\mu_x}{\sigma_x})^2 - 2\rho\frac{x(y-\mu_y)}{\sigma_x\sigma_y}]\right) \\ &= C_3 \exp\left(-\frac{1}{2\sigma_x^2(1-\rho^2)}[x - (\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y))]^2\right). \end{aligned}$$

Som funktion av x har vi alltså att $f_{X|Y}(x \mid y) = C_3 e^{-(x-a)^2/2b^2}$, där $a = \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y)$ och $b = \sigma_x\sqrt{1-\rho^2}$. En sådan täthet känner vi ju igen som den endimensionella normalfördelningstätheten. Vi behöver ju inte bry oss om C_3 eftersom denna måste göra så att vi får 1 om vi integrerar tätheten över hela reella axeln. Vi har således visat följande sats.

SATS 3.30

Antag att $(X, Y) \sim \mathrm{N}(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$. Då är den betingade fördelningen för X givet att $Y = y$ normalfördelad med väntevärde $\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y)$ och standardavvikelse $\sigma_x\sqrt{1-\rho^2}$, dvs.

$$X \mid Y = y \sim \mathrm{N}\left(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y),\ \sigma_x\sqrt{1-\rho^2}\right).$$

I Figur 3.29 nedan illustreras satsens resultat grafiskt. Den betingade fördelningen för X betingat av att $Y = y$ fås genom att betrakta den endimensionella fördelningen i x-led för fixt y.

Bild saknas

Figur 3.29. *Illustration av att den betingade fördelningen för X, givet att $Y = y$, är normalfördelad.*

Låt oss som avslutning tolka detta resultat. Den betingade fördelningen för X är således normalfördelad. Vad gäller väntevärdet har detta förändrats från det ursprungliga (μ_x) med en faktor som är proportionell mot hur mycket y avvek från sitt väntevärde. Proportionalitetsfaktorn $\rho\sigma_x/\sigma_y$ kan tolkas som att vi lägger större vikt vid hur mycket y avviker från μ_y ju mer X och Y är korrelerade. Om korrelationen är positiv förskjuts väntevärdet åt samma håll som y:s avvikelse från μ_y och omvänt.

Vad gäller den nya standardavvikelsen har den minskat med en faktor som beror av korrelationen: ju mer korrelerade X och Y är, dvs. ju större $|\rho|$ är, ju mindre standardavvikelse har X betingat på att vi observerat $Y = y$. Standardavvikelsen beror däremot inte på observationens värde y.

ÖVNING 3.75

Låt (X_1, X_2, X_3, X_4) vara multinomialfördelad med parametrar $n = 8$, $p_1 = 0.1$, $p_2 = 0.2$, $p_3 = 0.3$ och $p_4 = 0.4$.

a) Beräkna $p_{X_1,X_2,X_3,X_4}(0, 2, 2, 4)$.

b) Bestäm $E(X_3)$.

c) Bestäm $D(X_4)$.

d) Bestäm $C(X_1, X_4)$.

ÖVNING 3.76

Tidigare högskolebetyg var U (=underkänd), G (=godkänd) respektive VG (= väl godkänd). Kurser i matematiska ämnen brukade på vid ordinarie tentamenstillfällen ha ca 20% underkända, 55% godkända och 25% väl godkända. Antag att det vid en "typisk kurs" (dvs. inte svårare eller lättare än genomsnittet) tenterade 25 studenter vid ordinarie tentamenstillfället. Låt (X_U, X_G, X_{VG}) specificera hur många tenterande som fick betygen U, G respektive VG.

a) Vilken fördelning har (X_U, X_G, X_{VG})?
b) Bestäm $p_{X_U,X_G,X_{VG}}(4, 15, 6)$.
c) Bestäm $E(X_U)$.
d) Bestäm $D(X_G)$.
e) Bestäm $\rho(X_U, X_{VG})$.

ÖVNING 3.77

Fortsättning från Övning 3.76. Antag att antalet underkända blev 5, dvs. $X_U = 5$. Betinga med avseende på detta och beräkna följande

a) $p_{X_G|X_U}(15|5)$.
b) $E(X_{VG}|X_U = 5)$.
c) $D(X_{VG}|X_U = 5)$.

ÖVNING 3.78

Låt $(X, Y) \sim \mathrm{N}(\mu_x = 1, \mu_y = 2, \sigma_x = 0.5, \sigma_y = 1, \rho = 0)$.

a) Beräkna $P(X \leq 2)$.
b) Beräkna $P(X \leq 2, 1 \leq Y \leq 3)$.
c) Beräkna $P(1 \leq X \leq 2, Y > 3.5)$.
d) Beräkna $P(X > 0.5, Y < 1.5)$.

ÖVNING 3.79

Låt $(X, Y) \sim \mathrm{N}(\mu_x = 1, \mu_y = 2, \sigma_x = 0.5, \sigma_y = 1, \rho = 0.5)$. Antag att $Y = 3$ har observerats.

a) Bestäm fördelningen för X betingat av att $Y = 2$.
b) Bestäm $P(X > 1|Y = 3)$ och dess obetingade motsvarighet $P(X > 1)$.

ÖVNING 3.80

I beviset för Sats 3.28 på sidan 135 utelämnades ett par steg i manipulationen av en viss integral. Utför dessa.

3.11 Funktioner av slumpvariabler

I det här avsnittet ska vi konstruera nya slumpvariabler från givna slumpvariabler med hjälp av funktioner. Vi börjar med funktioner av en slumpvariabel.

3.11.1 Funktioner av en slumpvariabel

Antag att vi har en slumpvariabel X med fördelningsfunktion $F_X(x)$. Definiera nu en ny slumpvariabel $Y = g(X)$ där $g(x)$ är en reell funktion. En naturlig fråga är då vilken fördelningsfunktion Y har. I de allra flesta tillämpningarna är $g(x)$ en kontinuerlig och strikt växande eller strikt avtagande funktion och för detta fall blir fördelningsfunktionen lättare att uttrycka. Vi inskränker oss därför till detta fall nedan men tar senare upp några exempel då detta inte gäller. Anledningen till att formlerna blir lättare om $g(x)$ är monoton beror på att $g(x)$ då har en invers funktion $g^{-1}(y)$. Den inversa funktionen $g^{-1}(y)$ definieras som

$$g^{-1}(y) = \{x;\ g(x) = y\},$$

dvs. det x för vilket $g(x) = y$. Den strikta monotoniciteten tillsammans med kontinuiteten försäkrar om att det finns exakt ett sådant x. Vi har följande resultat.

SATS 3.31

Låt X vara en slumpvariabel med fördelningsfunktion $F_X(x)$ och antag att $g(x)$ är en kontinuerlig strikt monoton funktion. Då gäller att $Y = g(X)$ har fördelningsfunktion $F_Y(y) = F_X(g^{-1}(y))$ om $g(x)$ är växande. Om $g(x)$ är avtagande gäller $F_Y(y) = 1 - F_X(g^{-1}(y))$ om X är kontinuerlig och $F_Y(y) = 1 - F_X(g^{-1}(y) - 1)$ om X är diskret.

BEVIS

Det gäller att $F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$. Vidare gäller att om $g(x)$ är växande så är $g(x) \leq y$ om och endast om $x \leq g^{-1}(y)$. Så mäng-

den $A \in \Omega$ för vilket $g(X(u)) \leq y$ är således samma mängd som mängden för vilket $X(u) \leq g^{-1}(y)$. Alltså har mängderna även samma sannolikheter: $P(g(X) \leq y) = P(X \leq g^{-1}(y))$. Detta är i sin tur detsamma som $F_X(g^{-1}(y))$ vilket var satsens första påstående.

Om $g(x)$ är avtagande gäller i stället $g(x) \leq y$ om och endast om $x \geq g^{-1}(y)$. Man får därför i stället $F_Y(y) = P(X \geq g^{-1}(y))$. Om X är diskret (heltalsvärd) blir detta detsamma som $1 - F_X(g^{-1}(y) - 1)$ medan det blir $1 - F_X(g^{-1}(y))$ om X är kontinuerlig.

När man härlett fördelningsfunktionen kan man lätt få fram sannolikhetsfunktionen respektive täthetsfunktionen. Vi visar endast resultatet för fallet att $g(x)$ är växande, och överlåter det avtagande fallet åt läsaren (Övning 3.87 på sidan 157). Det gäller att $p_Y(k) = F_Y(k) - F_Y(k-1)$ för det diskreta fallet och $f_Y(y) = \frac{d}{dy} F_Y(g^{-1}(y))$ i det kontinuerliga fallet. För $g(x)$ växande gäller därför $p_Y(k) = F_X(g^{-1}(k)) - F_X(g^{-1}(k-1))$, respektive $f_Y(y) = f_X(g^{-1}(y))/g'(g^{-1}(y))$. Det senare resultatet följer av att derivatan av en invers funktion är inversen av derivatan av den ursprungliga funktionen, tagit i motsvarande punkt, dvs $\frac{d}{dy} g^{-1}(y) = 1/g'(g^{-1}(y))$.

EXEMPEL 3.44

Låt $X \sim \text{Exp}(2)$, dvs. X har täthetsfunktion $f_X(x) = e^{-2x}$, $x \geq 0$. Låt vidare $g(x) = 1 - e^{-2x}$ vilket är en kontinuerlig strikt växande funktion. Vi ska nu härleda fördelnings- och täthetsfunktionerna för $Y = g(X) = 1 - e^{-2X}$.

Fördelningsfunktionen för X ges av $F_X(x) = \int_0^x 2e^{-2t} dt = 1 - e^{-2x}$. Inversa funktionen till $g(x)$ är $g^{-1}(y) = -\ln(1-y)/2$. Vi får således följande fördelningsfunktion

$$F_Y(y) = F_X(g^{-1}(y)) = F_X(-\ln(1-y)/2) = 1 - e^{-2(-\ln(1-y)/2)} = y.$$

Eftersom X antar endast positiva värden så antar $Y = 1 - e^{-2X}$ värden i intervallet $(0, 1)$. Om vi vill beräkna täthetsfunktionen direkt, utan att gå "via" fördelningsfunktionen, behöver vi $g'(x)$ som blir $2e^{-2x}$, vilket ju är detsamma som $f_X(x)$. Vi får således

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))} = 1,$$

för $0 \leq y \leq 1$ (och $f_Y(y) = 0$ i övrigt). Eftersom vi redan visat att $F_Y(y) = y$, $0 \leq y \leq 1$ får vi ju detta enklare genom att derivera detta. Vi ser således att vi erhållit den likformiga $U(0, 1)$-fördelningen som

behandlades i Avsnitt 3.8.1 på sidan 100. Genom att transformera en exponentialfördelad slumpvariabel har vi således erhållit en likformigt fördelad slumpvariabel. Att genom transformationer erhålla "nya" slumpvariabler är en mycket användbar teknik inom simulering vilket vi diskuterar vidare i Kapitel **??**.

Vi avslutar med ett tips till läsaren: När man vill studera egenskaper hos transformerade slumpvariabler är det bäst att under räkningarnas gång använda sig av fördelningsfunktionen snarare än sannolikhetsfunktion-/täthetsfunktion. Övergå till de senare först på slutet.

3.11.2 Funktioner av flera slumpvariabler

Ibland är man intresserad av funktioner av flera slumpvariabler. Vanligaste exemplen är summor, men också t.ex. maximum och minimum vill man ibland känna fördelningen för. Om vi t.ex. har två oberoende slumpvariabler X och Y kan man vara intresserad av fördelningen för summan $Z = X + Y$.

Vi börjar allmänt och låter (X, Y) vara en tvådimensionell slumpvariabel och inför $Z = g(X, Y)$ för någon funktion $g(x, y)$. Fördelningsfunktionen för Z ges då av

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(g(X, Y) \leq z) \\ &= \begin{cases} \sum_{(j,k);\ g(j,k)\leq z} p_{X,Y}(j, k) \\ \int_{(x,y);\ g(x,y)\leq z} f_{X,Y}(x, y)dxdy. \end{cases} \end{aligned}$$

Så mycket mer exakta resultat kan man inte säga en godtycklig funktion $g(x, y)$ (se dock Avsnitt 3.11.4 för approximativa resultat). Om vi är intresserade av summan $g(x, y) = x + y$ kan vi däremot komma lite längre. Om man studerar summan av två slumpvariabler säger man att man *faltar* deras fördelningar. Om X och Y är diskreta gäller nämligen att sannolikhetsfunktionen för $Z = X + Y$ ges av

$$\begin{aligned} p_Z(z) &= P(Z = z) = P(X + Y = z) = \sum_{(j,k);\ j+k=z} p_{X,Y}(j, k) \\ &= \sum_j p_{X,Y}(j, z - j), \end{aligned}$$

den sista likheten erhålls genom substitution av summationsindex. Man kan på motsvarande sätt visa att för det kontinuerliga fallet gäller att

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)dx.$$

De s.k. faltningsformlerna blir ofta enklare att använda i fallet att X och Y är oberoende. Man har då

$$p_Z(z) = \sum_j p_X(j)p_Y(z-j), \text{ respektive}$$
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx.$$

EXEMPEL 3.45

Vi ska i detta exempel bevisa att Sats 3.13 på sidan 83 verkligen är sann med hjälp av faltningsformlerna. Låt alltså $X \sim \text{Bin}(n_1, p)$ och $Y \sim \text{Bin}(n_2, p)$ vara oberoende binomialfördelade slumpvariabler med samma p och definiera $Z = X + Y$ som summan av de två slumpvariablerna. Vi ska nu bestämma $p_Z(z)$. För att summan skall vara lika med z så måste X ligga mellan det största av 0 och $z - n_2$ (eftersom Y högst kan vara n_2 och summan ska vara z) och det minsta av n_1 och z. Vi har således

$$\begin{aligned}
P(Z = k) = P(X + Y = k) &= \sum_{j=\max(0,z-n_2)}^{\min(n_1,z)} P(X = j, Y = k - j) \\
&= \sum_{j=\max(0,z-n_2)}^{\min(n_1,z)} P(X = j)P(Y = k - j) \\
&= \sum_{j=\max(0,z-n_2)}^{\min(n_1,z)} \binom{n_1}{j} p^j (1-p)^j \binom{n_2}{k-j} p^{k-j}(1-p)^{n_2-(k-j)} \\
&= p^k(1-p)^{n_1+n_2-k} \sum_{j=\max(0,z-n_2)}^{\min(n_1,z)} \binom{n_1}{j}\binom{n_2}{k-j} \\
&= \binom{n_1+n_2}{k} p^k(1-p)^{n_1+n_2-k}.
\end{aligned}$$

Den sista likheten är en välkänd kombinatorisk likhet. Vi hoppar därför över beviset av den. Intuitivt säger den att, för att välja k bland $n_1 + n_2$ måste vi välja j bland de n_1 första och resten $(k - j)$ bland de n_2 övriga, för något j. Slutsatsen är alltså att $Z = X + Y \sim \text{Bin}(n_1 + n_2, p)$, vilket var just vad satsen påstod.

EXEMPEL 3.46

Låt $X \sim \text{Po}(5)$ och $Y \sim \text{Po}(3)$ vara oberoende Poissonfördelade slumpvariabler (se Avsnitt 3.7.7). Vi år då att $Z = X+Y$ har sannolikhetsfunktion

$$\begin{aligned} p_Z(z) &= \sum_{j=-\infty}^{\infty} p_X(j)p_Y(z-j) = \sum_{j=0}^{z} p_X(j)p_Y(z-j) \\ &= \sum_{j=0}^{z} \frac{5^j e^{-5}}{j!}\frac{3^{z-j}e^{-3}}{(z-j)!} = \frac{8^z e^{-8}}{z!}\sum_{j=0}^{z}\binom{z}{j}\left(\frac{5}{8}\right)^j\left(\frac{3}{8}\right)^{z-j} \\ &= \frac{8^z e^{-8}}{z!}. \end{aligned}$$

Om vi hade ersatt 5 med ett godtyckligt a och 3 med ett godtyckligt b hade vi erhållit $a + b$ i stället för ”8” ovan. Detta betyder alltså att om man har två oberoende Poissonfördelade slumpvariabler så är deras summa också Poissonfördelad med en parameter som är summan av de enskildas parametrar.

Om man vill härleda fördelningen för det största eller minsta av två slumpvariabler kan detta göras relativt enkelt. Vi behandlar endast fallet att variablerna är oberoende. Sätt $U = \min(X, Y)$ och $V = \max(X, Y)$. Vi kan nu härleda fördelningsfunktionerna för U och V genom att observera att $U > u$ om och endast om $X > u$ *och* $Y > u$. På motsvarande sätt gäller att $V \leq v$ om och endast om $X \leq v$ *och* $Y \leq v$. Av detta kan man lätt härleda fördelningsfunktionerna för U och V.

SATS 3.32

Låt X och Y vara oberoende slumpvariabler med fördelningsfunktioner $F_X(x)$ respektive $F_Y(y)$, och definiera $U = \min(X, Y)$ och $V = \max(X, Y)$. Då gäller att

$$\begin{aligned} F_U(u) &= 1 - (1 - F_X(u))(1 - F_Y(u)) \\ F_V(v) &= F_X(v)F_Y(v). \end{aligned}$$

Om $X_1, \ldots, X_n$ är oberoende och likafördelade med fördelningsfunktion $F_X(x)$ har $Y = \min(X_1, \ldots, X_n)$ och $Z = \max(X_1, \ldots, X_n)$ fördelningsfunktionerna

$$\begin{aligned} F_Y(y) &= 1 - (1 - F_X(y))^n \\ F_Z(Z) &= (F_X(z))^n. \end{aligned}$$

BEVIS

Resultaten följer direkt från antagandena om oberoende och definitionen av fördelningsfunktion:

$$\begin{aligned} F_U(u) &= P(min(X,Y) \leq u) = 1 - P(min(X,Y) > u) \\ &= 1 - P(X > u,\ Y > u) = 1 - (1 - F_X(u))(1 - F_Y(u)) \\ F_V(v) &= P(max(X,Y) \leq v) = P(X \leq v,\ Y \leq v) = F_X(v)F_Y(v). \end{aligned}$$

Beviset för flera likafördelade slumpvariabler är helt analogt.

EXEMPEL 3.47

En motor upphör först att fungera helt när dess samtliga 4 tändstift gått sönder. Antag vidare att cylindrarnas livslängd är oberoende och likafördelade Exp($\lambda = 1/7$) vilket alltså betyder att cylindrarna i genomsnitt håller i sju år. Om vi beräknar tiden T tills motorn upphör att fungera helt blir detta således $T = \max(X_1, X_2, X_3, X_4)$, där X_i är de enskilda livslängderna. Fördelningsfunktionen för de enskilda tändstiften blir då $F_X(x) = 1 - e^{-x/7}$. Vi får således $F_T(t) = F_X^4(t) = (1 - e^{-t/7})^4$. Om vi i stället är intresserad av tiden S då motorns funktion blir nedsatt pga att något tändstift inte fungerar så kan S skrivas som $S = \min(X_1, X_2, X_3, X_4)$. Fördelningsfunktion blir $F_S(s) = 1 - (1 - F_X(s))^4 = 1 - (e^{-s/7})^4 = 1 - e^{-(4/7)s}$ vilket alltså betyder att det minsta av ett antal oberoende och likafördelade exponentialfördelade slumpvariabler också är exponentialfördelad, fast med en parameter som är den ursprungliga multiplicerat antalet slumpvariabler.

3.11.3 Väntevärden och högre moment

Om vi t.ex. vill beräkna väntevärde och varians för den transformerade variabeln $Y = g(X)$ kan detta göras genom att först beräkna sannolikhetsfunktionen/täthetsfunktionen för Y och därefter beräkna väntevärdet enligt definition: $E(Y) = \sum k p_Y(k)$ eller motsvarande integral. Ett betydligt snabbare sätt är emellertid att använda Sats 3.4 på sidan 59. Där visades det att $E(Y)$ även kan beräknas genom följande ekvation

$$E(Y) = E(g(X)) = \sum_k g(k) p_X(k),$$

eller motsvarande integral. Samma formel kan även användas för variansen:

$$V(Y) = E(Y^2) - (E(Y))^2 = E(g^2(X)) - (E(g(X)))^2,$$

där $E(g^2(X)) = \sum_k g^2(k)p_X(k)$. Ibland är man intresserad av högre potenser och har för detta ändamål infört begreppet moment som vi nu definierar.

DEFINITION 3.37 (MOMENT AV EN SLUMPVARIABEL)

Låt X vara en slumpvariabel. Slumpvariabelns k:te *moment* definieras då som $E(X^k)$. Mer noggrant är detta k:te *nollpunktsmomentet* och $E((X - \mu)^k)$, där $\mu = E(X)$ är väntevärdet, kallas för det k : *te centralmomentet*.

Vi ska nu studera ett antal olika val av funktioner för att på så sätt härleda ett antal användbara resultat. I räkningarna antar vi att våra slumpvariabler är diskreta, men exakt samma teknik kan användas för att visa att resultaten även gäller kontinuerliga slumpvariabler. Vi börjar med *linjärtransformation*, $g(X) = aX + b$ för några konstanter a och b. Vi får direkt att

$$\begin{aligned} E(aX + b) &= \sum_k (ak + b)p_X(k) = a\sum_k kp_X(k) + b\sum_k p_X(k) \\ &= aE(X) + b, \end{aligned}$$

vilket visas lika lätt för det kontinuerliga fallet. Om vi betraktar två slumpvariabler (X, Y) och betraktar en linjärkombination av dessa, plus en konstant, får vi på motsvarande sätt

$$\begin{aligned} E(aX + bY + c) &= \sum_{j,k} (aj + bk + c)p_{X,Y}(j, k) \\ &= a\sum_j jp_X(j) + b\sum_k kp_Y(k) + c \\ &= aE(X) + bE(Y) + c, \end{aligned}$$

där andra likheten baseras på att marginalfördelningarna $p_X(j) = \sum_k p_{X,Y}(j, k)$ och $p_Y(k) = \sum_j p_{X,Y}(j, k)$. Sambandet $E(aX + bY + c) = aE(X) + bE(Y) + c$ brukar uttryckas som att väntevärdet är en *linjär operator*. Med hjälp av induktion kan man lätt visa att motsvarande samband gäller för en linjärkombination av godtyckligt många slumpvariabler,

$$E(a_1X_1 + \ldots a_nX_n) = a_1E(X_1) + \ldots + a_nE(X_n). \tag{3.6}$$

Notera att vi inte gör några förutsättningar om oberoende.

Vi gör nu motsvarande variansberäkningar. Vi börjar alltså med $V(aX + b)$. Vi beräknar först $E((aX + b)^2)$:

$$\begin{aligned} E((aX + b)^2) &= \sum_k (ak + b)^2 p_X(k) = \sum_k (a^2k^2 + 2abk + b^2)p_X(k) \\ &= a^2E(X^2) + 2abE(X) + b^2. \end{aligned}$$

Variansen blir därför

$$a^2E(X^2) + 2abE(X) + b^2 - (aE(X) + b)^2 = a^2V(X).$$

För linjärkombinationen $aX + bY + c$ får vi

$$\begin{aligned} E((aX + bY + c)^2) &= \sum_{j,k} (aj + bk + c)^2 p_{X,Y}(j, k) \\ &= \sum_{j,k} (a^2j^2 + b^2k^2 + c^2 + 2abjk + 2acj + 2bck)p_{X,Y}(j, k) \\ &= a^2E(X^2) + b^2E(Y^2) + c^2 \\ &\quad + 2abE(XY) + 2acE(X) + 2bcE(Y). \end{aligned}$$

Från tidigare vet vi att $E(aX+bY+c) = aE(X)+bE(Y)+c$, så dess kvadrat blir

$$\begin{aligned} (E(aX + bY + c))^2 &= a^2(E(X))^2 + b^2(E(Y))^2 + c^2 \\ &\quad + 2abE(X)E(Y) + 2acE(X) + 2bcE(Y). \end{aligned}$$

Om vi tar differensen av dessa två uttryck får vi således variansen

$$V(aX + bY + c) = a^2V(X) + b^2V(Y) + 2abC(X, Y).$$

Även här kan man lätt generalisera till flera variabler. Man får då

$$V(a_1X_1 + \ldots a_nX_n) = \sum_{k=1}^{n} a_k^2 V(X_k) + 2 \sum_{i<j} a_i a_j C(X_i, X_j).$$

I fallet att variablerna är parvis oberoende försvinner kovarianstermerna. Vi sammanfattar våra resultat i följande sats.

SATS 3.33

Låt X och Y, respektive $X_1, \ldots X_n$ vara slumpvariabler och $a,\ b,\ c$, respektive $a_1, \ldots, a_n$ vara godtyckliga konstanter. Då gäller

$$\begin{aligned} E(aX + bY + c) &= aE(X) + bE(Y) + c, \\ V(aX + bY + c) &= a^2V(X) + b^2V(Y) + 2abC(X, Y), \\ E(a_1X_1 + \ldots a_nX_n) &= a_1E(X_1) + \ldots + a_nE(X_n), \\ V(a_1X_1 + \ldots a_nX_n) &= \sum_{k=1}^{n} a_k^2V(X_k) + 2\sum_{i<j} a_ia_jC(X_i, X_j). \end{aligned}$$

Motsvarande formel för linjär transformation av en slumpvariabel fås genom att sätta $b = 0$ i formeln ovan.

EXEMPEL 3.48 (*Väntevärde och varians för negativ binomialfördelning*)

I Avsnitt 3.7.8 definierades en negativ binomialfördelad slumpvariabel X som antalet försök som krävs för att uppnå r lyckade, där försöken lyckas oberoende med sannolikhet p. Om vi låter Z_1 vara antal försök som krävs till det första lyckade, Z_2 antal försök därefter som behövs tills det andra lyckade försöket, och så vidare upp till Z_r, så gäller att $X = Z_1 + \ldots + Z_r$. Vidare är $Z_1, \ldots, Z_r$ oberoende (de har inga försök gemensamt) och de är alla ffg eftersom de räknar antal försök till nästa lyckade. Från Sats 3.17 på sidan 96 vet vi således att $E(Z_i) = 1/p$ och $V(Z_i) = q/p^2$. Från Sats 3.33 med $a_i = \ldots = a_n = 1$ gäller således att $E(X) = r/p$ och $V(X) = nq/p^2$, och därför att $D(X) = \sqrt{nq}/p$.

EXEMPEL 3.49

Två personer betraktar följande spel. Person A singlar en slant tills dess att han får en klave. Person A vinner 10 kr per kastad ”krona” som föregår första klaven. Spelet kostar 15 kr att delta i. Låt Y vara antal ”krona” som föregår första klaven. Då är $Y \sim \text{Geo}(p = 1/2)$. Person A:s nettoresultat Z blir $Z = 10Y - 15$. Detta resultat har väntevärde $E(10Y - 15) = 10E(Y) - 15$. Från tidigare vet vi att en geometriskt fördelad slumpvariabel gäller $E(Y) = (1 - p)/p = 1$ varför väntevärdet blir $E(Z) = 10 - 15 = -5$, ett negativt förväntat netto som i de flesta spel ...

Vad gäller variansen blir den $V(10Y-15) = 100V(Y) = 100(1-p)/p^2 = 200$ och standardavvikelsen blir således $\sqrt{200} \approx 14.1$. Slutsatsen är alltså att spelet i genomsnitt ger förlust, men eftersom standardavvikelsen är relativt stor finns det icke-försumbar chans att gå med vinst.

Ett viktigt specialfall av Sats 3.33 är när slumpvariablerna är oberoende och lika fördelade, eller åtminstone har samma väntevärde och varians, och då man betraktar summan av variablerna eller variablernas medelvärde.

FÖLJDSATS 3.2

Låt $X_1, \ldots X_n$ vara oberoende, alla med samma väntevärde $E(X_i) = \mu$ och standardavvikelse $D(X_i) = \sigma$. Låt $\bar{X} = \sum_{i=1}^{n} X_i/n$ beteckna medelvärdet av slumpvariablerna. Då gäller

$$E\left(\sum_{i=1}^{n} X_i\right) = n\mu \quad V\left(\sum_{i=1}^{n} X_i\right) = n\sigma^2 \quad D\left(\sum_{i=1}^{n} X_i\right) = \sqrt{n}\sigma$$

$$E(\bar{X}) = \mu \qquad V(\bar{X}) = \frac{\sigma^2}{n} \qquad D(\bar{X}) = \frac{\sigma}{\sqrt{n}}.$$

BEVIS

Resultatet följer direkt från Sats 3.33 genom att välja $a_i = 1$ respektive $a_i = 1/n$.

Låt X ha väntevärde μ och standardavvikelse σ. Från Sats 3.33 inser man då att man genom att dra ifrån väntevärdet och dividera med standardavvikelsen, dvs. transformerar till $(X-\mu)/\sigma$, får en slumpvariabel med väntevärde 0 och standardavvikelse 1. Denna transformation är vanligt förekommande, bl.a. för normalfördelningen, varför den har fått ett eget namn.

DEFINITION 3.38 (STANDARDISERAD SLUMPVARIABEL)

Låt Y vara en slumpvariabel med väntevärde $\mu = E(Y)$ och standardavvikelse $\sigma = D(Y)$. Då är $Z = (X-\mu)/\sigma$ motsvarande *standardiserade slumpvariabel*.

ANMÄRKNING 3.40

En standardiserad slumpvariabel har således alltid väntevärde 0 och standardavvikelse 1.

Man kan även välja funktioner med egna parametrar. Detta kan ibland vara användbart om man vill studera en slumpvariabel som kanske har en svårtillgänglig fördelningsfunktion. Vi definierar nu några sådana.

DEFINITION 3.39 (SANNOLIKHETSGENERERANDE, MOMENTGENERERANDE OCH KARAKTERISTISK FUNKTION)

Låt X vara en slumpvariabel. Dess *sannolikhetsgenererande funktion* $\rho_X(s)$ definieras som

$$\rho_X(s) = E(s^X)$$

för sådana s där väntevärdet är definierat. Dess *momentgenererande funktion* $\phi_X(t)$ definieras som

$$\phi_X(t) = E(e^{tX})$$

för sådana t där väntevärdet är definierat. Slutligen definieras den *karakteristiska funktionen* $\psi_X(t)$ av

$$\psi_X(t) = E(e^{itX})$$

där $i = \sqrt{-1}$ är det imaginära talet.

ANMÄRKNING 3.41

Sannolikhetsgenererande funktioner används oftast för positiva slumpvariabler. Den momentgenererande funktionen existerar inte alltid, men den är i alla fall väldefinierad för alla negativa t för positiva slumpvariabler (och för alla positiva t för negativa slumpvariabler). Fördelen med den karakteristiska funktionen är att den alltid existerar eftersom $|e^{itx}| \leq 1$ *för alla reella tal t och x. Eftersom vi inte förutsätter kunskaper om de imaginära talen kommer vi dock inte använda denna funktion ytterligare.*

ANMÄRKNING 3.42

Dessa funktioner är s.k. transformer av fördelningarna. Inom matematiken finns även andra transformer som t.ex. Fouriertransformen och Laplacetransformen.

EXEMPEL 3.50

Låt X vara geometriskt fördelad, dvs. $p_X(k) = q^k p$, $k = 0, 1, 2, \ldots$. Dess sannolikhetsgenererande funktion ges då av

$$\rho_X(s) = E(s^X) = \sum_{k=0}^{\infty} s^k q^k p = p \sum_{k=0}^{\infty} (qs)^k = \frac{p}{1 - qs},$$

för $-1 < s < 1$.

EXEMPEL 3.51

Antag att $Y \sim \mathrm{N}(\mu, \sigma^2)$. Då är dess momentgenererande funktion

$$\begin{aligned}\phi_Y(t) = E(e^{tY}) &= \int_{-\infty}^{\infty} e^{ty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-(\mu+\sigma^2 t))^2}{2\sigma^2} + t\mu + \sigma^2 t^2} dy \\ &= e^{t\mu+\sigma^2 t^2}.\end{aligned}$$

Den näst sista likheten erhålls med kvadratkomplettering och den sista följer av att $e^{t\mu+\sigma^2 t^2}$ kan brytas ut ur integralen och att det som då återstår är en normalfördelningstäthet (med väntevärde $\mu+\sigma^2 t$ och standardavvikelse σ) som därmed integrerar sig till 1.

Den momentgenererande funktionen har fått sitt namn eftersom man kan erhålla slumpvariabelns samtliga moment från den på följande sätt. Antag t.ex. att X är kontinuerlig. Momentgenererande funktionen blir då $\phi_X(t) = \int e^{tx} f_X(x)dx$. Om vi deriverar $\phi_X(t)$ med avseende på t och dessutom antar att vi får byta ordning på integration och derivering så får vi $\phi'_X(t) = \int x e^{tx} f_X(x)dx$. Om vi speciellt tittar på fallet $t = 0$ får vi $\phi'_X(0) = \int x f_X(x)dx = E(X)$. Vi kan alltså erhålla slumpvariabelns väntevärde genom att derivera den momentgenererande funktionen en gång. Om vi i stället deriverar momentgenererande funktionen k gånger (skrivs $\phi_X^{(k)}(\cdot)$) och observerar derivatan i nollan får vi slumpvariabelns k:te moment: $\phi_X^{(k)}(0) = E(X^k)$ vilket förklarar varför funktionen kalas momentgenererande funktionen. Byta ordning på derivering och integration (alternativt summation) får man göra om momentgenererande funktionen existerar i något intervall kring $t = 0$. Vi har således skissat beviset för följande resultat.

SATS 3.34

Om $\phi_X(t)$ är momentgenererande funktion till en slumpvariabel X och $\phi_X(t)$ är definierad i ett intervall som täcker in $t = 0$ så gäller $\phi_X^{(k)}(0) = E(X^k)$, där $\phi_X^{(k)}(0)$ är k:te derivatan av $\phi_X(t)$ evaluerad i $t = 0$.

ANMÄRKNING 3.43

Man kan visa flera andra resultat för de olika transformerna. Huvudidén är att, under olika regularitetsvillkor, så definierar transformen fördelningen. Detta kan vara användbart om fördelningen är krånglig medan lämplig transform är mer hanterlig.

Vi visar nu en olikhet som kan vara användbar när man vill göra uppskattningar av sådant man kanske inte beräkna exakt. Låt $f(x)$ vara en *konvex* funktion. För den som inte minns är f konvex om det för alla x och y samt för $0 < \alpha < 1$ gäller att $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ vilket illustrerar i Figur 3.30.

Bild saknas

Figur 3.30. *Bild på en konvex funktion $f(x)$ och den räta linjen $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ mellan två godtyckliga punkter x och y.*

För funktioner med andraderivata gäller att $f(x)$ är konvex om och endast om $f''(x) \geq 0$ för alla x.

SATS 3.35 (JENSENS OLIKHET)

Låt X vara en slumpvariabel med ändligt väntevärde och $f(x)$ en konvex funktion. Då gäller att

$$f(E(X)) \leq E(f(X)).$$

BEVIS

Vi ger endast en skiss av beviset för det diskreta fallet. Den viktiga insikten är att det för en konvex funktion följer från definitionen att olikheten även gäller för fler ”viktade punkter” dvs. $f(\sum_k k p_X(k)) \leq \sum_k p_X(k) f(k)$. Men detta är just detsamma som Jensens olikhet påstår!

EXEMPEL 3.52

Det gäller att $f(x) = x^2$ och $g(x) = e^x$ bägge är konvexa funktioner. Från Jensens olikhet kan vi således dra slutsatsen att $E(X^2) \geq (E(X))^2$ och att $E(e^X) \geq e^{E(X)}$. Den första olikheten är detsamma som att $V(X) \geq 0$.

Vi avslutar detta avsnitt med att härleda två räkneregler för väntevärden och varianser, som kan vara användbara då dessa är svåra att beräkna, men där man lättare kan beräkna dessa moment om man får betinga på någon annan variabel. Låt för detta ändamål (X, Y) vara en två-dimensionell slumpvariabel. Betrakta nu funktionen $g(k) := E(X \mid Y = k)$ som ju enligt definitionen av betingade väntevärden ges av

$$g(k) = E(X \mid Y = k) = \sum_j j p_{X|Y}(j \mid k) = \sum_j j \frac{p_{X,Y}(j,k)}{p_Y(k)}.$$

Om vi nu vill beräkna väntevärdet av denna funktion när vi betraktar Y som slumpvariabel får vi således

$$\begin{aligned} E(g(Y)) &= \sum_k g(k) p_Y(k) = \sum_{j,k} j \frac{p_{X,Y}(j,k)}{p_Y(k)} p_Y(k) \\ &= \sum_j j p_X(j) = E(X). \end{aligned}$$

Ofta skrivs denna relation på formen $E(X) = E(E(X \mid Y))$ eftersom $g(Y) = E(X \mid Y)$.

Vi ska nu härleda en liknande formel för variansen. Inför

$$\begin{aligned} h(k) &= E(X^2 \mid Y = k) - (E(X \mid Y = k))^2 \\ &= \sum_j j^2 p_{X|Y}(j \mid k) - \left(\sum_j j p_{X|Y}(j \mid k) \right)^2, \end{aligned}$$

vilket betyder att $h(k) = V(X \mid Y = k)$. Låt oss nu beräkna $V(g(Y)) + E(h(Y)) = V(E(X \mid Y)) + E(V(X \mid Y))$. Vi tar termerna var för sig

$$\begin{aligned} V(g(Y)) &= E(h^2(Y)) - (E(h(Y))^2 \\ &= E\left([E(X \mid Y)]^2\right) - (E(E(X \mid Y)))^2, \\ E(h(Y)) &= E(E(X^2 \mid Y)) - E\left([E(X \mid Y)]^2\right). \end{aligned}$$

Från detta får vi således att $V(g(Y)) + E(h(Y)) = E(E(X^2 \mid Y)) - (E(E(X \mid Y)))^2$. Vi har ovan visat att den andra termen är detsamma som $(E(X))^2$. Den första termen blir

$$\begin{aligned} E(E(X^2 \mid Y)) &= \sum_k \left(E(X^2 \mid Y = k\right) p_Y(k) \\ &= \sum_k \left(\sum_j j^2 p_{X|Y}(j \mid k) \right) p_Y(k) \\ &= \sum_k \sum_j j^2 \frac{p_{X,Y}(j,k)}{p_Y(k)} p_Y(k) = E(X^2). \end{aligned}$$

Detta medför att $V(E(X \mid Y)) + E(V(X \mid Y)) = V(X)$. Vi sammanfattar våra resultat i följande sats.

SATS 3.36

Låt (X, Y) vara en tvådimensionell slumpvariabel med ändliga varianser. Då gäller att

$$\begin{aligned} E(X) &= E(E(X|Y)), \\ V(X) &= E(V(X \mid Y)) + Var(E(X \mid Y)). \end{aligned}$$

ANMÄRKNING 3.44

För den ovane kan det vara något oklart vilket slumpvariabel som man ska beräkna väntevärde respektive varians med avseende på i uttrycket ovan. Vill man förtydliga detta kan man skriva $E_y(E_x(X|Y))$, $E_y(Var_x(X \mid Y))$, *och* $Var_y(E_x(X \mid Y)$.

EXEMPEL 3.53

Antag att ett slumpförsök upprepas ett Poisson-fördelat (Po(λ)) antal gånger, och att vart och ett av försöken ”lyckas” med sannolikhet p, samt

att försöken sker oberoende av varandra. Om vi låter Y beteckna antalet försök som görs och X antalet lyckade försök så gäller att $X \mid Y = y \sim \text{Bin}(y, p)$. Från detta drar vi slutsatsen att $E(X \mid Y = y) = yp$ och $V(X \mid Y = y) = yp(1-p)$. Således får vi $E(X) = E(E(X \mid Y)) = E(Yp) = \lambda p$. För variansen får vi

$$\begin{aligned} V(X) &= E(V(X \mid Y)) + V(E(X \mid Y) = E(Yp(1-p)) + V(Yp) \\ &= \lambda p(1-p) + p^2\lambda = \lambda p, \end{aligned}$$

eftersom $Y \sim \text{Po}(\lambda)$ har varians λ så $V(Yp) = p^2\lambda$. I själva verket gäller ett starkare resultat, nämligen att $X \sim \text{Po}(\lambda p$ vilket överlåts till läsaren att visa (Övning 3.84

3.11.4 Felfortplantningsformlerna

Som nog har framgått tidigare i innevarande avsnitt är det inte alltid så lätt att bestämma fördelningen för en funktion $g(X)$ av en slumpvariabel X med känd fördelning. Även att endast bestämma $E(g(X))$ och $V(g(X))$ kan vara svårt. Ibland kanske man dessutom inte ens känner fördelningen för den ursprungliga slumpvariabeln X och då har vi ingen möjlighet att bestämma exakt väntevärde och varians. Vi ska i det här avsnittet redogöra för hur man beräknar approximativa moment av funktioner av slumpvariabler, en metod som ofta kan vara mycket användbar. Av denna anledning har metoden flera olika benämningar, bl.a. *Gauss approximationsformel* och på engelska kallas metoden ofta för ”the δ-method”. Den kanske vanligaste benämningen är *felfortplantingsformlerna*.

Approximation baseras på att vi Taylorutvecklar funktionen runt väntevärdet $\mu = E(X)$. Från andra ordningens approximation får vi då

$$g(x) \approx g(\mu) + (x-\mu)g'(\mu).$$

Från denna approximation får vi $E(g(X)) \approx E\left(g(\mu) + g'(\mu)(X-\mu)\right) = g(\mu)$ respektive $V(g(X)) \approx V(g(\mu) + (x-\mu)g'(\mu)) = [g'(\mu)]^2 V(X)$, där vi använt oss av räkneregler i Sats 3.34 på sidan 151.

Om vi har en funktion av flera slumpvariabler $h(x_1, \ldots, x_n)$ kan vi på motsvarande sätt härleda en approximativ formel för väntevärde och varians genom Taylorutveckling runt väntevärdespunkten $(\mu_1, \ldots, \mu_n)$ där $\mu_i = E(X_i)$. Det gäller nämligen att

$$h(x_1, \ldots, x_n) \approx h(\mu_1, \ldots, \mu_n) + \sum_{i=1}^{n} (x_i - \mu_i) h^{(i)}(\mu_1, \ldots, \mu_n),$$

där $h^{(i)}(\mu_1, \ldots, \mu_n)$ symboliserar partiella derivatan av h med avseende på komponent i evaluerad i punkten $(\mu_1, \ldots, \mu_n)$. Med hjälp av räknereglerna i Sats 3.34 kan vi även här använda denna approximation till att få fram approximativt väntevärde och varians. Vi sammanfattar våra resultat.

METOD 3.1 (FELFORTPLANTNINGSFORMLERNA)

Låt X respektive $X_1, \ldots, X_n$ vara slumpvariabler med väntevärden μ respektive $\mu_1, \ldots, \mu_n$, och låt $g(x)$ och $h(x_1, \ldots, x_n)$ vara funktioner med kontinuerliga derivatorer. Då är följande approximationer användbara

$$\begin{aligned} E(g(X)) &\approx g(\mu) \\ V(g(X)) &\approx [g'(\mu)]^2 V(X) \\ E(h(X_1, \ldots, X_n)) &\approx h(\mu_1, \ldots, \mu_n) \\ V(h(X_1, \ldots, X_n)) &\approx \sum_i [h^{(i)}(\mu_1, \ldots, \mu_n)]^2 V(X_i) \\ &\quad + 2 \sum_{i<j} h^{(i)}(\mu_1, \ldots, \mu_n) h^{(j)}(\mu_1, \ldots, \mu_n) C(X_i, X_j). \end{aligned}$$

ANMÄRKNING 3.45

Om slumpvariablerna är oberoende försvinner kovarianstermerna i det nedersta uttrycket.

För linjära funktioner $g(x) = ax + b$ och $h(x_1, \ldots, x_n) = a_1 x_1 + \ldots + a_n x_n$ gäller formlerna exakt vilket följer av Sats 3.34 (se Övning 3.94).

EXEMPEL 3.54

Att beräkna väntevärde och varians av $1/X$ exakt analytiskt är svårt för de flesta slumpvariablerna. Med hjälp av felfortplantningsformlerna får man däremot, genom att välja $g(x) = 1/x$, $E(1/X) \approx 1/\mu$ och $V(1/X) \approx V(X)/(E(X))^4$. Standardavvikelsen blir således $D(1/X) \approx D(X)/[E(X)]^2$.

ÖVNING 3.81

I Övning 3.3, sidan 42, definierades en slumpvariabel Y som hur mycket för lite en ”100 g” chokladkaka vägde. Antag att vikten X på en slumpvis vald chokladkaka kan beskrivas av en normalfördelning med väntevärde 103g (producenter vill nog ligga lite i överkant i genomsnitt) och

standardavvikelse 2g. Bestäm funktionen $f(x)$ för vilket $Y = f(X)$. Bestäm även $P(Y = 0)$, $P(1 < Y < 2)$ och $F_Y(2)$.

ÖVNING 3.82

Låt $X \sim \text{Exp}(\lambda)$, dvs. X är exponentialfördelad med täthetsfunktion $f_X(x) = \lambda e^{-\lambda x}$, $x \geq 0$. Bestäm täthet och fördelningsfunktionen för $Y = X^2$ samt beräkna $E(Y)$.

ÖVNING 3.83

Låt $X \sim \text{Re}[0, 10]$, dvs. X har täthetsfunktion $f_X(x) = 1/10$, $0 < x < 10$ (och $f_X(x) = 0$ annars). Bestäm X:s tredje och fjärde moment $E(X^3)$ och $E(X^4)$.

ÖVNING 3.84

Antag som i Exempel 3.53 på sidan 153 att $Y \sim \text{Po}(\lambda)$ och att $X \mid Y = k \sim \text{Bin}(k, p)$. Visa att $X \sim \text{Po}(\lambda p)$. (L)

ÖVNING 3.85

Låt X vara en diskret slumpvariabel med sannolikhetsfunktion $p_X(j)$. Definiera Y som $Y = aX + b$ där vi kan anta att a och b är heltal så att även Y blir heltalsvärd.

a) Bestäm $p_{X,Y}(j, k)$.
b) Beräkna $C(X, Y)$.
c) Beräkna $\rho(X, Y)$.

ÖVNING 3.86

Antag att $X_1, X_2, \ldots$ är oberoende och likafördelade slumpvariabler, med $E(X_i) = 10$ och $D(X) = 2$. Låt $\bar{X}_n = \sum_{i=1}^{n} X_i/n$.

a) Beräkna $E(\bar{X}_1)$, $E(\bar{X}_{10})$ och $E(\bar{X}_{100})$.
b) Beräkna $D(\bar{X}_1)$, $D(\bar{X}_{10})$ och $D(\bar{X}_{100})$.
c) Vad måste n minst vara för att $D(\bar{X}_n) \leq 0.1$?

ÖVNING 3.87

Härled formlerna för sannolikhetsfunktion respektive täthetsfunktion för $Y = g(X)$ för fallet att $g(x)$ är strikt avtagande. Fallet med $g(x)$ växande gjordes efter Sats 3.31 på sidan 139.

ÖVNING 3.88

Antag att ett flygplan har två motorer som går sönder oberoende av varandra och med samma fördelning Exp(1) (vi tänker oss att tidsenheten är just ”förväntad livslängd”). Flygplanet havererar först när bägge motorerna är utslagna. Låt T beskriva denna tid. Jämför T med tiden S tills att ett plan med endast en motor, fast med dubbel genomsnittlig livslängd och fortfarande exponentialfördelning, havererar.

a) Bestäm $F_T(t)$ respektive $F_S(t)$.
b) Visa att det enmotoriga flygplanet har större risk att haverera för små t (dvs. för rimliga flygtider för relativt säkra plan).

ÖVNING 3.89

En teknisk mätinstrument går sönder varje enskild mätning med sannolikheten p. Innan mätinstrumentet gått sönder görs oberoende mätningar av antalet passerade elektroner, och varje sådan mätning är Poissonfördelad med väntevärde λ. Låt X ange totalt antalet elektroner som passerat innan apparaten gått sönder, och Y anger hur många mätningar som gjordes.

a) Bestäm fördelningen för Y.
b) Bestäm $E(X)$ och $V(X)$. (L)

ÖVNING 3.90

Låt $X \sim \text{Po}(\lambda)$. Bestäm den sannolikhetsgenererande funktionen $\rho_X(s) = E(s^X)$.

ÖVNING 3.91

Låt $Y \sim \text{Exp}(\beta)$. Bestäm den momentgenererande funktion $\phi_Y(t) = E(e^{tY})$.

ÖVNING 3.92

Låt X och Y vara oberoende positiva slumpvariabler. Definiera kvoten $Z = X/Y$ och beräkna approximativt $E(Z)$ och $V(Z)$ uttryckt i moment av X och Y.

ÖVNING 3.93

Antag att en befolknings vuxna män har normalfördelad längd med väntevärde 180 cm och standardavvikelse 5 cm, och befolkningens vuxna kvinnor också har normalfördelad längd, men med väntevärde 168 cm och standardavvikelse 5 cm. En vuxen person väljs på måfå ur befolkningen (som har lika många män och kvinnor). Låt Y beteckna personens längd. Y sägs då ha en blandad normalfördelning. Bestäm väntevärde, varians och standardavvikelse för Y. (L)

ÖVNING 3.94

Visa med hjälp av Sats 3.34 på sidan 151 att felfortplantningsformlerna gäller med likhet för linjära funktioner $g(x) = ax + b$ och $h(x_1, \ldots, x_n) = a_1x_1 + \ldots + a_nx_n$.

3.12 Stora talens lag

Vi ska nu gå igenom ett av de viktigaste resultaten inom sannolikhetsteorin, ”Stora talens lag”. Vi har redan nämnt detta resultat vid ett par tillfällen, och det utgör också en grundsten för den s.k. frekvenstolkningen av begreppet sannolikhet, se Avsnitt 2.3. Denna tolkning gick ut på att om man upprepar ett försök många gånger så bör den relativa andelen av försöken som resulterar i händelsen A ligga nära sannolikheten för A, $P(A)$. Vi visar resultatet mer allmänt för medelvärden av slumpvariabler som inte behöver ha samma fördelning utan bara samma väntevärde och standardavvikelse.

SATS 3.37 (STORA TALENS LAG)

Låt $X_1, X_2, \ldots$ vara en följd av oberoende slumpvariabler, alla med samma väntevärde $E(X_i) = \mu$ och standardavvikelse $D(X_i) = \sigma < \infty$.

Låt $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ vara medelvärdet av de n första variablerna. Då gäller, för alla $\epsilon > 0$, att

$$P(|\bar{X}_n - \mu| > \epsilon) \to 0, \quad \text{då } n \to \infty.$$

ANMÄRKNING 3.46

Det faktum att $P(|\bar{X}_n - \mu| > \epsilon) \to 0$ för alla $\epsilon > 0$ brukar formuleras som att $\bar{X}_n$ konvergerar i sannolikhet mot μ och skrivs ofta som $\bar{X}_n \xrightarrow{p} \mu$ (där p syftar på engelskans "probability"). Samma formulering och notation kan gälla för andra sekvenser av slumpvariabler. Således betyder $Y_n \xrightarrow{p}$ a detsamma som att $P(|Y_n - a| > \epsilon) \to 0$ för alla $\epsilon > 0$.

ANMÄRKNING 3.47

Ett viktigt specialfall är förstås när slumpvariablerna är likafördelade. Om man speciellt inför indikatorvariabeln X som är 1 om händelse A inträffar och 0 annars har man att $\mu = E(X) = 1 \cdot P(A) + 0 \cdot P(A^c) = P(A)$, och $\bar{X}_n$ är den relativa andelen som resulterar i händelsen A av n försök. Satsen säger i detta fall att den relativa frekvensen A-händelser konvergerar i sannolikhet mot $P(A)$, då $n \to \infty$.

BEVIS

Betrakta slumpvariabeln $\bar{X}_n$. Från Följdsats 3.2 på sidan 148 vet vi att $E(\bar{X}_n) = \mu$ och $D(\bar{X}_n) = \sigma/\sqrt{n}$. Från Chebyshevs olikhet (Sats 3.7 på sidan 67) får vi därför att

$$P(|\bar{X}_n - \mu| > \epsilon) \leq \frac{V(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

När $n \to \infty$ går detta mot 0 vilket alltså bevisar satsen

Detta viktiga resultat är en av grundstenarna inom mer eller mindre all empirisk vetenskap. Antag nämligen att man gör många oberoende observationer av någon slumpvariabel, t.ex. hållfastheten hos en metall, blodtrycket hos patienter behandlade med en ny medicin, eller livslängden hos personer i ett försäkringskollektiv. Då kommer medelvärdet av dessa observationer att ligga nära det sanna väntevärdet. Om vi inte vet det sanna väntevärdet kan vi

alltså gissa, eller skatta, väntevärdet med medelvärdet. Detta förfarande är en viktig ingrediens i statistisk inferensteori och behandlas senare i boken, bl.a. i Kapitel **??**.

ÖVNING 3.95

Antag att $X_1, X_2, \ldots$ är oberoende och likafördelade med väntevärde $E(X_i) = 10$ och standardavvikelse $D(X_i) = 1.5$.

a) Uppskatta $P(|\bar{X}_{20} - 10| > 0.5)$.

b) Hur många observationer måste man ta medelvärdet av för att motsvarande sannolikhet inte ska överstiga 0.1.

ÖVNING 3.96

Antag att $X_1, X_2, \ldots$ är oberoende och likafördelade Bernoulli-variabler med sannolikhet för lyckat utfall lika med 0.4, dvs $X_i \sim Be(p = 0.4)$. Med andra ord har X_i sannolikhetsfunktion $p(0) = 0.6$ och $p(1) = 0.4$ vilket medför att $E(X_i) = 0.4$ och $V(X_i) = 0.4 \cdot 0.6 = 0.24$. Det gäller att antalet lyckade bland $n = 20$ försök, $Y = \sum_{i=1}^{20} X_i$, är Bin$(n = 20, p = 0.4)$, och $\bar{X}_{20} = Y/20$.

a) Beräkna $P(|\bar{X}_{20} - 0.4| > 0.2) = P(|Y - 8| > 4)$ exakt med hjälp av Tabell 2.

b) Uppskatta $P(|\bar{X}_{20} - 0.4| > 0.1)$ med hjälp av Chebyshevs olikhet.

ÖVNING 3.97

Som vi tidigare visat kan en Poissonfördelad slumpvariabel Y_n med väntevärde n (Po(n)) beskrivas som summan av n st oberoende Po(1) variabler. Visa med hjälp av detta att Y_n/n konvergerar i sannolikhet mot en konstant, och vad denna konstant är. (L)

ÖVNING 3.98

Låt $Y_1, Y_2, \ldots$ vara en följd av slumpvariabler med tvåpunktsfördelning sådana att Y_n är lika med n med sannolikhet $1/n$ och 0 med resterande sannolikhet (dvs $P(Y_n = 0) = 1 - 1/n$ och $P(Y_n = n) = 1/n$).

a) Beräkna $E(Y_n)$.

b) Visa att det likväl gäller att $Y_n \xrightarrow{p} 0$. (L)

3.13 Centrala gränsvärdessatsen

I föregående avsnitt visades att medelvärdet $\bar{X}_n$ av en följd av oberoende slumpvariabler $X_1, X_2, \ldots$, alla med samma väntevärde μ och standardavvikelse σ, konvergerade i sannolikhet mot μ när $n \to \infty$. I det här avsnittet ska vi studera fördelningen för $\bar{X}_n - \mu$ när n växer, och vi gör detta under antagandet att slumpvariablerna inte bara har samma väntevärde och standardavvikelse, utan att de har samma fördelning så att vi har ett stickprov från en fördelning F.

Eftersom $\bar{X}_n - \mu$ går mot 0 i sannolikhet måste vi således skala upp $\bar{X}_n - \mu$ för att finna något intressant. Vi studerar därför i stället den standardiserade slumpvariabeln av $\bar{X}_n$. Från Följdsats 3.2 vet vi att $E(\bar{X}_n) = \mu$ och $D(\bar{X}_n) = \sigma/\sqrt{n}$. Den standardiserade slumpvariabeln av $\bar{X}_n$ ges således av $\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu)$, som alltså har väntevärde 0 och standardavvikelse 1 för alla n. Vi måste alltså förstora, eller ”blåsa upp” avvikelserna från väntevärdet på $\sqrt{n}$-skalan. Det märkliga är att denna standardiserade slumpvariabel konvergerar mot en och samma fördelning (standardiserad normalfördelning) oavsett vilken fördelning de ursprungliga slumpvariablerna, $X_1, X_2, \ldots$, har. Detta resultat kallas “Centrala gränsvärdessatsen” och får nog sägas vara sannolikhetsteorins allra viktigaste resultat.

SATS 3.38 (CENTRALA GRÄNSVÄRDESSATSEN)

Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade slumpvariabler med $E(X_i) = \mu$ och $D(X_i) = \sigma$, där $0 < \sigma < \infty$, och låt $\bar{X}_n = \sum_{i=1}^{n} X_i/n$. För godtyckliga $a < b$ gäller då att

$$P(a < \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) < b) \to \Phi(b) - \Phi(a), \quad \text{då } n \to \infty,$$

där $\Phi(\cdot)$ är fördelningsfunktionen för $\mathrm{N}(0, 1)$.

ANMÄRKNING 3.48

Det gäller alltså att fördelningen för $\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu)$ alltmer liknar den standardiserade normalfördelningen. Ett alternativt sätt att skriva detta är $\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \xrightarrow{d} N(0, 1)$, där d står för engelskans ”distribution” som betyder fördelning. Man säger att slumpvariabeln ifråga konvergerar i fördelning mot $N(0, 1)$.

ANMÄRKNING 3.49

Centrala gränsvärdessatsen finns i många olika versioner. T.ex. gäller den även om slumpvariablerna inte är likafördelade under vissa bivillkor, om slumpvariablerna är beroende förutsatt att beroendet är tillräckligt svagt, och med ett svagare villkor än ändlig varians. Ett viktigt forskningsområde inom sannolikhetsteorin består i själva verket av att försöka utreda huruvida det existerar en central gränsvärdessats för den studerade situationen eller inte.

BEVIS

Vi ger endast en skiss av beviset eftersom ett fullständigt bevis är lite för omfattande. Det vi kommer att visa är att den momentgenererande funktionen (mgf) för $Y_n = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu)$ konvergerar mot den mgf för en standardiserad normalfördelning. Den senare har mgf $\phi(t) = e^{t^2/2}$ vilket visades i Exempel 3.51 på sidan 150. Det gäller nämligen, men visas inte, att om en mgf konvergerar mot en annan mgf så konvergerar även fördelningen mot den andra fördelningen.

Om vi standardiserar de ursprungliga variablerna $X_1, X_2, \ldots$ blir dessa $Z_i = (X_i - \mu)/\sigma$. Det gäller då att

$$Y_n = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) = \sum_{i=1}^{n} Z_i/\sqrt{n}.$$

Vi ska alltså visa att mgf för $Y_n = \sum_{i=1}^{n} Z_i/\sqrt{n}$ konvergerar mot $e^{t^2/2}$. Men eftersom $X_1, X_2, \ldots$ är oberoende så är även $Z_1, Z_2, \ldots$ oberoende och då gäller att

$$\phi_{Y_n}(t) = E\left(e^{\sum_{i=1}^{n} tZ_i/\sqrt{n}}\right) = \prod_{i=1}^{n} E\left(e^{tZ_i/\sqrt{n}}\right) = \left(\phi_Z(t/\sqrt{n})\right)^n,$$

där den andra likheten visas i Övning 3.101 och den tredje beror på att alla Z_i har samma fördelning. Om vi Taylorutvecklar $\phi_Z(s)$ runt $s_0 = 0$ får vi

$$\phi_Z(s) = \phi_Z(0) + s\phi'_Z(0) + \frac{s^2}{2}\phi''(0) + O(s^3).$$

Av definitionen för mgf gäller alltid att $\phi(0) = 1$, och från Sats 3.34 på sidan 151 vet vi att $\phi_Z^{(k)}(0) = E(Z^k)$, där $\phi_Z^{(k)}(0)$ är k:te derivatan evaluerad i $t = 0$. Eftersom Z är standardiserad gäller att $E(Z) = 0$ och

$V(Z) = E(Z^2) = 1$, så $\phi_Z(s) = 1 + s^2/2 + O(s^3)$. Om vi väljer $s = t/\sqrt{n}$, för givet t, får vi

$$\phi_{Y_n}(t) = \left(\phi_Z(t/\sqrt{n})\right)^n = \left(1 + \frac{t^2}{2n} + O\left(\frac{t^3}{n^{3/2}}\right)\right)^n.$$

Om vi låter $n \to \infty$ får vi därför att $\phi_{Y_n}(t) \to e^{t^2/2}$ vilket är mgf för den standardiserade normalfördelningen.

Centrala gränsvärdessatsen uttalar sig om fördelningen då antalet observationer n går mot oändligheten. Det som gör resultatet så viktigt är emellertid vad som gäller *innan* n har gått mot oändligheten. Om vi har bildat medelvärdet $\bar{X}_n$, där n är ”någorlunda” stort, så gäller att medelvärdet är *approximativt* normalfördelat. Detta approximativa resultat kallas för normalfördelningsapproximation och baseras alltså på den *asymptotiska* centrala gränsvärdessatsen. Från beviset såg vi att $E(\bar{X}_n) = \mu$ och $D(\bar{X}_n) = \sigma/\sqrt{n}$, och för summan $\sum_{i=1}^n X_i$ gäller $E\left(\sum_{i=1}^n X_i\right) = n\mu$ och $D\left(\sum_{i=1}^n X_i\right) = \sigma\sqrt{n}$.

METOD 3.2 (NORMALFÖRDELNINGSAPPROXIMATION)

Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade slumpvariabler med $E(X_i) = \mu$ och $D(X_i) = \sigma$. Då gäller det att $\bar{X}_n$ är approximativt $\mathrm{N}(\mu, \sigma/\sqrt{n})$, och $\sum_{i=1}^n X_i$ är approximativt $\mathrm{N}(n\mu, \sigma\sqrt{n})$.

Således har vi approximationerna

$$P(a \le \bar{X}_n \le b) \approx \Phi\left(\frac{b-\mu}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{a-\mu}{\sigma/\sqrt{n}}\right), \tag{3.7}$$

$$P(c \le \sum_{i=1}^n X_i \le d) \approx \Phi\left(\frac{d-n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right). \tag{3.8}$$

ANMÄRKNING 3.50

I fallet att X är kontinuerlig spelar det ingen roll om olikheterna ovan är strikta eller ej eftersom sannolikheten att få exakt likhet är 0. För det diskreta fallet spelar det däremot roll och approximationerna kan förbättras med s.k. halvkorrektion som tas upp i nästa avsnitt (Avsnitt 3.14).

Vi illustrerar detta viktiga resultat med ett exempel.

EXEMPEL 3.55

Antag att $X_1, X_2, \ldots$ är exponentialfördelade med intensitet 3 ($X_i \sim$ Exp(3)). Vi vet från tidigare att detta medför att $E(X_i) = 1/3$ och $D(X_i) = 1/3$. Om vi gör 10 observationer gäller alltså att $\sum_{i=1}^{10} X_i$ är approximativt normalfördelad med väntevärde $10/3 \approx 3.33$ och standardavvikelse $\sqrt{10}/3 \approx 1.054$ medan medelvärdet är normalfördelat med väntevärde $1/3$ och standardavvikelse $1/(3\sqrt{n}) \approx 0.1054$. Således gäller t.ex. att

$$P(3.00 \leq \sum_{i=1}^{10} X_i \leq 4.00) \approx \Phi\left(\frac{4.00 - 3.33}{1.054}\right) - \Phi\left(\frac{3.00 - 3.33}{1.054}\right)$$
$$\approx \Phi(0.64) - \Phi(-0.32) = 0.7389 - (1 - 0.6255) = 0.364.$$

För medelvärdet får vi t.ex. att

$$P(0.2 \leq \bar{X}_{10} \leq 0.4) \approx \Phi\left(\frac{0.4 - 0.33}{0.1054}\right) - \Phi\left(\frac{0.20 - 0.33}{0.1054}\right)$$
$$\approx \Phi(0.64) - \Phi(-1.26) = 0.7389 - (1 - 0.8962) = 0.635.$$

Varför är då centrala gränsvärdessatsen och dess approximation så viktig? Den kanske allra viktigaste orsaken är att många slumpvariabler som dyker upp i verkligheten i sin tur påverkas av ett flertal andra slumpvariabler (en individs längd beror på ett flertal gener samt näringsintag under uppväxten, ett fordons livslängd beror på alla dess körningar och väderpåverkan m.m.). Centrala gränsvärdessatsen säger då att om inflytandet från andra slumpvariabler är någorlunda linjärt så blir den studerade slumpvariabeln normalfördelad. En annan viktig orsak till centrala gränsvärdessatsens betydelse är när man vill dra slutsatser om en slumpvariabels väntevärde och man gör detta genom att göra ett antal oberoende observationer från fördelningen ifråga och bildar medelvärdet av dessa, mer om detta i Kapitel **??**.

ÖVNING 3.99

Betrakta slumpvariabeln X med täthetsfunktion $f(x) = 2x$, $0 \leq x \leq 1$, och $f(x) = 0$ annars. Låt $X_1, \ldots X_{20}$ vara oberoende och likafördelade med denna fördelning.

a) Beräkna $E(X)$ och $D(X)$.

b) Beräkna $P(10.0 \leq \sum_{i=1}^{20} X_i \leq 15.0)$ approximativt.

ÖVNING 3.100

Antag att $Y \sim \text{Re}(0, 4)$, dvs $f_Y(y) = 1/4$, $0 \leq y \leq 4$. Låt $Y_1, \ldots Y_{10}$ vara oberoende med denna fördelning.

a) Beräkna $P(1.8 \leq Y \leq 2.2)$.

b) Beräkna $E(Y)$ och $D(Y)$.

c) Beräkna $P(1.8 \leq \bar{Y}_{10} \leq 2.2)$ approximativt.

ÖVNING 3.101

Visa att om $X_1, \ldots, X_n$ är oberoende slumpvariabler så gäller det att den momentgenererande funktionen $\phi_Y(t)$ för $Y = \sum_{i=1}^n a_i X_i$ satisfierar $\phi_Y(t) = \prod_{i=1}^n \phi_{X_i}(a_i t)$. (L)

3.14 Approximationer av fördelningar

I de tidigare avsnitten när vi behandlat specifika fördelningar har vi på ett par ställen nämnt att fördelningen ifråga för vissa val av parametrar liknar andra fördelningar. I detta avsnitt ska vi sammanfatta dessa approximationer av fördelningar. De kan vara till nytta när man ska beräkna sannolikheter och har problem att göra detta, men de kan också bidra till ökad förståelse genom att de bidrar till insikten om egenskaper hos fördelningar som gör att de liknar andra fördelningar. Innan vi går igenom dessa approximationer går vi först igenom ett generellt ”knep” som förbättrar approximationen i fallet att man approximerar en diskret fördelning med en kontinuerlig fördelning.

3.14.1 Halvkorrektion

Antag att vi har en diskret (heltalsvärd) slumpvariabel X vars fördelning F_X är svår att beräkna, men att vi i stället ska approximera den med en kontinuerlig fördelning F_Y. Det första man skulle göra är förstås att sätta $F_X(x) \approx F_Y(x)$. Eftersom X är diskret gäller dock t.ex. att $F_X(k + 0.9) = F_X(k)$ så denna approximation skulle ge $F_X(k) = F_X(k + 0.9) \approx F_Y(k + 0.9)$. Vi skulle alltså med resonemanget ovan approximera $F_X(k)$ med allt mellan $F_Y(k)$ upp till $F_Y(k + 0.999\ldots)$. Eftersom vi bör ha en entydig rekommendation bör vi välja ett unikt värde. Det förefaller rimligt att välja mittvärdet vilket innebär att vi gör följande approximation $F_X(k) \approx F_Y(k + 0.5)$. Detta är den s.k. halvkorrektionen. Denna korrigering visar sig ofta innebära en

kraftigt förbättrad approximation. I enstaka fall kan approximation förbättras ännu mer, men vi nöjer oss med att presentera denna viktigaste förbättring av approximationen.

METOD 3.3 (HALVKORREKTION)

Låt X vara en diskret slumpvariabel och antag att dess fördelning kan approximeras av en kontinuerlig fördelning F_Y. Då förbättras approximationen med *halvkorrektion* som definieras av

$$F_X(k) \approx F_Y(k + 0.5).$$

EXEMPEL 3.56

Antag att X är diskret med väntevärde 3 och standardavvikelse 0.8, och att $X_1, \ldots, X_{10}$ är 10 observationer från denna fördelning. Från centrala gränsvärdessatsen kan vi alltså approximera fördelningen för $\sum_{i=1}^{10} X_i$ med normalfördelningen med väntevärde $n\mu = 30$ och standardavvikelse $\sigma\sqrt{n} = 0.8\sqrt{10} \approx 2.53$. Antag att vi vill approximera $P(\sum_{i=1}^{10} X_i \leq 32)$. Om vi inte använder halvkorrektion får vi

$$P\left(\sum_{i=1}^{10} X_i \leq 32\right) \approx \Phi\left(\frac{32 - 30}{2.53}\right) = \Phi(0.79) = 0.7852.$$

Med halvkorrektion blir approximationen däremot

$$P\left(\sum_{i=1}^{10} X_i \leq 32\right) \approx \Phi\left(\frac{32.5 - 30}{2.53}\right) = \Phi(0.98) = 0.8365.$$

Som synes skiljer sig approximationerna åt ganska mycket och man kan utgå från att halvkorrektionen ger en bättre approximation.

3.14.2 Approximationer för några vanliga fördelningar

Vi ska nu presentera några fördelningar som kan approximeras av andra enklare fördelningar i vissa specialfall. Bakom varje föreslagen approximation finns en gränsvärdessats som motiverar den, men vi presenterar bara approximationerna. Approximationerna som nämns baseras på följande 3 huvudprinciper:

1. En stor ändlig population är ungefär detsamma som en oändlig population.
2. Om ett försök som upprepas har liten sannolikhet, p, att lyckas i varje enskilt försök, så är det förväntade antal lyckade försök en mycket viktigare parameter än hur många försök som utfördes.
3. Om många oberoende observationer görs, och det totala förväntade värdet och standardavvikelsen är relativt stora så är summan och medelvärdet approximativt normalfördelade.

Påstående 3 ovan är helt enkelt normalapproximationen från föregående avsnitt, medan de andra inte presenterats tidigare.

Den första principen är av intresse när man har att göra med hypergeometrisk fördelning. En slumpvariabel X sades vara hypergeometriskt fördelad med parametrar N, n och m, där N var populationens storlek, m antalet i populationen som har en viss egenskap av intresse, och slumpvariabeln X anger hur många bland n utan återläggning dragna element som har egenskapen av intresse. Som nämndes när den hypergeometriska fördelningen presenterades uppstår ett beroende mellan dragningar eftersom den kvarvarande populationsandelen med egenskapen efterhand ändras. Om å andra sidan populationsstorleken N är relativt stor och vi inte drar alltför många element n är detta beroende mycket svagt vilket motiverar att vi kan approximera den hypergeometriska fördelningen med fördelningen då vi i varje dragning har samma sannolikhet att få egenskapen och där dragningarna sker oberoende, dvs med binomialfördelningen med parametrar n och $p = m/N$ (som är sannolikheten för lyckat, dvs att vi får egenskapen av intresse).

Argumentet ovan är inte matematiskt stringent men syftar till att motivera approximationen. Rent matematiskt baseras approximation på det faktum att när N är stor och n/N liten så gäller

$$p_X(k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \approx \binom{n}{k}\left(\frac{m}{N}\right)^k\left(\frac{N-m}{N}\right)^{n-k},$$

och det senare är sannolikhetsfunktionen för just $\text{Bin}(n, m/N)$. Approximation gäller med likhet asymptotiskt om k och n är fixa, medan $N \to \infty$ och $m \to \infty$ på ett sådant sätt att $m/N \to p$ där $0 < p < 1$. Tumregeln som brukar användas för att få använda approximation är att $n/N < 0.1$.

METOD 3.4 (APPROXIMATION AV HYPERGEOMETRISK FÖRDELNING)

Antag att $X \sim \text{Hyp}(N, n, m)$ där $n/N < 0.1$. Då gäller att X kan approximeras med $\text{Bin}(n, m/N)$, dvs att

$$p_X(k) \approx \binom{n}{k} \left(\frac{m}{N}\right)^k \left(\frac{N-m}{N}\right)^{n-k}.$$

ANMÄRKNING 3.51

Att en approximation över huvud taget behövs beror på att binomialuttrycken i den hypergeometriska fördelningen är numeriskt instabila och även moderna datorer kan ha problem när $N > 10^6$ vilket ofta förekommer i tillämpningar (t.ex. i opinionsmätningar då man brukar dra ett par tusen individer ur Sveriges röstberättigade befolkning.

Ovan approximeras den hypergeometriska fördelningen med binomialfördelningen, vars fördelning i denna bok betecknas med $\text{Bin}(n, p)$. Även denna fördelning kan dock behöva approximeras om n (och k) är stort. Det som är lite svårare i detta fall är att det finns två alternativa approximationer beroende på värdet av p.

Vi börjar med fallet att p är litet (eller nära 1) och n är någorlunda stort. Vi har alltså ett försök som lyckas med liten sannolikhet p och som upprepas n oberoende gånger, och Y anger antalet försök som lyckas. Sannolikhetsfunktionen ges av $p_Y(k) = \binom{n}{k} p^k (1-p)^{n-k}$. Förväntat antal lyckade försök är np och variansen är $np(1-p) \approx np$ eftersom p är litet. När Poissonfördelningen presenterades (Avsnitt 3.7.7 på sidan 90) nämndes att Poissonfördelningen kan användas för att approximera binomialfördelningen då p är litet. I själva verket kan Poissonfördelningen $\text{Po}(\lambda)$ sägas vara gränsfördelningen av $\text{Bin}(n, p)$ när $n \to \infty$ och $p \to 0$ så att $np \to \lambda$. Vi visar inte detta strikt, men det följer om man betraktar sannolikhetsfunktionen för binomialfördelningen för ett fixt k och antar att n är stort och $p = \lambda/n$:

$$p_Y(k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{n^k}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda},$$

och högerledet är sannolikhetsfunktionen för $\text{Po}(\lambda)$. Det viktigaste för att approximation skall fungera är att p är litet eftersom Poissonsannolikheten att överstiga n (vilket är omöjligt i binomialfallet!) är mycket liten. Parametern n bör dock även vara ganska stor eftersom det annars inte är några problem att

räkna med binomialfördelningen exakt. Om $p > 0.9$ kan även då fördelningen approximeras med Poissonfördelningen genom att i stället för antal lyckade Y betrakta antalet misslyckade $n - Y$ som då blir Bin$(n, 1 - p)$

METOD 3.5 (POISSONAPPROXIMATION AV BINOMIALFÖRDELNING)

Antag att $Y \sim \text{Bin}(n, p)$ där $p < 0.1$. Då gäller att Y kan approximeras med Po$(\lambda = np)$, dvs att

$$p_Y(k) \approx \frac{(np)^k}{k!} e^{-np}.$$

Om $Y \sim \text{Bin}(n, p)$ där $p > 0.9$ kan fördelningen för Y approximeras med talet n minus en Po$(\lambda = n(1 - p))$-fördelning. Närmare bestämt kan sannolikhetsfunktionen för Y approximeras med

$$p_Y(k) \approx \frac{(n(1 - p))^{n-k}}{(n - k)!} e^{-n(1-p)}.$$

ANMÄRKNING 3.52

Poissonapproximationen av binomialfördelningen är ett exempel på vad som brukar kallas de små talens lag. Det är i själva verket så att det finns flera andra liknande situationer då man kan approximera med Poissonfördelningen, t.ex. om man betraktar många försök (inte nödvändigtvis oberoende eller likafördelade), men där varje försök har liten chans att lyckas. Att man säger "små talens lag" syftar alltså på att sannolikheterna är små.

EXEMPEL 3.57

Varje enskilt år löper individer i Sverige en (mycket) liten sannolikhet att dö av blixtnedslag. Om vi intar att denna sannolikhet är $1/2 \cdot 10^6$ och Sverige består av $9 \cdot 10^6$ individer betyder det att ett enskilt år så kommer antalet som dör av blixtnedslag således att bli Bin$(n = 9 \cdot 10^6,\ p = 1/2 \cdot 10^6)$. Detta kan med väldigt hög noggrannhet (eftersom p är mycket litet och n mycket stort) approximeras med Po$(9 \cdot 10^6/2 \cdot 10^6 = 4.5)$. Om vi t.ex. vill beräkna sannolikheten att minst 10 personer omkommer av blixtnedslag ett enskilt år så ges den av $P(Y \geq 10) = 1 - P(Y \leq 9)$ och motsvarande sannolikhet för Po(4.5) fås från Tabell 3 och blir $1-0.9829 = 0.0171$.

Antag nu i stället att varken p eller $1-p$ är så små, eller närmare bestämt att både np och $n(1-p)$ är någorlunda stora. Från tidigare vet vi att $Y \sim \text{Bin}(n, p)$ kan skrivas som summan av n Bernoullifördelade slumpvariabler ($Y = \sum_{i=1}^{n} X_i$), alla med samma p. Från normalapproximationen i föregående avsnitt vet vi därför redan att Y är approximativt normalfördelad. Väntevärdet är np och standardavvikelsen $\sqrt{np(1-p)}$. Med halvkorrektion får vi följande approximation.

METOD 3.6 (NORMALAPPROXIMATION AV BINOMIALFÖRDELNINGEN)

Antag att $Y \sim \text{Bin}(N, p)$ där $np(1-p) > 10$. Då gäller att X kan approximeras med $\text{N}(np, \sqrt{np(1-p)})$, med användning av halvkorrektion. Fördelningsfunktionen approximeras således av

$$F_Y(k) \approx \Phi\left(\frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right),$$

där $\Phi(\cdot)$ är fördelningsfunktionen för den standardiserade normalfördelningen som finns tabulerad i Tabell 4.

EXEMPEL 3.58

Vid en bilaffär efterfrågas färgen silvergrå av 30% kunderna i det långa loppet. Under en vecka har affären 97 besök som resulterar i beställning av bil. Antalet Y av dessa som beställer silvergrå är då $Y \sim \text{Bin}(97, 0.3)$ vilket alltså kan approximeras med $\text{N}(97 \cdot 0.3 = 29.1,\ \sqrt{97 \cdot 0.3 \cdot 0.7} \approx 4.51)$ eftersom $np(1-p) = 97 \cdot 0.3 \cdot 0.7 = 20.37 > 10$. Om vi t.ex. vill beräkna sannolikheten att mellan 20 och 40 bilar beställs med silvergrå färg får vi

$$\begin{aligned} P(20 \le Y \le 40) &= F_Y(40) - F_Y(19) \\ &\approx \Phi\left(\frac{40.5 - 29.1}{4.51}\right) - \Phi\left(\frac{19.5 - 29.1}{4.51}\right) \approx 0.978. \end{aligned}$$

Vi avslutar med en sista approximation som redan nämnts och som är ett specialfall av normalapproximation, nämligen för fallet med Poissonfördelningen då λ är ganska stort. Som tidigare nämnt kan ju t.ex. Poissonfördelning med väntevärde $\lambda = n$ skrivas som summan av n oberoende Poissonvariabler, alla med väntevärde 1. Av detta följer att normalapproximation kan

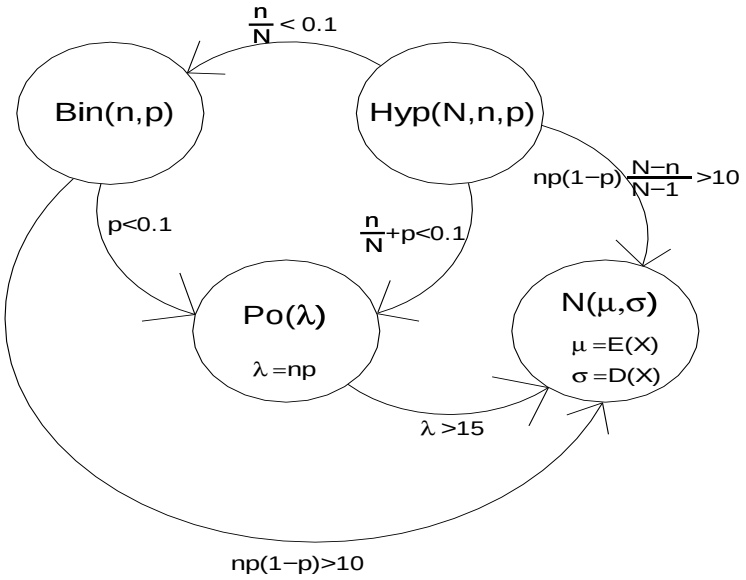
användas, helst i kombination med halvkorrektion eftersom vi approximerar den diskreta Poissonfördelningen med den kontinuerliga normalfördelningen.

METOD 3.7 (NORMALAPPROXIMATION AV BINOMIALFÖRDELNINGEN)

Antag att $Y \sim \text{Po}(\lambda)$ där $\lambda > 15$. Då kan Y approximeras med $\text{N}(\lambda, \sqrt{\lambda})$, med användning av halvkorrektion. Fördelningsfunktionen för Y approximeras således av

$$F_Y(k) \approx \Phi\left(\frac{k + 0.5 - \lambda}{\sqrt{\lambda}}\right).$$

Det går att kombinera approximationerna ovan. Tumregler för när detta är tillåtet och vilka approximationer som gäller illustreras i Figur 3.31. Det



Figur 3.31. *Illustration av ett antal fördelningsapproximationer och när de är tillåtna.*

är värt att påpeka att man inte skall approximera ”onödigt mycket”. Antag t.ex. att man är intresserad av en hypergeometriskt fördelad slumpvariabel $X \sim \text{Hyp}(N, n, p = m/N)$ där N är stort och $n/N < 0.1$, men där likväl $np(1 - p) > 10$. Vi kan då approximera den hypergeometriska fördelning-

en med Bin(n, p) som i sin tur kan approximeras med N$(np, \sqrt{np(1-p)})$. Men, eftersom vi vet att för en hypergeometriskt fördelad slumpvariabel gäller $E(X) = np$ och $D(X) = \sqrt{np(1-p)(N-n)/(n-1)}$ (se Sats 3.15 på sidan 88) så bör man använda denna standardavvikelse i stället för $\sqrt{np(1-p)}$ vid normalapproximationen. På så sätt minskas felet i approximationen. Detta illustreras i Figur 3.31 med att vi ska normalapproximera med $E(X)$ och $D(X)$ som moment.

ÖVNING 3.102

Antag att $X \sim \text{Hyp}(N = 10000, n = 10, p = 0.3)$. Beräkna approximativt
a) $P(X \leq 3)$,
b) $P(X \geq 5)$.

ÖVNING 3.103

Låt $Y \sim \text{Bin}(n = 100, p = 0.02)$.
a) Beräkna $P(X \leq 2)$ exakt.
b) Beräkna $P(X \leq 2)$ approximativt.

ÖVNING 3.104

Antag att $Y \sim \text{Bin}(n = 100, p = 0.4)$. Beräkna approximativt med halvkorrektion
a) $P(Y \leq 50)$,
b) $P(35 \leq Y \leq 45)$.

ÖVNING 3.105

Antag att $X_1, \ldots X_{15}$ är oberoende och Po(1) fördelade. Som tidigare visats är då $Y \sum_{i=1}^{15} X_i \sim \text{Po}(15)$.
a) Beräkna $P(\sum_{i=1}^{15} X_i \leq 15)$ exakt.
b) Beräkna $P(\sum_{i=1}^{15} X_i \leq 15)$ med normalapproximation utan halvkorrektion.
c) Beräkna $P(\sum_{i=1}^{15} X_i \leq 15)$ med normalapproximation med halvkorrektion.
d) Jämför approximationerna.

ÖVNING 3.106

Antag att $X \sim \text{Hyp}(N = 500, n = 40, p = 0.025)$. Beräkna approximativt
a) $P(X \leq 1)$ genom bara binomialapproximation,
b) $P(X \leq 1)$ även med Poissonapproximation.

ÖVNING 3.107

Antag att $X \sim \text{Hyp}(N = 1200, n = 100, p = 0.6)$. Beräkna approximativt
a) $P(X \geq 65)$ utan halvkorrektion,
b) $P(X \geq 65)$ med halvkorrektion.

ÖVNING 3.108

Antag att $X \sim \text{Hyp}(N = 10000, n = 10, p = 0.7)$. Beräkna approximativt $P(X > 8)$.

ÖVNING 3.109

Antag att $Y \sim \text{Bin}(n = 200, p = 0.99)$. Beräkna approximativt
a) $P(Y \leq 195)$,
b) $P(197 \leq Y \leq 199)$.

ÖVNING 3.110

Antag att en skola har 789 elever varav 401 är flickor och resten (388) är pojkar. Vid ett skollotteri delas 5 vinster ut till 5 slumpvis valda olika elever. Oberoende av detta väljs slumpvis 50 elever ut för intervju av hur de trivs i skolan. Låt Y ange antalet flickor som vinner på lotteriet och X antalet flickor som intervjuas.
a) Bestäm approximativt $P(Y > 3)$ (använd ej Tabell).
b) Bestäm approximativt $P(X > 30)$ (med halvkorrektion).

3.15 Blandade problem

301. Antag att ett slumpförsök upprepas ett Poisson-fördelat (Po(λ)) antal gånger, och att vart och ett av försöken ”lyckas” med sannolikhet p, samt att försöken sker oberoende av varandra (se Exempel 3.53 på sidan 153. Låter X beteckna antalet lyckade försök. Bestäm fördelningen för X. (L)

302. En något överoptimistisk pokerspelare satsar enbart på att få royal straight imperial flush, dvs. hjärter 10 upp till hjärter ess. Spelarna får byta kort en gång och hon (pokerspelaren) byter alla kort som inte ingår i denna följd. Vad är sannolikheten att hon for Royal straight imperial flush? (L)

303. En roulette ger rött (R) och svart (S) vardera med sannolikhet 1/2, och olika spelomgångar är oberoende av varandra. Bestäm det förväntade antalet spelomgångar tills sekvensen
    a) (R, R) inträffar,
    b) (R, S) inträffar.
    (L)

Ledningar till vissa av övningarna

Kapitel 2

2.7 $\Omega = A \cup A^c$.

2.8 $\emptyset = \Omega^c$

2.9 Utnyttja Sats 2.1 och Axiom 1.

2.10 Visa att $A \subseteq B$ medför att $P(A) \leq P(B)$ och utnyttja Axiom 2.

2.11 a) Utnyttja att $\sum_{n\geq 1} \frac{1}{k^2} = \pi^2/6$ och att Kolmogorovs Axiom 2 kräver att $\sum_{n\geq 1} P(A_n) = 1$.

2.12 Använd induktion och Kolmogorovs tredje axiom (Definition 2.2).

2.13 Utnyttja att $A \cup B \cup C = (A \cup B) \cup C$.

2.14 a) Utnyttja att $A \cup B \cup C = (A \cup B) \cup C$, b) Induktion.

2.22 Förenkla högerledet!

2.24 Det räcker att visa det första påståendet.

2.26 Testet ger rätt utslag om man är smittad *och* testet är positivt ($S \cap +$), samt om man inte är smittad *och* testet är negativt ($S^c \cap -$).

2.27 Rita ett träddiagram i två nivåer. Ringa in respektive händelse i träddiagrammet och beräkna sannolikheter för respektive händelse med hjälp av summaregeln och produktregeln.

2.31 a) Studera sannolikheten för att *inte* se ett stjärnfall och beräkna denna numeriskt.

2.33 Låt $A_i :=$ ”Del i innehåller exakt ett ess” och beräkna $P(A_1 \cap A_2 \cap A_3 \cap A_4)$ med hjälp av upprepad betingning.

Kapitel 3

3.10 Definiera lämpliga händelser och använd Kolmogorovs axiomsystem.

3.18 Beräkna $E(Y)$ på traditionellt sätt, härled $F_Y(y)$ och beräkna integralen.

3.20 För väntevärdet: Det gäller alltid att $E(X) = a + \int (x-a) f_X(x)\, dx$. Visa att intgralen blir 0.
För medianen: Jämför $\int_{-\infty}^{a} f_X(x)\, dx$ och $\int_{a}^{\infty} f_X(x)\, dx$.

3.23 a) $W = \sum_{k=1}^{X} Y_k$, där $\{Y_k\}$ är oberoende och exponentialfördelade.
b) $E(W) = E(X) \cdot E(Y)$.

3.34 Beräkna $E(X^2)$ på liknande sätt som i beviset för $E(X)$ genom att i summan skriva k^2 som $k(k-1)+k$ och summera uttrycken var för sig.

3.47 Använd exakt samma bevis som användes för att härleda väntevärde och varians för en ffg-variabel.

3.51 Använd att $V(U) = E(U^2) - (E(U))^2$ och beräkna $E(U^2)$.

3.63 Använd Tabell 3.

3.64 x_α definieras som lösningen till $P(X > x_\alpha) = \alpha$. Men $P(X > x_\alpha) = 1 - \Phi(\frac{x_\alpha - 100}{10})$.

3.69 Skriv ut $E(e^{tX})$ som en integral och kvadratkomplettera i exponenten så att allt som beror av integranden x finns inne i kvadraten.

3.73 Visa alltså att det för alla endimensionella mängder A gäller att $P(X \in A) = \int_A \left(\int_0^\infty f_{X,Y}(x,y) dy \right) dx$ genom att använda vad $f_{X,Y}(x,y)$ måste satisfiera för att vara en täthetsfunktion.

3.74 Uttryck $p_{X,Y}(x,y)$ i termer av $F_{X,Y}(x,y)$.

3.84 Beräkna sannolikhetsfunktionen $p_X(k)$ genom betingning: $P(X = j) = \sum_{k=j}^{\infty} P(X = j \mid Y = k) P(Y = k)$ och skriv om denna summa så att det innanför summation blir en Poissonsumma som summerar sig till 1.

3.89 Betinga med avseende på Y och använd formlerna för väntavärden och varians via betingning på annan variabel.

3.93 Betinga med avseende på om personen var en man eller en kvinna ($X = 0$ kan symbolisera man och $X = 1$ symbolisera kvinna)och använd formlerna för väntevärde och varians medelst betingning av en annan variabel.

3.95 Använd Chebyshevs olikhet.

3.97 Beräkna $E(Y_n/n)$ och $V(Y_n/n)$ och använd Chebyshevs olikhet.

3.98 Använd Chebyshevs olikhet.

3.101 Använd definitionen $\phi_Y(t) = E\left(e^{\sum_{i=1}^{n} a_i X_i t} \right)$ och utnyttja oberoendet för att skriva detta som en produkt av väntevärden.

Svar till övningarna

Kapitel 2

2.1 a) $\Omega = \mathcal{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, $A = \mathcal{N} = \{1, 2, \ldots\}$ och $B = \{1401, 1402, \ldots\}$.
b)$A \cup B = \{1, 2, \ldots\} \cup \{1400, 1401, \ldots\} = \{1, 2, \ldots\} = A$ och $A \cap B = \{1, 2, \ldots\} \cap \{1400, 1401, \ldots\} = \{1401, 1402, \ldots\} = B$.
c) $A^{\mathrm{c}} = \{\ldots, -2, -1, 0\}$, dvs ett negativt eller nollresultat, och $A \setminus B = \{1, 2, \ldots 1400\}$.

2.2 a) $\Omega = \{1, 2, \ldots\}$, $A = \{1, 2, \ldots, 10\}$, och $B = \{2, 4, 6, \ldots\}$.
b) $A \cup B = \{1, 2, \ldots 10, 12, 14, \ldots\}$, $A \cap B = \{2, 4, 6, 8, 10\}$.
c) $A^{\mathrm{c}} = \{11, 12, \ldots\}$ och $A \setminus B = \{1, 3, 5, 7, 9\}$.

2.3 a) $\Omega = \{1/1, 2/2, \ldots, 31/1, 1/2, \ldots, 31/12\}$, b) $S = \{1/9, \ldots 30/9\}$, $O = \{1/10, \ldots 31/10\}$, c) $V = \{24/9, \ldots 30/9, 1/10, \ldots 23/10\}$,
d) $S \setminus V = \{1/9, \ldots 23/9\}$, $O \cap V = \{1/10, \ldots, 23/10\}$ respektive $S \cup V = \{1/9, \ldots 30/9, 1/10, \ldots 23/10\}$.

2.4 $A = \{2, 4, 6\}$, $P(A) = 3/6 = 1/2$, $B = \{3, 6\}$, $P(B) = 2/6 = 1/3$, $A \cap B = \{6\}$, $P(A \cap B) = 1/6$, $A \cup B = \{2, 3, 4, 6\}$, $P(A \cup B) = 4/6 = 2/3$.

2.5 $P(A \cap B) = 0.3$.

2.6 $P(\text{vinst}) = P(H) + P(N) + P(T) = 0.111$.

2.11 a) $c = 6/\pi^2$, b) $P(B) = \frac{49}{6\pi^2} \approx 0.827$, c) $P(A_1^c) = 1 - 6/\pi^2 \approx 0.392$.

2.13 $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

2.14 a) $P(A \cup B \cup C) \leq P(A) + P(B) + P(C)$, b) $P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P(A_k)$.

2.15 Sannolikheten att det blir Marie är $1/31 \approx 0.03226$, och sannolikheten att det blir en flicka är $17/31 \approx 0.5484$.

2.16 a) Frekventistisk, b) subjektiv, c) frekventistisk.

2.17 a) Skattad, b) axiomatisk, c) beräknad (utifrån andra axiomatiska sannolikheter, nämligen att alla kortgivar har samma sannolikhet).
2.18 Sannolikheten för någon vinst är 0.102.
2.19 a) 1/10, b) 1/9.
2.20 a) 20, b) 41.
2.21 a) 70, b) 1/35, 12/35, 18/35, 4/35.
2.23 a) Nej ty $P(A \mid B) = 2/3 \neq 1/2 = P(A)$.
b) Ja, ty $P(A \mid C) = 1/2 = P(A)$. c) Ja, ty $P(C \mid B) = 1/3 = P(C)$.
2.25 Sannolikheten för båda felen är $3 \cdot 10^{-8}$.
2.26 0.74.
2.27 a) 1/3, b) 2/3, c) 2/3.
2.28 a) $P(E) = 5/100$, $P(F \mid E) = 4/99$ och $P(F \mid E^{\mathrm{c}}) = 5/99$.
b) $P(F) = P(F \mid E)P(E) + P(F \mid E^{\mathrm{c}})P(E^{\mathrm{c}}) = 5/100$.
c) Nej, ty $P(F \mid E) = 4/99 \neq 5/100 = P(F)$.
2.29 a) S för samhällsvetare, H för humanist, N för naturvetare och T för teknolog, samt P för privatanställd, O för offentliganställd och A för arbetslös.
b)$P(P \mid H) = 0.61$, $P(O \mid H) = 0.31$, $P(A \mid H) = 0.08$, $P(P \mid S) = 0.54$, $P(O \mid S) = 0.40$, $P(A \mid S) = 0.06$, $P(P \mid N) = 0.67$, $P(O \mid N) = 0.29$, $P(A \mid N) = 0.04$, $P(P \mid T) = 0.73$, $P(O \mid T) = 0.23$, $P(A \mid T) = 0.04$, $P(H) = 0.20$, $P(S) = 0.37$, $P(N) = 0.21$ och $P(T) = 0.22$.
c) $P(P) = 0.6231$, d) $P(T \mid A) = 0.159$.
2.30 a) O är händelsen att en slumpvis vald förare är onykter, och D är händelsen att en förare dör i en bilolocka.
b) $P(O \mid D) = 1/5$, c) $P(O) = 1/5$. d) Det bör gälla att $P(O) < 1/5$.
2.31 a) 0.000612, b) $P(A_1) \approx 0.6079$, $P(A_1|S) \approx 0.000993$.
2.32 a) $8/16575 \approx 0.000483$, b) $16/5525 \approx 0.00290$.
2.33 $2197/20825 \approx 0.105$.
2.34 $Q(S) = 0.429$, $Q(S^{\mathrm{c}}) = 0.571$.

Kapitel 3

3.1 $\Omega = \{0, 1, \ldots\}$ respektive $\Omega = \mathcal{R}$.
3.2 Låt $\Omega = \{1, 2, \ldots, 100\}$ och $Y(u) = u$. Eftersom alla utfall har samma sannolikhet, som därmed måste vara 1/100, så har vi likformig sannolikhetsfördelning (Definition 2.3 på sidan 11). Sannolikheten att vinsten överstiger 90 kr, blir då antalet gynnsamma utfall dividerat med antalet utfall totalt, alltså $10/100 = 0.1$. Det gäller alltså att $P(X > 90) = 0.1$.
3.3 Låt $\Omega = \mathcal{R}^+$, dvs. den positiva reella talaxeln, som anger chokladkakans vikt i gram (om man vill kan man möjligen inskränka utfallsrummet till (80, 120)

eller något annat intervall inom vilket man säkert vet att alla chokladkakors vikt ligger). Slumpvariabeln $X(u)$ som anger hur mycket för lite en chokladkaka med vikt u väger blir då $X(u) = 100 - u$ för $u < 100$ och $X(u) = 0$ för $u \geq 0$.

3.4 $p_Y(0) = 0.765$, $p_Y(1) = 0.096$, $p_Y(2) = 0.098$ och $P(Y \geq 3) = 0.041$. $P(Y \leq 2) = 0.959$, $P(1 \leq Y \leq 2) = 0.194$, $P(Y > 1) = 0.1386$.

3.5 a) Från sannolikhetsaxiomen följer att $p_X(4) = 1-(p_X(1)+p_X(2)+p_X(3)) = 0.1$. $P(X \leq 2) = p_X(1) + p_X(2) = 0.5$.
b) Kravet $\sum_{x=1}^{4} p_Y(x) = 1$ ger att $(0.2 + 0.3 + 0.4 + 0.5)c = 1$, dvs. $c = 1/1.4 \approx 0.714$. $P(Y > 2) = 1 - P(Y \leq 1) = 1 - (0.2 + 0.3)c \approx 0.357$.

3.6 $P(Z \leq 49) = 0.3$, $P(Z > 51) = 0.1$.

3.7 $p_X(0) = 0.002$, $p_X((1) = 0.038$, $p_X(2) = 0.279$, $p_X(3) = 0.682$.

3.8 a) 0.25, b) 0.56.

3.9 $F_X(x) = 0$, $x < 0$, $F_X(x) = 0.765$, $0 \leq x < 1$, $F_X(x) = 0.861$, $1 \leq x < 2$, $F_X(x) = 0.959$, $2 \leq x < 3$, $F_X(x) = 1$, $x \geq 3$.

3.10 Sätt $A_a = \{u; X(u) \leq a\}$ och $A_b = \{u; X(u) \leq b\}$ och $A_{(a,b]} = \{u; a < X(u) \leq b\}$. Då gäller att $A_a \cap A_{(a,b]} = \emptyset$ och $A_a \cup A_{(a,b]} = A_b$. Från Kolmogorovs axiomsystem har vi därför $P(A_b) = P(A_a \cup A_{(a,b]}) = P(A_a) + P(A_{(a,b]})$, dvs. $F_X(b) = F_X(a) + P(a < X \leq b)$ vilket är vad som skulle visas.

3.11 a) $F_X(x) = 2x - 5$ för $2.5 \leq x \leq 3$, $F_X(x) = 0$ för $x < 2.5$, $F_X(x) = 1$ för $x > 3$, b) 0.4, c) 1, d) 0.4, e) 3.33

3.12 a) $c = 2$, c) $F_Y(y) = y^2$ för $0 \leq y \leq 1$, d) $\lambda_Y(y) = 2y/(1 - y^2)$

3.13 a) $f_Y(y) = 1/\ln(5)y$, $1 \leq y \leq 5$, b) $\ln(y)/\ln(5)$, c) $\ln(2)/\ln(5)$, d) $(\ln(4) - \ln(3))/\ln(5)$.

3.14 $f_Y(y) = 2y$ för $0 \leq y \leq 1$ och 0 för övrigt.

3.15 $E(X) = 3.2$, $D(X) = \sqrt{12.4} \approx 1.47$.

3.16 a) 2.75, b) $x_{0.5} = 2.75$, c) 0.021, d) 0.144, e) 0.052.

3.17 a) 2/3, b) 0.707, c) 1/18, d) 0.736.

3.21 $P(X \geq 200) \leq 0.05$, så $x_{0.05} \leq 200$.

3.22 $P(|Y| \geq 4) \leq 4/4^2 = 0.25$.

3.23 a) $W = \sum_{k=1}^{X} Y_k$, där $\{Y_k\}$ är oberoende och exponentialfördelade. Om $X = 0$ blir även $W = 0$, vilket inträffar med sannolikhet 1/3. W kan därför skrivas som

$$W = \begin{cases} 0 & \text{med sannolikhet } 1/3, \\ \sum_1^Z Y_k & \text{med sannolikhet } 2/3, \end{cases}$$

där $Z = X|X > 0 \sim \text{ffg}(1/3)$.
b) $E(W) = 2$.

3.25 $p_Z(-1) = 0.995$ och $p_Z(100) = 0.005$. $E(Z) = -0.495$. $V(Z) \approx 50.75$.
3.26 $p_X(k) = 1/100$, $k = 1, \ldots, 100$, $E(X) = 50.5$, $D(X) = \sqrt{833.25} \approx 28.87$.
3.27 a) $E(X) = 1.8$ och $D(X) = 1.24$, b) 0.4435, c) 0.2642, d) 0.0683.
3.28 a) $E(Y) = 12$ och $D(Y) = \sqrt{3} \approx 1.73$, b) 0.4050, c) 0.9204.
3.29 $p_X(3) \approx 0.129$, $P(X \geq 3) = 1 - F_X(2) = 0.1841$.
3.31 a) 0.1480, b) 0.3277, c) 0.2749.
3.32 a) $E(X) = 1.2$, $D(X) = 0.748$, b) 0.3.
3.33 0.866.
3.35 $P(Y = 2) = 0.309$ och $P(X = 2) = 0.316$, dvs ganska lika varandra.
3.36 a) 0.353, b) 0.380, c) 1.32 respektive 1.15.
3.37 a) 0.0859, b) 0.4240.
3.38 $1/\sqrt{\lambda}$.
3.39 0.083.
3.40 a) 0.0782, b) 0.2517 (mer exakt värde =??).
3.41 a) 0.1029, b) 0.343, c) 0.51, d) $E(X)$=3.33 och $D(X)$=2.789.
3.42 a) 0.16, b) 0.8, c) 0.992, d) $E(Y)$=0.25 och $D(X)$=0.5590.
3.43 a) 0.488, b) 4.
3.44 $R(X) = \sqrt{q}$, $R(Y) = 1/\sqrt{q}$, och $R(Z) = \sqrt{q/r}$.
3.45 a) 0.5625, b) 0.0659, c) 0.9624, d) $E(X) = 4$, $D(X) = 1.155$.
3.46 a) 0.1073, b) 0.1890, c) $k = 12$.
3.48 a) 0.3 b) 0.8, c) $E(U) = 15$, $D(U) = 10/\sqrt{12} \approx 2.89$.
3.49 a) $f_U(u) = 1/10$ för $-5 \leq u \leq 5$ och $f_U(u) = 0$ för övriga u.
b) $F_U(u) = 0$, $u < -5$, $F_U(u) = (u + 5)/10$ för $-5 \leq u \leq 5$ och $F_U(u) = 1$, $u > 5$.
c) $\lambda_U(u) = 1/(5 - u)$ för $-5 < u < 5$ och $\lambda_U(u) = 0$ annars.
3.50 a) 0.1, b) 0.1, c) $1/\sqrt{3} \approx 0.577$.
3.52 a) $1 - e^{-2} \approx 0.865$, b) $e^{-0.8} - e^{-1.6} \approx 0.247$, $E(Y) = 0.25$ och $V(Y) = 0.0625$.
3.53 a) $e^{-0.4} - e^{-0.8} \approx 0.221$, b) $e^{-0.8} - e^{-1.2} \approx 0.148$, c) $R(T) = 1$.
3.54 a) $e^{-3} \approx 0.0498$, b) $1 - e^{-3/12} \approx 0.221$, 1/3 år.
3.56 a) 0.9332, b) 0.0919, c) 0.3085.
3.57 a) 0.3821, b) 0.7224, c) 0.8181.
3.58 a) 0.6915, b) 0.4938, c) 0.0401.
3.59 a) 0.0668, b) 0.9599, c) 0.4649.
3.60 a) 0.8849, b) 0.7517, c) 0.2666.
3.61 a) $\Phi(a/\sigma)$, b) $2\Phi(a/\sigma) - 1$, c) $\Phi(b/\sigma) + \Phi(a/\sigma) - 1$.
3.62 a) $2\Phi(a) - 1$, b) $\Phi(b) - 0.5$, c) $\Phi(b) - \Phi(a)$.
3.63 a) 0.47, b) 1.48, c) 2.05.
3.64 a) 116.45, b) 119.6, c) 125.76.
3.65 $x_\alpha = \mu + \sigma\lambda_\alpha$.

3.66 a) 0.0548, b) 0.3446, c) 0.6449.
3.67 a) 0.92, b) 0.998, c) 1.0000 (med minst 4 siffrors noggrannhet).
3.68 Barn: 15.9%, vuxna 3.6%.
3.69 $\phi_X(t) = e^{mut+\sigma^2t^2/2}$.
3.70 Fanns tidigare men nu uppdelat i a), b), c).
3.71 $E(X) = E(Y) = 7/12$, $V(X) = V(Y) = 11/144 \approx 0.0764$,
$C(X, Y) = 1/3 - (7/12)^2 \approx -0.00694$ och $\rho(X, Y) \approx -0.091$.
3.72 a) $f_{X|Y}(x \mid y) = (x+y)/(x+0.5)$, b) $E(X \mid Y = y) = (2+3y)/(3+6y)$,
$E(X \mid Y = y)$ avtar i y så väntevärdet är störst för $E(X \mid Y = 0) = 2/3$ och minimum vid $E(X \mid Y = 1) = 5/9$.
3.75 a) 0.0387, b) 2.4, c) 1.386, d) -0.32.
3.76 a) Multinomial $n = 25$, $p_U = 0.2$, $p_G = 0.55$ och $p_{VG} = 0.25$.
b) 0.0342, c) 2, d) 2.487, e) -0.289.
3.77 a) 0.167, b) 3.25, c) 4.30.
3.78 0.9772, b) 0.667, c) 0.0319, d) 0.260.
3.79 a) N($\mu = 1.25, \sigma = 0.433$),
b) $P(X > 1|Y = 3) \approx 0.719$ och $P(X > 1) = 0.5$.
3.81 $f(x) = 100 - x$, $x < 100$ och $f(x) = 0$ annars.
$P(Y = 0) = P(X > 100) \approx 0.9332$.
$P(1 < Y < 2) = P(98 < X < 99) \approx 0.0166$.
$F_Y(2) = P(X > 98) \approx 0.9938$.
3.82 $F_Y(y) = 1 - e^{-\lambda\sqrt{y}}$, $f_Y(y) = \lambda e^{-\lambda\sqrt{y}}/2\sqrt{y}$ och $E(Y) = 2\lambda^2$.
3.83 $E(X^3) = 250$, $E(X^4) = 2000$.
3.85 a) $p_{X,Y}(j, aj + b) = p_X(j)$ och $p_{X,Y}(j, k) = 0$ för övriga k.
b) $C(X, Y) = aV(X)$, $\rho(X, Y) = 1$ om $a > 0$ och $\rho(X, Y) = -1$ om $a < 0$.
3.86 a) $E(\bar{X}_1) = E(\bar{X}_{10}) = E(\bar{X}_{100}) = 10$.
b) $D(\bar{X}_1) = 2$, $D(\bar{X}_{10}) \approx 0.632$ och $D(\bar{X}_{100}) = 0.2$.
c) $n \geq 400$.
3.87 $p_Y(k) = 1-F_X(g^{-1}(k))-F_X(g^{-1}(k)-1)$, $f_Y(y) = f_X(g^{-1}(y))/|g'(g^{-1}(y))|$.
3.88 a) $F_T(t) = (1 - e^{-1})^2$, $F_S(t) = 1 - e^{-t/2}$.
b) följer av att $t^2 < t/2$ för små t.
3.89 a) $Y \sim$ Geo(p), b) $E(Y) = \lambda q/p$, $V(X) = \lambda q/p + \lambda^2 q/p^2$, där $q = 1 - p$.
3.90 $\rho_X(s) = e^{-\lambda(1-s)}$.
3.91 $\phi_Y(t) = \beta(\beta - t)$ för $t < \beta$.
3.92 $E(Z) \approx E(X)/E(Y)$, $V(Z) \approx V(X)/(E(Y))^2 + (E(X))^2V(Y)/(E(Y))^4$.
3.93 $E(Y) = 174$, $V(Y) = 61$, $D(Y) \approx 7.81$.
3.95 a) 0.45, b) $n \geq 90$.
3.96 a) 0.037, b) 0.3.
3.98 a) $E(Y_n) = 1$ för alla n.

3.99 a) $E(X) = 2/3$, $D(X) = 1/\sqrt{18}$.
b) 0.9421
3.100 a) 0.1, b) $E(Y) = 2$ och $D(Y) = 4/\sqrt{12}$, c) 0.418.
3.102 a) 0.6496, b) 0.8497.
3.103 a) 0.6767, b) 0.6767.
3.104 a) 0.9838, b) 0.7372.
3.105 a) 0.5681, b) 0.5, c) 0.5517,
d) Approximationen med halvkorrektionen är väsentligt bättre.
3.106 a) 0.7358, b) 0.7358.
3.107 a) 0.7934, b) 0.8212.
3.108 0.1493.
3.109 a) 0.9473, b) 0.7218.
3.110 a) 0.198, b) 0.9521.