Päätöksiä ja Paatoksia

Tommi Sottinen,
tommi.sottinen@uwasa.fi

28. marraskuuta 2011

Sisältö

Esipuhe

Nämä ovat luentomuistiinpanot syksyn 2011 kurssille *ORMS2020 Päätöksenteko epävarmuuden vallitessa* Vaasan yliopistossa. Kurssi on 5 op laajuinen sisältäen jotakuinkin 36 tuntia luentoja ja 12 tuntia harjoituksia.

Nämä luentomuistiinpanot ovat laajennettu, supistettu ja teoretisoitu versio edellisten vuosien luentomuistiinpanoista. Erityinen uutuus on lyhyt luku 0, jossa esitellään lyhyesti logiikkaa ja joukko-oppia.

Sisällöstä

Luvussa 0 esitämme pikaisesti logiikan ja joukko-opin merkinnät ja perusmääritelmät. Luku 1 on johdanto todennäköisyyslaskentaan. Enemmän todennäköisyyslaskennasta ja -teoriasta kiinnostunut löytää lisätietoa esimerkiksi lähteistä [5] ja [8]. Luvussa 2 tarkastelemme päätöksentekoa "pysäytetyssä ajassa", jolloin päätöksillä ei ole pitkän aikavälin syitä eikä seurauksia. Luvussa 3 tarkastelemme päätöksentekoa "ajassa", jolloin aikaisemmat päätökset ja tapahtumat vaikuttavat myöhempiin päätöksiin ja tapahtumien todennäköisyyksiin. Luku 4 palaa todennäköisyysteemaan ja todennäköisyyksien arviointiin. Luku 5 käsittelee järkevää päätöksentekoa ja luku 6 sen kritiikkiä!

Henkisestä omaisuudesta

Tämä kurssikirja on pitkälti koottu kirjallisuusluettelossa mainituista lähteistä — ja monista muista lähteistä, jotka kirjoittaja on unohtanut. Kirjoittaja ei ole jaksanut tai muistanut mainita, mistä mikäkin esimerkki, määritemä tmv. on poimittu. Kirjoittaja toivoo, ettei hän ole rikkonut pyhiä ©-lakeja liiaksi, ja pyytää varmuuden vuoksi anteeksi kaikilta, joiden kiltaprivilegioita hän on tullut loukanneeksi!

Vaasassa 28. marraskuuta 2011 T.S.

Luku 0

Logiikkaa ja joukko-oppia

> With a few brackets it is easy enough to see that 5 + 4 is 9. What is not easy to see is that 5 + 4 is not 6. – Carl Linderholm

> We all have a tendency to think that the world must conform to our prejudices. The opposite view involves some effort of thought, and most people would die sooner than think – in fact they do so. – Bertrand Russel

> Life is a tragedy for those who feel, and a comedy for those who think.
> – Jean de La Bruyère

Tämän luvun tarkoitus on, logiikan ja joukko-opin esittelyn lisäksi, esittää formaaleja merkintätapoja eli notaatiota. Toivottavasti tämä luku myös antaa alkusysäyksen täsmällisen ajattelun harjoittamiseen. Nimittäin täsmällinen notaatio ja täsmällinen ajattelu kulkevat käsi kädessä.

0.1 Logiikka

Logiikka on kaiken tiedon, päättelyn ja päätöksenteon perusta. Siksi onkin hämmästyttävää, kuinka vähan logiikkaa opetetaan suomalaisessa koulutuskoneistossa. Tämän luvun tarkoitus on, vähäisessä määrin, korjata tätä epäkohtaa.

Propositiologiikka

Propositio- eli lauselogiikassa yhdistetään väitteitä, eli *propositioita*, toisiinsa loogisilla *konnektiiveilla*. Propositiologiikan idea on se, että jos annettujen niin sanottujen atomilauseiden totuusarvot tiedetään, niin niistä konnektiiveilla johdettujen yhdistettyjen lauseiden totuusarvot voidaan laskea.

Esitämme useimmin käytetyt konnektiivit ja niiden *totuustaulut*. Totuustauluissa konnektiivit määritellään luettelemalla kaikki mahdolliset totuusvaihtoehdot. Tauluissa 0 tarkoittaa valhetta ja 1 totta.

0.1.1 Määritelmä (Disjunktio ja konjunktio)**.** *Disjunktio* $\vee$ tarkoittaa “tai” ja *konjunktio* $\wedge$ tarkoittaa “ja”. Niiden totuustaulut ovat (samassa taulussa):

P	Q	$P \vee Q$	$P \wedge Q$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	1	1

0.1.2 Huomautus (Disjunktion ja konjunktion symmetria)**.** Sekä disjunktio että konjunktio ovat symmetrisiä: $A \vee B = B \vee A$ ja $A \wedge B = B \wedge A$.

0.1.3 Huomautus. Disjunkio on “pelkästään tai”. Se ei ole “joko-tai”: esimerkiksi loogisesti väite “Pekka on mies tai marjanpoimija” on totta, vaikka Pekka olisi sekä mies että marjanpoimija.

0.1.4 Määritelmä (Negaatio)**.** *Negaatio* $\neg$ kääntää väitteen: $\neg A$ on totta jos ja vain jos A ei ole totta:

P	$\neg P$
0	1
1	0

0.1.5 Huomautus (Negaation negaatio)**.** Kieltämisen kieltäminen on sallimista: $\neg\neg P = P$.

0.1.6 Esimerkki. Jos P tarkoittaa “kaikki autot ovat punaisia”, niin $\neg P$ tarkoittaa “ei pidä paikkaansa, että kaikki autot ovat punaisia” eli “on olemassa auto, joka ei ole punainen”.

0.1.7 Määritelmä (Implikaatio ja ekvivalenssi)**.** *Implikaatio* $\rightarrow$ tarkoittaa loogista seurausta ja *ekvivalenssi* $\leftrightarrow$ tarkoittaa implikaatioita molempiin suuntiin. Implikaation ja ekvivalenssin totuustaulut ovat (samassa taulussa):

P	Q	$P \rightarrow Q$	$P \leftrightarrow Q$
0	0	1	1
0	1	1	0
1	0	0	0
1	1	1	1

0.1.8 Huomautus (Ekvivalenssi on symmetrinen, implikaatio ei ole)**.** Selvästi $P \leftrightarrow Q = Q \leftrightarrow P$, mutta $P \rightarrow Q \neq Q \rightarrow P$.

0.1.9 Huomautus. Implikaatio $P \to Q$ luetaan "P:stä seuraa Q" tai "P implikoi Q:n" tai "jos P, niin Q" tai "Q on välttämätön P:lle" tai "P on riittävä Q:lle". Itse asiassa loogisesti $P \to Q$ on sama kuin $\neg P \vee Q$! Tämän näkee helposti esimerkiksi totuustaulusta

P	Q	$P \to Q$	$\neg P$	Q	$\neg P \vee Q$
0	0	1	1	0	1
0	1	1	1	1	1
1	0	0	0	0	0
1	1	1	0	1	1

Jos P ja Q ovat väitteitä, niin ekvivalenssi $P \leftrightarrow Q$ tarkoittaa $(P \to Q) \wedge (Q \to P)$. Toisin sanoen $P \leftrightarrow Q$ tarkoittaa, että väite P on totta jos ja vain jos väite Q on totta.

Lopuksi huomattakoon, että $P \to Q$ tarkoittaa "Q jos P" ja $P \leftrightarrow Q$ tarkoittaa "Q jos ja vain jos P".

0.1.10 Esimerkki. Olkoot

$$\begin{aligned} P &= \text{“Pekka rakastaa Liisaa”}, \\ Q &= \text{“Pekka rakastaa Sigríðuriä”}. \end{aligned}$$

Tällöin väite, tai kaava,

$$(P \vee Q) \wedge \neg(P \wedge Q)$$

sanoo "Pekka rakastaa joko Liisaa tai Sigríðuriä".

0.1.11 Lause (De Morganin kaavat)**.** *Disjuktiot ja konjunktiot voidaan negaation avulla vaihtaa toisikseen kaavoilla*

$$\neg(P \wedge Q) = (\neg P) \vee (\neg Q), \tag{0.1.12}$$

$$\neg(P \vee Q) = (\neg P) \wedge (\neg Q). \tag{0.1.13}$$

Todistus. Väitteet on helppo todistaa tarkistamalla kaikki mahdolliset totuusvaihtoehdot totuustaulujen avulla. Todistamme vain kaavan (0.1.12). Kaavan (0.1.13) todistus menee samalla tavalla.

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$(\neg P) \wedge (\neg Q)$
0	0	0	1	1	1	1
0	1	1	0	1	0	0
1	0	1	0	0	1	0
1	1	1	0	0	0	0

Väite seuraa siitä, että neljäs ja viimeinen sarake ovat samoja. □

Predikaattilogiikka

Predikaattilogiikkaan päästään propositiologiikasta, kun väitteisiin lisätään *muuttujia* ja kaavoihin lisätään *kvanttoreita*.

Predikaattilogiikassa väitteessä kaksi termiä: *subjekti* ja *predikaatti*. Subjektista väitetään jokin predikaatti, eli ominaisuus. Esimerkiksi väitteessä “7 on alkuluku” predikaatti “on alkuluku” väitetään subjektista “7”.

0.1.14 Esimerkki. Olkoon

$$P(x) \quad = \quad \text{“}x\text{ on alkuluku”}.$$

Nyt siis x on muuttuja ja P on jokin ominaisuus joka kertoo jotakin x:stä, eli *predikoi* x:ää. Tällöin $P(x)$ on väite, joka voi olla totta tai olla olematta riippuen muuttujan x arvosta. Esimerkiksi $P(3)$ on totta ja $P(4)$ on valhetta.

0.1.15 Määritelmä (Eksistenssikvanttori)**.** *Eksistenssikvanttori* $\exists$ tarkoittaa “on olemassa” tai “jollekin”.

0.1.16 Esimerkki. Jos $P(x)$ on kuten esimerkissä 0.1.14, niin kaava $\exists x P(x)$ tarkoittaa “on olemassa x, jolle $P(x)$ pätee”, eli “on olemassa alkuluku”. Väite on muuten totta ☺.

0.1.17 Määritelmä (Universaalikvanttori)**.** *Universaalikvanttori* $\forall$ tarkoittaa “kaikille” tai “jokaiselle”.

0.1.18 Esimerkki. Jos $P(x)$ on kuten esimerkissä 0.1.14, niin kaava $\forall x P(x)$ tarkoittaa “$P(x)$ pätee kaikille x”, eli “kaikki luvut ovat alkulukuja”. Väite muuten ei ole totta ☹.

0.1.19 Esimerkki. Olkoot

$$\begin{aligned} P(x,y) &= \text{“}x\text{ rakastaa }y\text{:tä,} \\ Q(x) &= \text{“}x\text{:ää rakastetaan.”.} \end{aligned}$$

Tällöin kaava

$$\forall x P(x, \text{Pekka}) \rightarrow \exists z Q(z)$$

sanoo “Jos kaikki rakastavat Pekkaa, niin on olemassa joku, jota rakastetaan.”

Existenssi- ja universaalikvanttori voidaan korvata toisillaan *de Morganin kaavojen* 0.1.11 laajennuksilla:

$$\neg\forall x P(x) \iff \exists x \neg P(x), \tag{0.1.20}$$

$$\neg\exists x P(x) \iff \forall x \neg P(x). \tag{0.1.21}$$

0.1.22 Huomautus (Predikaattilogiikka äärettömänä propositiologiikkana)**.** Jos tarkastelun kohteena oleva "universumi" on äärellinen, eli subjekteja x on olemassa vain äärellinen määrä, niin predikaattilogiikka ja propositiologiikka ovat samoja asioita.

Tällöin esimerkiksi laajennetut de Morganin kaavat (0.1.20)–(0.1.21), jotka pitää äärettömässä universumissa olettaa aksiomaattisesti, voidaan todistaa propositiologiikan de Morganin kaavoista (0.1.12)–(0.1.13). Nimittäin, jos universumissa on vain n subjektia, $x_1, \ldots, x_n$, niin tällöin esimerkiksi kaava (0.1.20) sanoo, että

$$\neg\left(P(x_1) \wedge \cdots \wedge P(x_n)\right) \iff \left(\neg P(x_1)\right) \vee \cdots \vee \left(\neg P(x_n)\right). \tag{0.1.23}$$

Kaava (0.1.23) voidaan johtaa kaavasta (0.1.12) *induktiolla* seuraavasti:

1. Jos $n = 2$, on (0.1.23) sama kuin kaava (0.1.12). Kaava (0.1.23) siis pätee n:n arvolla 2.

2. Teemme sitten *induktio-oletuksen*, että kaava (0.1.23) pätee yleisellä n:n arvolla ja

3. osoitamme, että tästä seuraa, että kaava pätee arvolla $n+1$. Pitää siis osoittaa, että kaava

$$\begin{aligned} &\neg\left(P(x_1) \wedge \cdots \wedge P(x_n) \wedge P(x_{n+1})\right) \\ &\iff \left(\neg P(x_1)\right) \vee \cdots \vee \left(\neg P(x_n)\right) \vee \left(\neg P(x_{n+1})\right) \end{aligned} \tag{0.1.24}$$

pätee. Merkitsemällä

$$\begin{aligned} A &= P(x_1) \wedge \cdots \wedge P(x_n), \\ B &= P(x_{n+1}) \end{aligned}$$

ja käyttämällä de Morganin kaavaa (0.1.12) saamme

$$\begin{aligned} &\neg\left(P(x_1) \wedge \cdots \wedge P(x_n) \wedge P(x_{n+1})\right) \\ &\iff \neg(A \wedge B) \\ &\iff (\neg A) \vee (\neg B) \\ &\iff \left(\neg(P(x_1) \wedge \cdots \wedge P(x_n))\right) \vee \left(\neg P(x_{n+1})\right). \end{aligned}$$

Väite seuraa nyt käyttämällä induktio-oletusta tähän. Nimittäin

$$\begin{aligned} &\left(\neg(P(x_1) \wedge \cdots \wedge P(x_n))\right) \vee \left(\neg P(x_{n+1})\right) \\ &\iff \left((\neg P(x_1)) \vee \cdots \vee (\neg P(x_n))\right) \vee \left(\neg P(x_{n+1})\right) \\ &\iff \left(\neg P(x_1)\right) \vee \cdots \vee \left(\neg P(x_n)\right) \vee \left(\neg P(x_{n+1})\right) \end{aligned}$$

0.2 Joukko-oppi

Matematiikka, ja sitä kautta kaikki tiede, perustuu joukko-oppiin. Kuten logiikkaa, joukko-oppiakaan ei opeteta juurikaan suomalaisessa koululaitoksessa (tosin tähän saattaa olla hyvä syy 70-luvun ylilyönneissä). Korjaamme siis tätäkin epäkohtaa hieman.

0.2.1 Määritelmä (Joukko, kuuluu joukkoon)**.** *Joukko* on “järkevä”[1] kokoelma alkioita. Jos alkio a *kuuluu joukkoon* A, merkitsemme $a \in A$. Jos alkio a ei kuulu joukkoon A, niin merkitsemme $a \notin A$.

0.2.2 Huomautus. Loogisesti siis $a \notin A$ on väite $\neg(a \in A)$ samalla tavalla kuin väite $1 \neq 2$ on väite $\neg(1 = 2)$.

0.2.3 Määritelmä (Inkluusio, osajoukko)**.** Joukko A on joukon B *osajoukko*, jos

$$a \in A \implies a \in B.$$

Tällöin merkitsemme $A \subset B$. Relaatiota $\subset$ kutsutaan *inkluusioksi.*

0.2.4 Huomautus (Osajoukko ja aito osajoukko)**.** Merkintä $\subset$, sekaantuessaan merkintään $<$, antaa ymmärtää hieman liikaa. Nimittäin $A \subset B$ on totta, vaikka olisi $A = B$. Tämän takia joskus käytetäänkin merkinnän $\subset$ sijasta merkintää $\subseteq$. Merkintä $\subsetneq$ tarkoittaa *aitoa osajoukkoa*: $A \subsetneq B$, jos $A \subset B$ ja $A \neq B$.

0.2.5 Määritelmä (Komplementti ja poisto)**.** Joukon A *komplementti* on

$$A^c \;=\; \{a \in \Omega; a \notin A\}.$$

Tässä Ω on *universaalijoukko*, jonka suhteen komplementointi ymmärretään. Joukon B *poisto* joukosta A on joukko

$$A \setminus B \;=\; \{a; a \in A \,\wedge\, a \notin B\}.$$

Tällöin siis $A^c = \Omega \setminus A$.

0.2.6 Määritelmä (Yhdiste ja leikkaus)**.** Joukkojen A_1 ja A_2 *yhdiste* on joukko

$$A_1 \cup A_2 \;=\; \{a; a \in A_1 \,\vee\, a \in A_2\}.$$

Joukkojen A_1 ja A_2 *leikkaus* on joukko

$$A_1 \cap A_2 \;=\; \{a; a \in A_1 \,\wedge\, a \in A_2\}.$$

[1] Kokoelma $\{A; A \notin A\}$ ei ole “järkevä”.

0.2.7 Määritelmä (Yleistetty yhdiste ja leikkaus)**.** Olkoon J jokin joukko, ja olkoon A_j joukko jokaisella $j \in J$. Yleinen *yhdiste* ja *leikkaus* ovat joukkoja

$$\bigcup_{j\in J} A_j = \{a\,;\, \exists\, j \in J : a \in A_j\}\,,$$
$$\bigcap_{j\in J} A_j = \{a\,;\, \forall\, j \in J : a \in A_j\}\,.$$

0.2.8 Huomautus (Yhdisteen ja leikkauksen symmetria)**.** Yhdiste on symmetrinen: $A_1 \cup A_2 = A_2 \cup A_1$. Myös yleistetty yhdiste on symmetrinen:

$$\bigcup_{j\in J} A_j = \bigcup_{j\in J} A_{k_j},$$

missä $(k_j) : J \to J$ on mikä tahansa joukon J *permutaatio* eli uudelleenjärjestys.

Samat symmetriat pätevät myös leikkaukselle ja yleistetylle leikkaukselle.

0.2.9 Huomautus (Muita merkintöjä yleistetylle yhdisteelle ja leikkaukselle)**.** Jos $J = \mathbb{N} = \{1, 2, \ldots\}$, niin merkitsemme myös

$$\bigcup_{j\in J} A_j = \bigcup_{j=1}^{\infty} A_j = A_1 \cup A_2 \cup \cdots\,,$$
$$\bigcap_{j\in J} A_j = \bigcap_{j=1}^{\infty} A_j = A_1 \cap A_2 \cap \cdots\,.$$

Jos $J = \{1, 2, \ldots, n\}$, niin merkitsemme myös

$$\bigcup_{j\in J} A_j = \bigcup_{j=1}^{n} A_j = A_1 \cup \cdots \cup A_n,$$
$$\bigcap_{j\in J} A_j = \bigcap_{j=1}^{n} A_j = A_1 \cap \cdots \cap A_n.$$

0.2.10 Määritelmä (Karteesinen tulo)**.** Jos A ja B ovat joukkoja, niin *karteesinen tulo* $A \times B$ on (järjestettyjen parien) joukko

$$A \times B = \{(a, b)\,;\, a \in A, b \in B\}\,.$$

Käytämme lyhennysmerkintää $A^2 = A \times A$ ja yleisemmin $A^n = A \times \cdots \times A$ (n kertaa).

0.2.11 Huomautus. Tarkalleen ottaen edellisen määritelmän nojalla esimerkiksi joukon $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$ alkiot olisivat muotoa $((x_1, x_2), x_3)$, missä reaalilukupari (x_1, x_2) on paritettu reaaliluvun x_3 kanssa. Yleisesti kuitenkin tehdään ilmiselvä samaistus $((x_1, x_2), x_3) \simeq (x_1, x_2, x_3)$.

0.3 Harjoitustehtäviä lukuun 0

0.1 Harjoitustehtävä. Kaava on *tautologia*, jos se on aina totta. Kaava on *ristiriita*, jos se on aina valhetta. Mitkä seuraavista kaavoista ovat tautologioita tai ristiriitoja?

(a) $P \vee \neg P$.
(b) $(P \to Q) \vee (Q \to P)$.
(c) $\forall x P(x) \vee \exists x P(x)$.
(d) $(\forall x, y) P(x, y) \vee (\exists x, y)(\neg P(x, y))$.
(e) $P \wedge \neg P$.
(f) $(P \to Q) \wedge (Q \to P)$.
(g) $\exists x P(x) \wedge \forall x P(x)$.
(h) $(\forall x, y) P(x, y) \wedge (\exists x, y)(\neg P(x, y))$.

0.2 Harjoitustehtävä. Kirjoita esimerkin 0.1.19 merkinnöillä loogisin kaavoin seuraavat väitteet:

(a) Kaikki rakastavat Pekkaa, mutta Pekka ei rakasta ketään.
(b) Kaikki rakastavat kaikkia.
(c) Kukaan ei rakasta ketään.
(d) Jokaisella on joku, joka rakastaa häntä, mutta on olemassa rakkautta, joka ei ole molemminpuolista.

0.3 Harjoitustehtävä. (a) Looginen seuraus eli implikaatio ja arkikielen syy–seuraus eli kausaliteetti ovat kaksi eri asiaa. Havainnollista tätä keksimällä asiat eli tosiasiaväitteet P ja Q, joille $P \to Q$, mutta P seuraakin Q:sta kausaalisesti.

(b) Kausaliteetti ja implikaatio sekoitetaan usein. Mutta soppaa hämmentää vielä lisäksi *korrelaation* käsite. Karkeasti ottaen korrelaatio tarkoittaa tilastollista samanaikaisuutta: tapahtumat A ja B korreloivat positiivisesti, jos tieto tapahtuman A sattumisesta lisää uskottavuutta tapahtuman B sattumiseen.
Selitä, miksi korrelaatio, implikaatio ja kausaliteetti ovat kaikki täysin eri käsitteitä, mutta miksi korrelaatio on näistä käsitteistä päätöksenteon kannalta kuitenkin keskeisin.

0.4 Harjoitustehtävä. Osoita, että jokainen looginen kaava voidaan kirjoittaa käyttämällä pelkästään symboleja $\exists$, $\neg$ ja $\wedge$. Toisin sanoen symbolit $\forall$, $\to$, $\leftrightarrow$ ja $\vee$ voidaan eliminoida kaavasta.

0.5 Harjoitustehtävä. (a) Osoita, että symboleilla $\vee$ (tai) ja $\wedge$ (ja) voi laskea väitteitä formaalisti täsmälleen samaan tapaan kuin symboleilla $+$ (yhteenlasku) ja $\cdot$ (kertolasku) voi laskea formaalisti lukuja.

(b) Osoita, että symboleilla $\cup$ (yhdiste) ja $\cap$ (leikkaus) voi laskea väitteitä formaalisti täsmälleen samaan tapaan kuin symboleilla $+$ (yhteenlasku) ja $\cdot$ (kertolasku) voi laskea formaalisti lukuja.

0.6 Harjoitustehtävä. Tulkitse seuraavat joukot suorasanaisesti.

(a) $S^2(r) = \{(x, y, z) \in \mathbb{R}^3\,;\, x^2 + y^2 + z^3 = r^2\}$.
(b) $\mathbb{P} = \{n \in \mathbb{N}\,;\, \exists k, n \in \mathbb{N}, k \geq 2 : n = km\}^c$.
(c) $\mathrm{graph}(f) = \{(x, y) \in \mathbb{R}^2\,;\, y = f(x)\}$.
(d) $\mathbb{R}^* = \cup_{n\in\mathbb{N}}\mathbb{R}^n$.

0.7 Harjoitustehtävä. Esitä seuraavat joukot joukko-opillisin merkinnöin:

(a) Parillisten luonnollisten lukujen joukko.
(b) Parittomien alkulukujen joukko.
(c) Epäsymmetristen kääntyvien matiisien joukko.
(d) Niiden reaalilukujen joukko, jotka eivät ole minkään polynomifunktion nolla-kohtia.

0.8 Harjoitustehtävä. (a) Esitä joukko $\cup_{j\in J}A_j$ käyttämällä vain joukkoja A_j, komplementointia ja leikkausta.
(b) Esitä joukko $\cap_{j\in J}A_j$ käyttämällä vain joukkoja A_j, komplementointia ja yhdistämistä.

Vihje: De Morgan.

0.9 Harjoitustehtävä. Miksi $\{A; A \notin A\}$ ei ole joukko?

Vihje: Bertrand Russelin parturi.

Luku 1

Todennäköisyys

Probability is the very guide of life. – Marcus Tullius Cicero

Probability is expectation founded upon partial knowledge. A perfect acquaintance with all the circumstances affecting the occurrence of an event would change expectation into certainty, and leave neither room nor demand for a theory of probabilities. – George Boole

It is remarkable that probability theory, which originated in the consideration of games of chance, should have become the most important object of human knowledge. The most important questions of life are, for the most part, really only problems of probability. – Pierre-Simon Laplace

1.1 Todennäköisyyskäsitteet

Klassinen todennäköisyys

Klassinen todennäköisyys perustuu "yhtä todennäköisen" periaatteelle. Tilannetta jossa esiintyy satunnaisuutta, kutsutaan satunnaiskokeeksi. Satunnaiskokeen eri tulosmahdollisuuksia kutsutaan alkeistapauksiksi. Klassisessa todennäköisyydessä alkeistapauksia on äärellinen määrä ja ne kaikki ovat yhtä mahdollisia eli yhtä todennäköisiä. Tämä olettamus lausutaan sanomalla, että alkeistapaukset ovat symmetrisiä. Esimerkiksi kolikonheitossa on kaksi symmetristä alkeistapausta, kruuna ja klaava, ja nopanheitossa on kuusi symmetristä alkeistapausta, pisteluvut $1, 2, \ldots, 6$.

Tapahtuma on alkeistapausten joukko, erityisesti se voi olla tyhjä ($\varnothing$) tai kaikkien alkeistapausten joukko (Ω). Tapahtumia merkitään kirjaimilla A, B, C, jne., ja alkeistapauksia kirjaimella ω. Esimerkiksi nopanheitossa tapahtuma A voisi olla "nopanheiton tulos on vähintään neljä", siis $A = \{4, 5, 6\}$. Tapahtuma on varma, jos se sattuu välttämättä jokaisessa satunnaiskokeessa, ja mahdoton, jos se ei voi sattua yhdessäkään kokeessa. Nopanheitossa tapahtuma $B =$

"pisteluku on vähintään yksi" $= \Omega$ on varma, ja tapahtuma $C =$ "pisteluvuksi ei tule mitään" $= \varnothing$ on mahdoton.

1.1.1 Määritelmä (Klassinen todennäköisyys)**.** Olkoon n kaikkien alkeistapausten lukumäärä ja $n(A)$ joukon A alkioiden lukumäärä, jota kutsutaan A:lle suotuisien alkeistapausten lukumääräksi. Tapahtuman A *klassinen todennäköisyys* on

$$\mathrm{P}(A) = \frac{n(A)}{n}.$$

1.1.2 Esimerkki. Tapahtuman $A =$ "nopanheiton tulos on vähintään neljä" todennäköisyys on

$$\mathrm{P}(A) = \frac{n(A)}{n} = \frac{3}{6} = \frac{1}{2} = 0{,}5.$$

1.1.3 Huomautus. Alkeistapausten symmetrisyyttä ei voi perustella matemaattisesti, vaan tarvitaan epämääräinen käsite "umpimähkäinen valinta". Alkeistapausten symmetrisyyden voisi yrittää johtaa fysikaalisesta symmetriasta. Esimerkiksi kolikonheitossa kruuna ja klaava ovat symmetrisiä alkeistapauksia, jos kolikkoa ei ole painotettu. Symmetriaa ei voi kuitenkaan perustella sillä, että kolikko olisi fysikaalisesti täysin symmetrinen — silloinhan kruunaa ja klaavaa ei voisi erottaa toisistaan.

Geometrinen todennäköisyys

Symmetrisiin yhtä todennäköisiin tapahtumiin perustuva todennäköisyyden klassinen määritelmä on riittämätön. Yksi tapa laajentaa määritelmää on geometrisen todennäköisyyden idea. Tässäkin lähestymistavassa yhtä todennäköisen käsite on keskeisessä roolissa, mutta geometrista todennäköisyyttä voidaan kuitenkin hyvällä syyllä pitää klassisen todennäköisyyden yleistyksenä — esimerkiksi alkeistapauksia geometrisessa todennäköisyydessä on ääretön määrä.

Geometrista todennäköisyyttä voi soveltaa tilanteissa, joissa satunnaiskokeen tulos voidaan havainnollistaa geometrisella kuviolla ja kiinnostuksen kohteena oleva tapahtuma A tämän osakuviona. Tällaisia kuvioita ja niiden osakuvioita voivat olla esimerkiksi yksiulotteinen jana, kaksiulotteinen tasoalue tai kolmiulotteinen kappale. Tilanteen on oltava siinä mielessä symmetrinen, että A:n mahdollisuus esiintyä riippuu vain A:n geometrisesta mitasta (janalla pituus, tasoalueella pinta-ala ja kappaleella tilavuus), eikä lainkaan A:n muodosta tai sijainnista.

1.1.4 Määritelmä (Geometrinen todennäköisyys)**.** Tapahtuman A *geometrinen todennäköisyys* on

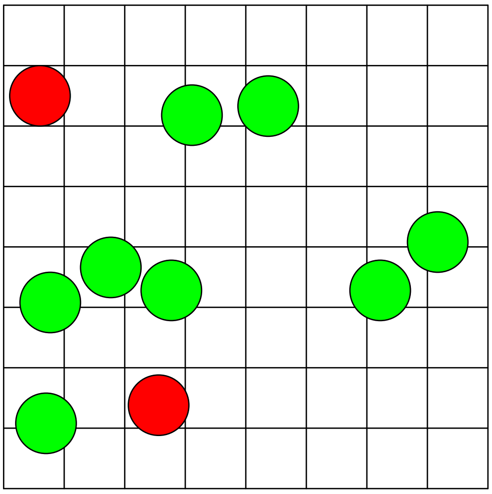
$$\mathrm{P}(A) \quad = \quad \frac{m(A)}{m},$$

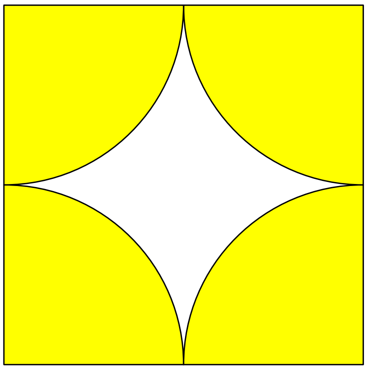
missä $m(A)$ on tapahtumaa A vastaavaan kuvion ja m koko kuvion geometrinen mitta.

Seuraava esimerkki 1.1.5, yhdistettynä suurten lukujen lakiin 1.3.19 ja seuraavaksi esitettävään todennäköisyyden frekvenssitulkintaan 1.1.7, on kuuluisa *Buffonin neulakoe*, jolla voidaan laskea π tilastollisesti.

1.1.5 Esimerkki. Lattialla on neliöruudukko, jossa neliön sivu = kolikon halkaisija = $2r$. Millä todennäköisyydellä lattialle heitetty kolikko peittää neliön kärjen?

Tutkimme kysytyn geometrisen todennäköisyyden selvittämiseksi kolikon keskipisteen sijaintia neliöruudukossa. Koska eri neliöt ovat toisiinsa nähden samassa asemassa, voimme tarkastella yhtä neliötä. Sen pinta-ala on $m = (2r)^2 = 4r^2$. Tarkastelemme tapahtumaa A = "lattialle heitetty kolikko peittää neliön kärjen", jota mallissamme edustaa kolikon keskipisteen sijainti neliössä. Suotuisissa tapauksissa kolikon keskipisteen etäisyys neliön kärjestä on pienempi kuin r





Näin ollen A:n pinta-ala on $m(A) = 4 \cdot \frac{\pi r^2}{4} = \pi r^2$, ja siten

$$\mathrm{P}(A) = \frac{\pi r^2}{4r^2} = \frac{\pi}{4} \approx 0{,}785.$$

1.1.6 Huomautus. Geometrisen todennäköisyyden määrittelyyn liittyy samoja periaatteellisia ongelmia kuin klassisenkin todennäköisyyden määrittelyyn. Vakavin puute kummassakin määritelmässä on, että ne kattavat vain hyvin suppean osan niistä satunnaiskokeista, joista olemme kiinnostuneet. Kummankaan määritelmän pohjalta on mahdotonta rakentaa alkeistapauksia, joiden avulla voisimme johtaa todennäköisyyden, että syntyvä lapsi on tyttö (0,487) tai että radioaktiivisen hiiliatomin ^{14}C elinikä on yli 1.000 vuotta (0,883).

Frekventistinen todennäköisyys

Perinteinen tilastollisen todennäköisyyden käsite perustuu frekvenssitulkintaan, joka puolestaan perustuu suurten lukujen lakiin.

Tarkastelemme satunnaiskoetta, jota voidaan toistaa samanlaisissa olosuhteissa rajattomasti. Olkoon A tähän kokeeseen liittyvä tapahtuma ja $F_n(A)$ tapahtuman A esiintymiskertojen lukumäärä n:ssä toistossa.

1.1.7 Määritelmä (Frekventistinen todennäköisyys)**.** Tapahtuma A:n *suhteellisen frekvenssi* on

$$f_n(A) \quad = \quad \frac{F_n(A)}{n}.$$

Suurten lukujen lain 1.3.19 nojalla toistojen lukumäärän n kasvaessa suhteellinen frekvenssi $f_n(A)$ keskittyy tietyn luvun läheisyyteen. Tapahtuman A *frekventistinen todennäköisyys* on juuri kyseinen luku:

$$\mathrm{P}(A) = \lim_{n\to\infty} f_n(A).$$

1.1.8 Esimerkki. Huomautuksen 1.1.6 tytön syntymän todennäköisyys 0,487 on saatu siitä, että historiallisesti jokaisesta tuhannesta syntyneestä lapsesta keskimäärin 487 on ollut tyttöjä. Radioaktiivisia hiiliatomeja tarkasteltaessa taas havaittiin, että keskimäärin 833 hiiliatomia tuhannesta saavutti kunnioitettavan tuhannen vuoden iän ☺.

1.1.9 Huomautus. Frekventistisen todennäköisyyden käsitteeseen liittyy useita filosofisia ongelmia. Von Mises esimerkiksi katsoi, että todennäköisyydet liittyvät vain ja ainoastaan äärettömiin toistokokeisiin. Tämä näkemys, vaikkakin filosofisesti hygieeninen, tekee todennäköisyydestä käytännön kannalta merkityksettömän. Käytännössä äärettömiä toistokokeita ei ole. Sen sijaan tulee tyytyä äärellisiin approksimaatioihin.

Bayesläinen todennäköisyys

Bayesläinen todennäköisyystulkinta on nykyään varmaankin suosituin eri todennäköisyystulkinnoista — tosin aikanaan sitä pidettiin yleisesti epäilyttävänä.

1.1.10 Määritelmä (Bayeläinen todennäköisyys)**.** *Bayesläisessä tulkinnassa* todennäköisyys on *uskomuksen aste* ja se koskee väitteitä. Toisin sanoen väitteen todennäköisyys on sitä lähempänä ykköstä, mitä uskottavampana sitä pidetään.

1.1.11 Huomautus. Määritelmä 1.1.10 on selvästi hämärä. Siitä onkin olemassa monia enemmän tai vähemmän hämäriä muotoiluja riippuen siitä, millainen on uskomisen subjekti.

Subjektiivisessa muodossa bayesläinen todennäköisyys on jokaiselle uskojalle omansa: kaksi eri henkilöä voi päätyä samoilla tiedoilla eri todennäköisyyksiin. Tämä on De Finettin todennäköisyyden filosofinen perusta ja sitä voitaneen pitää vähintäänkin yhtä epäonnistuneena kuin Von Misesin filosofiaa.

Objektiivisen tulkinnan mukaan “rationaalisten uskojien” tulee samalla informaatiolla päätyä samoihin todennäköisyyksiin. Tähänkin tulkintaan liittyy ongelmia. Suurin ongelma lienee se, että kukaan ei ole koskaan havainnut “rationaalista uskojaa”.

Käytännössä bayesläinen todennäköisyyslaskenta perustuu *priorin*, *uskottavuuden* ja *posteriorin*[1] yhdistämiseen *Bayesin kaavalla*: henkilöllä on priorinäkemys $\mathrm{P}(A)$ tapahtuman A todennäköisyydestä. Sitten sattuu tapahtuma E (evidenssi). Tällöin henkilön on päivitettävä priorinäkemyksensä $\mathrm{P}(A)$ posteriorinäkemyksekseksi $\mathrm{P}(A|E)$, missä tapahtuman E sattuminen on otettava huomioon. Päivittäminen onnistuu, jos henkilö pystyy määrittämään tapahtuman E uskottavuuden $\mathrm{P}(E|A)$ priorinäkemyksensä $\mathrm{P}(A)$ valossa. Nimittäin tällöin Bayesin kaava sanoo, että

$$\begin{aligned}\mathrm{P}(A|E) &= \frac{\mathrm{P}(A)\mathrm{P}(E|A)}{\mathrm{P}(E)} \\ &= \frac{\mathrm{P}(A)\mathrm{P}(E|A)}{\mathrm{P}(A)\mathrm{P}(E|A)+\mathrm{P}(A^c)\mathrm{P}(E|A^c)}.\end{aligned}$$

Todistamme Bayesin kaavan myöhemmin tässä luvussa.

1.1.12 Esimerkki. Mitkä todennäköisyyskäsitteistä *klassinen*, *geometrinen*, *frekventistinen* tai *bayesläinen* (sen eri muodoissaan) sopivat seuraaviin todennäköisyyttä koskeviin väitteisiin?

(a) Todennäköisyys että nopanheitossa saadaan silmäluku 5 on $1/6$.
(b) Todennäköisyys että nopanheitossa saadaan silmäluku 5 on $1/5$.
(c) Dinosaurukset kuolivat sukupuuttoon todennäköisesti meteorin törmättyä maahan.
(d) Todennäköisyys että Kokoomus on Suomen suurin puolue vuoden 2008 kuntavaaleissa on 99,9%.
(e) Kolikkoa on heitetty 7 kertaa ja joka kerralla saatiin klaava. Todennäköisyys saada klaava seuraavallakin heitolla on $8/9$.
(f) Kolikkoa on heitetty 7 kertaa ja joka kerralla saatiin klaava. Todennäköisyys saada klaava seuraavallakin heitolla on 1.
(g) Kolikkoa on heitetty 7 kertaa ja joka kerralla saatiin klaava. Todennäköisyys saada klaava seuraavallakin heitolla on $1/2$.
(h) Todennäköisyys sille, että tikanheittäjä N.N. osuu täsmälleen napakymppiin seuraavalla heitolla on 0.

Kohtaan (a) soveltuu mitä ilmeisimmin symmetriaan perustuva *klassinen* tulkinta: nopassa on kuusi sivua ja noppa on symmetrinen. Siten jokainen sivu on yhtä todennäköinen, ja silmäluku 5 vastaa yhtä sivua. Siten kysytty

[1] ‘Priori’ tarkoittaa ‘ennen’, ja ‘posteriori’ tarkoittaa ‘jälkeen’.

todennäköisyys on $1/6$. Kohtaan (a) sopii kyllä myös bayesläinenkin tulkinta. Bayesläisen ajatus voisi mennä vaikka näin: “Minulla ei ole mitään muuta tietoa nopasta, kun että siinä on kuusi tahkoa. Otanpa käyttöön — ilman parempaa tietoa tulosmahdollisuuksien mahdollisista eroista — *tasaisen priorin*, eli ajattelen että jokainen tulosmahdollisuus on yhtä todennäköinen.”

Kohtaan (b) klassinen tulkinta ei oikein sovi, ellei kyseessä ole symmetrinen viisisivuinen noppa (onko sellaisia). Sen sijaan *frekventistinen* ja *bayesläinen* tulkinta tulevat kysymykseen. Frekventistinen tulkinta voi nousta esimerkiksi tapauksessa, jossa ko. noppaa on heitetty 15 kertaa ja 3 kertaa on saatu silmäluku 5. Vaikka saatu tulos on täysin mahdollinen — eikä edes tavattoman harvinainen — symmetrisen nopan tapauksessa antaa saatu tulos frekventistille pienen epäilyksen siitä, että noppa on painotettu. Bayesläinen tulkinta on voi nousta esimerkiksi yksinkertaisesti siitä syystä että valitaan erilainen priori kuin kohdassa (a). Itse asiassa bayesläinen tulkinta, varsinkin sen subjektiivinen muoto, on niin yleinen, että se sopii käytännössä kaikkiin mahdollisiin tapauksiin.

Kohtaan (c) on vaikea kuvitella muuta kuin *bayesläistä* tulkintaa, joka sopii kaikkiin tilanteisiin.

Kohta (d) ei juurikaan kaipaisi tulkintaa, jos väitetty todennäköisyys olisi 100%. Kyseessähän on mennyt tapahtuma, jonka tiedämme varmasti sattuneeksi. Ainoa tapa selittää todennäköisyys 99,9% on subjektiivinen *bayesläinen* tapa. Bayesläinen voi esimerkiksi ajatella seuraavasti: “On 0,1% todennäköisyys että Kokoomus ei oikeasti ollutkaan vaalien suurin puolue, vaan että me kaikki elämme jonkinlaisessa harhassa.”

Kohtaan (e) sopii *bayesläinen* tulkinta. Itse asiassa esitetty luku $8/9$ tulee nimenomaan bayesläisestä tavasta, kun käytetään “tasaista prioria” (luennoija selittää laskut pyynnöstä).

Kohtaan (f) sopii (ääri)*frekventistinen* tulkinta. Frekventisti on nähnyt 7 toiston toistokokeen, jossa jokaisessa on tullut klaava. Siis tapahtuma “seuraavallakin heitollatulee klaava” on frekventistin mielestä varma. Toki frekventisti ymmärtää, että 7 toistoa on vielä kohtalaisen pieni määrä — erityisesti se on paljon pienempi kuin $+\infty$, jonka frekventisti tarvitsee voidakseen täydellä varmuudella määrätä todennäköisyydet. Siispä frekventistin mielestä tapahtuma “seuraavallakin heitolla tulee klaava” on varma, mutta frekventisti ei ole kovin varma siitä.

Kohtaan (g) sopii *klassinen* tulkinta. Klassisisti ajattelee, että kolikko on symmetrinen ja siten todennäköisyys saada klaava on aina $1/2$. Saatu data — 7 klaavaa, eikä yhtään kruunaa — ei horjuttanut klassisistin uskoa kolikon symmetrisyyteen. Ehkä hän ajattelee seuraavasti: “Jos kolikko on reilu, eli klaavan todennäköisyys on $1/2$, niin todennäköisyys saada 7 klaavaa putkeen on $(1/2)^7 = 0{,}78\%$. Tämä on toki harvinaista, mutta ei erityisen epäuskottavaa.”

Kohtaan (h) sopii *geometrinen* tulkinta sekä — kuten aina — *bayesläinen* tulkinta.

1.2 Todennäköisyyden laskusäännöt

Todennäköisyyden aksioomat

Olemme törmänneet vähintäänkin neljään eri todennäköisyystulkintaan. Nämä eri tulkinnat eivät ole kaikin puolin täysin yhteensovitettavissa. Mutta, oli todennäköisyyden tulkinta sitten mikä tahansa, sen on toteutettava seuraavat *Kolmogorovin aksioomat*:

1.2.1 Määritelmä (Todennäköisyyden aksioomat)**.** Joukkofunktio P on todennäköisyys, jos pätee

(i) $\mathrm{P}(A) \geq 0$ kaikilla tapahtumilla A.
(ii) $\mathrm{P}(\Omega) = 1$.
(iii) Jos tapahtumat A_k, $k \in \mathbb{N}$, ovat erillisiä, eli $A_i \cap A_j = \varnothing$ kun $i \neq j$, niin silloin
$$\mathrm{P}\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mathrm{P}\left(A_k\right).$$

1.2.2 Huomautus. Todennäköisyyden määritelmässä 1.2.1 aksioomat 1.2.1(i) (positiivisuus) ja 1.2.1(ii) (äärellisyys) ovat ilmiselviä. Sen sijaan kolmas aksiooma 1.2.1(iii) (*täys-* tai *sigma-additiivisuus*) on epätriviaali, jopa kiistanalainen. Se on kuitenkin välttämätön esimerkiksi todennäköisyyden frekvenssitulkintaa 1.1.7 varten.

Lause 1.2.3 esittää muutamia todennäköisyyden laskusääntöjä, jotka seuraavat loogisesti Kolmogorovin aksioomista.

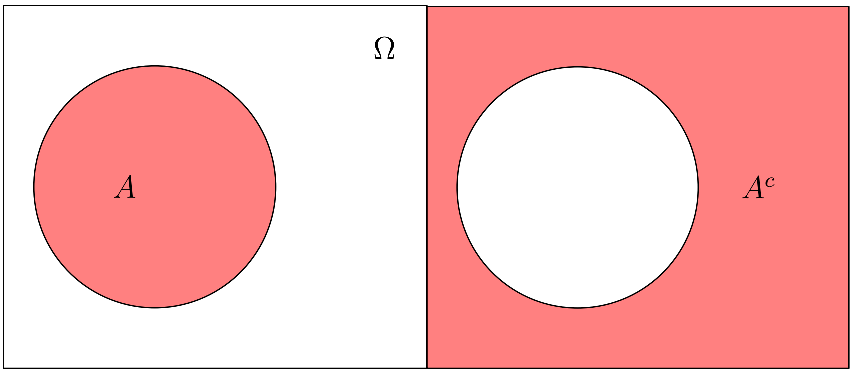
1.2.3 Lause (Todennäköisyyden laskusääntöjä)**.** *Olkoot* $A, B, A_1, \ldots, A_n$ *tapahtumia. Tällöin*

(i) $\mathrm{P}(A^c) = 1 - \mathrm{P}(A)$,
(ii) *jos* $A \subset B$, *niin* $\mathrm{P}(B \setminus A) = \mathrm{P}(B) - \mathrm{P}(A)$,
(iii) $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$,
(iv)
$$\begin{aligned}
\mathrm{P}(A_1 \cup \cdots \cup A_n) &= \sum_{i\,:\,1\leq i\leq n} \mathrm{P}(A_i) - \sum\sum_{i,j\,:\,1\leq i<j\leq n} \mathrm{P}(A_i \cap A_j) \\
&\quad + \sum\sum\sum_{i,j,k\,:\,1\leq i<j<k\leq n} \mathrm{P}(A_i \cap A_j \cap A_k) \\
&\quad - \sum\sum\sum\sum_{i,j,k,\ell\,:\,1\leq i<j<k<\ell\leq n} \mathrm{P}(A_i \cap A_j \cap A_k \cap A_\ell) \\
&\quad\quad \vdots \\
&\quad +(-1)^{n-1}\mathrm{P}(A_1 \cap \cdots \cap A_n).
\end{aligned}$$

1.2.4 Huomautus. Kohtaa (i) kutsutaan *komplementtikaavaksi*, kohtaa (ii) kutsutaan *vähenyskaavaksi*, kohtaa (iii) kutsutaan *summakaavaksi* ja kohtaa (iv) kutsutaan *yleiseksi summakaavaksi*, *inkluusio–ekskluusioperiaatteeksi* tai *seulaperiaatteeksi*.

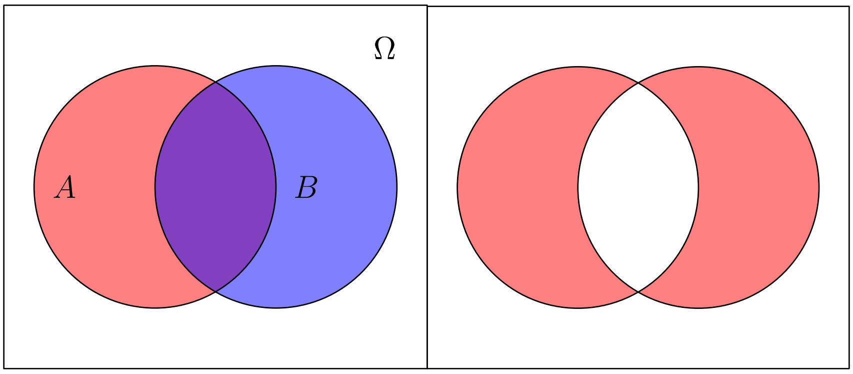
Ennen todistusta havainnollistamme lauseen 1.2.3 laskusääntöjä *Venn-diagrammien* avulla.

Kaavan 1.2.3(i) tapahtuman A *komplementille* näkee kuvasta



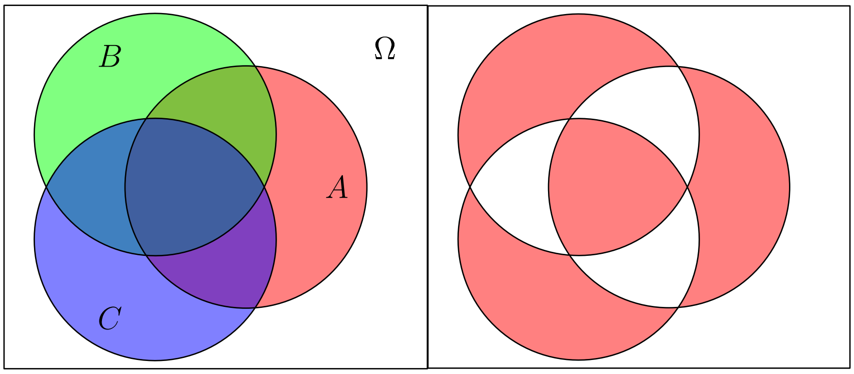
Kaavan 1.2.3(ii) näkee edellisestä kuvasta tulkitsemalla tapahtuman B koko avaruudeksi Ω.

Kahden tapahtuman A ja B summakaavan 1.2.3(iii) näkee kuvasta



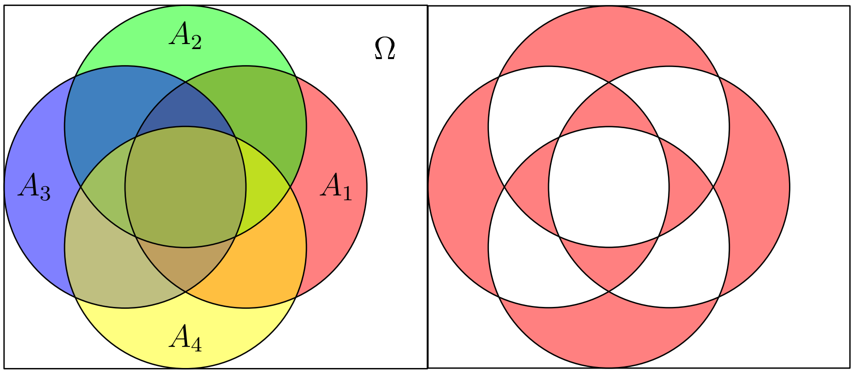
Yleisen summakaavan 1.2.3(iv) ymmärtämiseksi havainnollistamme aluksi kolmen tapahtuman A, B ja C tilannetta:

$$\begin{aligned} \mathrm{P}(A \cup B \cup C) &= \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C) \\ &\quad -\mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) \\ &\quad +\mathrm{P}(A \cap B \cap C). \end{aligned}$$



Neljälle tapahtumalle A_1, A_2, A_3 ja A_4 kaava 1.2.3(iv) sanoo:

$$
\begin{aligned}
&\mathrm{P}(A_1 \cup A_2 \cup A_3 \cup A_4) \\
&\quad = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \mathrm{P}(A_3) + \mathrm{P}(A_4) \\
&\qquad - \mathrm{P}(A_1 \cap A_2) - \mathrm{P}(A_1 \cap A_3) - \mathrm{P}(A_1 \cap A_4) \\
&\qquad\qquad - \mathrm{P}(A_2 \cap A_3) - \mathrm{P}(A_2 \cap A_4) - \mathrm{P}(A_3 \cap A_4) \\
&\qquad + \mathrm{P}(A_1 \cap A_2 \cap A_3) + \mathrm{P}(A_1 \cap A_2 \cap A_4) \\
&\qquad\qquad + \mathrm{P}(A_1 \cap A_3 \cap A_4) + \mathrm{P}(A_2 \cap A_3 \cap A_4) \\
&\qquad - \mathrm{P}(A_1 \cap A_2 \cap A_3 \cap A_4).
\end{aligned}
$$



Yleinen summakaava 1.2.3(iv) on nyt helppo arvata edellisistä kuvista.

Lauseen 1.2.3 todistus. (i) Koska $\Omega = A \cup A^c$ ja $A \cap A^c = \varnothing$, niin todennäköisyyden äärellisyydestä ja täysadditiivisuudesta seuraa, että

$$
\begin{aligned}
1 &= \mathrm{P}(\Omega) \\
&= \mathrm{P}(A \cup A^c) \\
&= \mathrm{P}(A) + \mathrm{P}(A^c).
\end{aligned}
$$

Väite seuraa tästä vähentämällä $\mathrm{P}(A^c)$ puolittain.

(ii) Koska $A \subset B$, niin A on erillinen yhdiste

$$B = (B \setminus A) \cup A.$$

Siten, aksiooman 1.2.1(iii) nojalla,

$$\mathrm{P}(B) = \mathrm{P}(B \setminus A) + \mathrm{P}(A).$$

Väite seuraa tästä vähentämällä $\mathrm{P}(A)$ puolittain.

(iii) Yhdiste $A \cup B$ voidaan esittää erillisenä yhdisteenä

$$(A \setminus B) \cup (A \cap B) \cup (B \setminus A).$$

Siten, käyttämällä aksioomaa 1.2.1(iii), saamme

$$\mathrm{P}(A \cup B) = \mathrm{P}(A \setminus B) + \mathrm{P}(A \cap B) + \mathrm{P}(B \setminus A).$$

Toisaalta harjoitustehtävän 1.3 nojalla

$$\begin{aligned}
\mathrm{P}(A \setminus B) &= \mathrm{P}(A) - \mathrm{P}(A \cap B), \\
\mathrm{P}(B \setminus A) &= \mathrm{P}(B) - \mathrm{P}(A \cap B).
\end{aligned}$$

Siten

$$\begin{aligned}
\mathrm{P}(A \cup B) &= \mathrm{P}(A \setminus B) \\
&\quad + \mathrm{P}(A \cap B) \\
&\quad + \mathrm{P}(B \setminus A) \\
&= \mathrm{P}(A) - \mathrm{P}(A \cap B) \\
&\quad + \mathrm{P}(A \cap B) \\
&\quad + \mathrm{P}(B) - \mathrm{P}(A \cap B) \\
&= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).
\end{aligned}$$

(iv) Harjoitustehtävä 1.4. □

1.2.5 Esimerkki. Uurnassa on 3 valkoista palloa ja 7 mustaa palloa. Uurnasta nostetaan 2 palloa. Mikä on todennäköisyys, että molemmat nostetut pallot ovat valkoisia, kun

(a) pallot nostetaan samanaikaisesti,
(b) ensin nostetaan yksi pallo, tarkistetaan sen väri, palautetaan pallo uurnaan, ja sitten nostetaan toinen pallo.

Ratkaisemme esimerkin 1.2.5 ensin viittaamatta yleiseen (kombinatoriseen) teoriaan, ja sitten kerromme yleisestä kombinatorisesta viitekehyksestä.

Kohdassa (a) symmetriset alkeistapaukset ovat kaikki 2 pallon nostot $3+7=10$ pallon joukosta. Tulkitsemme jokaisen noston yhtä todennäköiseksi. Eri tapoja nostaa 2 palloa 10 pallon joukosta on 45 kappaletta. Nimittäin voimme *kuvitella* nostot kahdessa vaiheessa. Ensiksi nostamme yhden pallon. Meillä on 10 eri tapaa tehdä tämä, sillä palloja on 10. Sitten nostamme toisen pallon. Meillä on 9 eri tapaa tehdä tämä, sillä uurnassa on enää jäljellä 9 palloa. Siis jokaista 10 ensimmäistä nostoa kohden meillä on 9 toista nostoa. Yhteensä tämä tekee $10{\cdot}9 = 90$ erilaista nostoa. Koska kuitenkin perättäiset nostot olivat kuviteltuja, emme voi erottaa ensimmäistä ja toista nostoa toisistaan. Siten nosto “ensin a ja sitten b” näyttää samalta kuin nosto “ensin b ja sitten a”. Näinollen eri nostoja on oikeasti vain $90/2=45$ kappaletta. Entä sitten suotuisat nostot? Kuinka monella tavalla voimme nostaa 2 valkoista palloa? *Kuvittelemalla* tilanne toistokokeeksi näemme, että ensimmäisellä nostolla uurnassa on 3 valkoista palloa. Voimme siis nostaa valkoisen pallon kolmella eri tavalla. Seuraavalla nostolla uurnassa on kaksi valkoista palloa jäljellä, sillä olemme jo nostaneet toisen valkoisen pallon sieltä. Siten toisella nostolla on siis kaksi nostaa valkoinen pallo. Yhteensä tämä tekee $3\cdot 2=6$ tapaa nostaa kaksi valkoista palloa toistokokeessa. Koska kyseessä ei kuitenkaan ollut toistokoe joudumme jakamaan tuloksen samalta näyttävien toistojen lukumäärällä. Saamme $6/2=3$ tapaa. Siten kysytty todennäköisyys on

$$\frac{3}{45} \quad = \quad \frac{1}{15} \quad = \quad 6{,}667\%$$

Kohdassa (b) voimme ajatella tilannetta toistokokeena. Nimittäin koska nostettu pallo palautetaan takaisin uurnaan, pysyy uurna samanlaisena toiston jälkeen. Nyt ensimmäisellä nostolla todennäköisyys saada valkoinen pallo on $3/10$, sillä tapoja nostaa valkoinen pallo on 3 ja tapoja nostaa jokin pallo on $3+7=10$. Klassisen tulkinnan mukaan jokainen nosto on yhtä todennäköinen. Toisella nostolla tilanne on sama. Nyt ainoa tapa nostaa kaksi valkoista palloa on, että molemmilla nostoilla nostetaan valkoinen pallo. Siten kysytty todennäköisyys on, riippumattomien toistojen periaatteen nojalla,

$$\frac{3}{10}\cdot\frac{3}{10} \quad = \quad \frac{9}{100} \quad = \quad 9\%.$$

Tarkastelemme sitten kohtia (a) ja (b) seuraavassa yleistetyssä viitekehyksessä: uurnassa on v valkoista palloa ja m mustaa palloa, nostamme n palloa, ja kysytty todennäköisyys on

> “Mikä on todennäköisyys, että nostamme k valkoista palloa?”

On selvää, että jos $k > v$, niin kysytty todennäköisyys on nolla. Siten jatkossa oletamme, että k on joko $0,\ 1,\ 2,\ \ldots,\ v-1$ tai v.

Kohdan (a) yleinen tarkastelu: Ensiksi meidän on laskettava kaikkien mahdollisten alkeistapausten lukumäärä. Eli kuinka monta tapaa on nostaa n palloa joukosta, jossa on $m+v$ palloa? Tälle kombinatoriselle suureelle käytetään merkintää

$$\binom{m+v}{n}$$

ja se lausutaan “$m+v$ yli n:n”. Tämä ei tietysti kerro paljoakaan siitä, miten ko. lukumäärä *lasketaan*. Kerromme nyt miten laskeminen onnistuu *kuvittelemalla*, että nostot tehdään peräkkäin toistokokeena. Aluksi voimme nostaa pallon $(m+v)$:llä eri tavalla. Seuraavaksi, nostolla nro 2, voimme nostaa pallon $(m+v-1)$:llä eri tavalla, sillä uurnassa on nyt $m+v-1$ palloa jäljellä. Nostolla nro 3 voimme nostaa pallon $(m+v-2)$:lla eri tavalla, jne. Siten n:llä nostolla voimme nostaa palloja

$$(v+m)(v+m\ -\ 1)(v+m\ -\ 2)\cdots(v+m\ -\ (n-1))$$

eri tavalla. Nyt kuitenkin pitää muistaa, että toistokoe oli puhtaasti kuviteltu. Emme näe n:n pituista jonoa nostoja, vaan n nostettua palloa. Siten jokainen saatu n:n pallon kokoelma voi vastata mitä tahansa tapaa laittaa n palloa jonoon. Eri tapoja laittaa n palloa jonoon on

$$n(n-1)(n-2)\cdots 2\cdot 1$$

kappaletta. Nimittäin ensimmäiseksi palloksi jonoon voidaan valita mikä tahansa n:stä eri pallosta. Toiseksi palloksi voidaan valita mikä tahansa $(n-1)$:stä jäljellä olevasta pallosta, jne. Tämä siis tarkoittaa, että[2]

$$\binom{m+v}{n} \quad = \quad \frac{(v+m)(v+m-1)(v+m-2)\cdots(v+m-n+1)}{n(n-1)(n-2)\cdots 2\cdot 1}. \tag{1.2.6}$$

Nyt tiedämme, miten $\binom{m+v}{n}$, joita myös *binomikertoimiksi* kutsutaan, lasketaan. Nyt siis tiedämme miten monta alkeistapausta on. Entäpä sitten suotuisat alkeistapaukset? Kuinka monta tapaa on valita k valkoista palloa, kun nostoja on n kappaletta, valkoisia palloja on v kappaletta, ja kaikkiaan palloja on $m+v$ kappalatta? Vastaus kysymykseen saadaan kuvittelemalla toisto: kuvittelemme, että ensiksi nostamme k valkoista palloa, ja sitten nostamme $n-k$ mustaa palloa. Koska jo tiedämme, kuinka monella eri tavalla nämä kaksi nostoa voidaan

[2] Usein näkee kaavaa

$$\binom{m+v}{n} \quad = \quad \frac{(m+v)!}{(m+v-n)!\,n!},$$

missä $n!$ on n:n *kertoma*:

$$n! \quad = \quad n\cdot(n-1)\cdot(n-2)\cdots 3\cdot 2\cdot 1.$$

Tämä on vain toinen tapa esittää kaava (1.2.6).

tehdä, voimme laskea kuinka monella tavalla nämä kaksi perättäistä nostoa voidaan tehdä:

$$\binom{v}{k}\binom{m}{n-k}.$$

Siispä vastaus kysymykseen

"Mikä on todennäköisyys, että nostamme k valkoista palloa?"

on

$$\binom{v}{k}\binom{m}{n-k}\bigg/\binom{m+v}{n}. \tag{1.2.7}$$

Todennäköisyyksiä (1.2.7) kutsutaan *hypergeometriseksi jakaumaksi* parametrein $m+v$, v ja n ja se liittyy *otantaan ilman takaisinpanoa*.

Kohdan (b) yleinen tarkastelu: Tarkastelemme sitten tilannetta, jossa nostettu pallo palautetaan uurnaan noston jälkeen. Nyt tilanne on aito toistokoe. Jokaisella toistolla on symmetrian perusteella todennäköisyys $v/(m+v)$ nostaa valkoinen pallo. Siten todennäköisyys nostaa täsmälleen k valkoista palloa on

$$\binom{n}{k}\left(\frac{v}{m+v}\right)^k\left(\frac{m}{m+v}\right)^{n-k}. \tag{1.2.8}$$

Nimittäin joka ikisen n:n mittaisen jonon, jossa on täsmälleen k valkoista palloa todennäköisyys on

$$\left(\frac{v}{m+v}\right)^k\left(\frac{m}{m+v}\right)^{n-k},$$

ja erilaisia n:n mittaisia jonoja, joissa on täsmälleen k valkoista palloa on $\binom{n}{k}$ kappaletta.

Todennäköisyyksiä (1.2.8) kutsutaan *binomijakaumaksi* parametrein n ja $v/(m+v)$ ja se liittyy *otantaan takaisinpanolla*.

1.2.9 Esimerkki. [3] Marilyn vos Savant osallistuu seuraavaan peliin: Pelissä on kolme ovea A, B ja C. Yhden oven takana on 10.000 euroa, jonka Marilyn saa, jos arvaa oven oikein. Kahden muun oven takana ei ole mitään.

Marilyn valitsee oven A. Nyt peluuttaja Monty Hall avaa oven B, jonka takana ei ollut mitään. Monty Hall antaa Marilynille mahdollisuuden vaihtaa valitsemansa oven A oveksi C.

Kannattaako Marilynin vaihtaa ovea?

Entä kannattaako Marilynin vaihtaa siinä tapauksessa, että ovia on 29 kappaletta: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, Å, Ä, Ö, ja Monty Hall on avannut kaikki muut ovat paitsi oven C (ja tietysti oven A)?

[3] Tämä on kuuluisa Monty Hall -ongelma. Sen värikkäästä historiasta löytyy tietoa sivulta www.marilynvossavant.com/articles/gameshow.html.

Tarkastelemme aluksi kolmen oven tapausta. Käytämme oville nimiä 1, 2 ja 3, ja oletamme, että voitto-ovi on ovi 3. Marilyn ei luonnollisestikaan tiedä mitkä ovista A, B ja C ovat ovia 1, 2, ja 3. Mahdollisia tilanteita on nyt kolme:

1 Marilyn on alun perin valinnut tyhjän oven numero 1. Hänelle avataan tyhjä ovi numero 2.
2 Marilyn on alun perin valinnut tyhjän oven numero 2. Hänelle avataan tyhjä ovi numero 1.
3 Marilyn on valinnut voitto-oven numero 3, ja hänelle avataan jompikumpi tyhjistä ovista 1 tai 2.

Jokainen näistä kolmesta tilanteesta on symmetrian perusteella yhtä todennäköinen. Jos Marilyn päättää vaihtaa ovensa, hän ei valitse tietämäänsä tyhjää ovea, jolloin tilanne 1 ja tilanne 2 johtavat molemmat siihen että hän voittaa. Tilanteessa 3 hän kuitenkin menettää voittonsa. Ilman vaihtoa Marilynin voittomahdollisuus on siksi $1/3$ ja vaihdon jälkeen $2/3$.

Tarkastelemme sitten 29 oven tapausta. Tilanne on olennaisesti sama, kuin 3 oven tapauksessa, mutta radikaalimpi. Nimittäin nyt symmetrisiä tilanteita on 29 kappaletta

1 Marilyn on alun perin valinnut tyhjän oven numero 1. Hänelle avataan tyhjät ovet $2 \ldots 28$.
2 Marilyn on alun perin valinnut tyhjän oven numero 2. Hänelle avataan tyhjät ovet $1,\ 3 \ldots 28$.
n $(n = 3, \ldots, 28)$ Marilyn on alun perin valinnut tyhjän oven numero n. Hänelle avataan tyhjät ovet $2 \ldots (n-1),\ (n+1) \ldots 28$.
29 Marilyn on valinnut voitto-oven numero 29, ja hänelle näytetään joku 27 oven kokoelma tyhjistä 28 ovesta.

Siten Marilynin kannattaa vaihtaa, sillä vain tapauksessa numero 29 hän häviää, eli todennäköisyys että palkinto on vaihdetun oven takana on $28/29$.

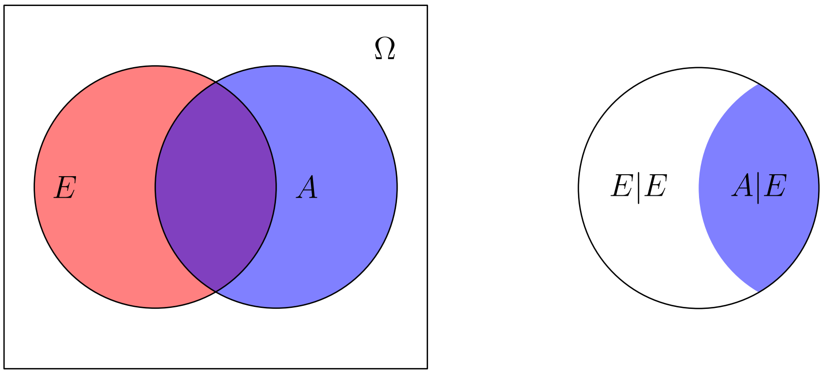
Ehdollinen todennäköisyys

Tapahtuman A ehdollinen todennäköisyys on todennäköisyys tapahtumalle A sillä ehdolla, että jokin tapahtuma E *tiedetään* tapahtuneen (tai tapahtuvan).

1.2.10 Määritelmä (Ehdollinen todennäköisyys)**.** Olkoon E tapahtuma, jolle $\mathrm{P}(E) > 0$. Tällöin tapahtuman A *ehdollinen todennäköisyys ehdolla* E on

$$\mathrm{P}(A \mid E) = \frac{\mathrm{P}(A \cap E)}{\mathrm{P}(E)}.$$

Jos ehtotapahtuman E todennäköisyys on 0, eli $\mathrm{P}(E) = 0$, niin $P(A|E)$ pitää määritellä tapauskohtaisesti.



Määritelmä 1.2.10 on yhtäpitävä seuraavan tulokaavan kanssa:

1.2.11 Lause (Tulokaava)**.** *Kaikille tapahtumille A ja B pätee*

$$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A).$$

Todistus. Huomaamme aluksi, että jos $\mathrm{P}(A)=0$, niin lause pätee muodossa $0=0$. Voimme siis olettaa nyt, että $\mathrm{P}(A)>0$. Tällöin lause seuraa suoraan ehdollisen todennäköisyyden määritelmästä 1.2.10:

$$\mathrm{P}(A)\mathrm{P}(B|A) \quad = \quad \mathrm{P}(A)\frac{\mathrm{P}(B\cap A)}{\mathrm{P}(A)} \quad = \quad \mathrm{P}(A\cap B).$$

□

Edellä oli formaali todistus. Lauseen 1.2.11 havaitsee "todeksi" myös tulkitsemalla

$$\begin{aligned} & A\cap B \\ &\quad = \ A \text{ sattuu ja } B \text{ sattuu} \\ &\quad = \ \text{ensin sattuu } A \text{ ja sitten sattuu } B \\ &\quad = \ \text{ensin sattuu } A \text{ ja sitten sattuu } B, \text{ kun tiedetään } A\text{:n sattuneen} \\ &\quad = \ A\cap B|A. \end{aligned}$$

Ketjuttamalla tätä ajatusta näemme yleisen tulokaavan:

1.2.12 Lause (Ylenen tulokaava)**.** *Kaikille tapahtumille $A_1,\ldots,A_n$ pätee*

$$\begin{aligned} & \mathrm{P}(A_1\cap\cdots\cap A_n) \\ &\quad = \ \mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\cdots\mathrm{P}(A_n|A_1\cap\cdots\cap A_{n-1}). \end{aligned}$$

Todistus. Jätämme formaalin todistuksen harjoitustehtäväksi 1.5. Vihjeenä sanottakoon, että kannattaa käyttää tulokaavalauseen 1.2.11 todistusta ja induktiota. □

1.2.13 Esimerkki. Pekka ja Jukka päättävät pelata perjantai-illan ratoksi venäläistä rulettia. Yhteisen sopimuksen mukaan Pekka aloittaa, eli Pekka vetää liipaisimesta ensin. Rulettia pelataan kunnes peli päättyy luonnollisella traagisella tavallaan.

Onko peli reilu, kun

(a) rullaa pyöräytetään ennen jokaista liipaisimenvetoa,
(b) rullaa pyöräytetään ainoastaan ennen ensimmäistä liipaisimenvetoa?

Kohdassa (a) pitää ajatella loputonta laukausten sarjaa, sillä peli voi kestää periaatteessa loputtomasti — joskaan ei kannata laskea sen varaan, että kuolee vanhuuteen. Olkoot

$$\begin{aligned} A &= \text{“Pekka häviää”}, \\ A_n &= \text{“Pekka häviää kierroksella nro } n\text{”}. \end{aligned}$$

Tällöin

$$\mathrm{P}(A) = \sum_{n=1}^{\infty} \mathrm{P}(A_n).$$

Ratkaistaan sitten todennäköisyys $\mathrm{P}(A_n)$. Selvästi $\mathrm{P}(A_1) = 1/6$. Entä $\mathrm{P}(A_2)$? Olkoon

$$B_n = \text{“Jukka häviää kierroksella nro } n\text{”}.$$

Nyt A_2 sattuu, jos ensiksi sattuu A_1^c, sitten sattuu B_1^c, ja sitten sattuu A_2. Siten

$$\begin{aligned} \mathrm{P}(A_2) &= \mathrm{P}(A_1^c \cap B_1^c \cap A_2) \\ &= \mathrm{P}(A_1^c)\mathrm{P}(B_1^c|A_1^c)\mathrm{P}(A_2|A_1^c \cap B_1^c) \\ &= 5/6 \cdot 5/6 \cdot 1/6. \end{aligned}$$

Samalla tavalla näemme, että

$$\begin{aligned} \mathrm{P}(A_3) &= \mathrm{P}(A_1^c \cap B_1^c \cap A_2^c \cap B_2^c \cap A_3) \\ &= \mathrm{P}(A_1^c) \cdot \mathrm{P}(B_1^c|A_1^c)\mathrm{P}(A_2^c|A_1^c \cap B_1^c) \cdot \mathrm{P}(B_2^c|A_1^c \cap B_1^c \cap A_2^c) \cdot \\ &\quad \mathrm{P}(A_3|A_1^c \cap B_1^c \cap A_2^c \cap B_2^c) \\ &= 5/6 \cdot 5/6 \cdot 5/6 \cdot 5/6 \cdot 1/6. \end{aligned}$$

Jatkamalla samalla tavalla näemme yleisen kuvion:

$$\mathrm{P}(A_n) = (5/6)^{2(n-1)} \cdot 1/6.$$

Siten käyttämällä *geometrisen sarjan* laskusääntöä

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}, \qquad \text{kun } |q| < 1,$$

saamme

$$\begin{aligned} P(A) &= \sum_{n=1}^{\infty} P(A_n) \\ &= \sum_{n=1}^{\infty} (5/6)^{2(n-1)} \cdot 1/6 \\ &= 1/6 \sum_{n=0}^{\infty} (5/6)^{2n} \\ &= 1/6 \sum_{n=0}^{\infty} (25/36)^{n} \\ &= 1/6 \cdot 1/(1-25/36) \\ &= 6/11 \\ &> 1/2. \end{aligned}$$

Siten peli ei ole reilu.

Kohdassa (b) kannattaa ajatella kuutta laukausta (enempää ei tarvita pelin loppuunsaattamiseksi). Mikäli luoti on pesässä laukauksella 1, 3 tai 5, kuolee Pekka. Mikäli luoti on pesässä laukauksella 2, 4 tai 6 kuolee Jukka. Koska luoti on rullan pyöräyttämisen jälkeen pesässä yhtä hyvin millä tahansa laukauksista $1, 2, 3, 4, 5$ tai 6, peli on reilu.

Riippumattomuus

Intuitiivisesti tapahtumat ovat riippumattomia, jos toisten tapahtumien sattumiset tai sattumatta jäämiest eivät vaikuta toisten tapausten todennäköisyyksiin. Formaali määritelmä on:

1.2.14 Määritelmä (Riippumattomuus)**.** Tapahtumat $A_1, \ldots, A_n$ ovat *riippumattomia*, jos päätee *yksinkertainen tulokaava*

$$P(A_{i_1} \cap \cdots \cap A_{i_m}) = P(A_{i_1}) \cdots P(A_{i_m}),$$

missä $A_{i_1}, \ldots, A_{i_m}$ on mikä tahansa osakokoelma joukoista $A_1, \ldots, A_n$.

1.2.15 Huomautus. Yleisen tulokaavan 1.2.12 nojalla näemme, että tapahtumien $A_1, \ldots, A_n$ riippumattomuus tarkoittaa sitä, että

$$P(A_k | A_{i_1} \cap \cdots \cap A_{i_m}) = P(A_k)$$

mille tahansa kokoelmalle $A_{i_1}, \ldots A_{i_m}$, joka ei tietenkään sisällä itse tapahtumaa A_k.

1.2.16 Esimerkki. Kolikkoa heitetään kaksi kertaa. Olkoot

$$\begin{aligned} A &= \text{ensimäinen heitto antaa klaavan,} \\ B &= \text{toinen heitto antaa klaavan,} \\ C &= \text{heitot menevät eri päin.} \end{aligned}$$

Esimerkin 1.2.16 tapahtumat A, B, C ovat pareittain riippumattomat:

$$\begin{aligned} P(A \cap B) &= P(A)P(B) &= 1/4, \\ P(A \cap C) &= P(A)P(C) &= 1/4, \\ P(B \cap C) &= P(B)P(C) &= 1/4. \end{aligned}$$

Kolmikko A, B, C ei kuitenkaan ole riippumaton, sillä

$$1/8 \quad = \quad P(A)P(B)P(C) \quad \neq \quad P(A \cap B \cap C) \quad = \quad 0.$$

Bayesin ja kokonaistodennäköisyyden kaavat

1.2.17 Lause (Bayesin kaava I). *Olkoon* $P(E) > 0$. *Tällöin tapahtumaan* A_k *liittyvät ehdolliset todennäköisyydet voidaan kääntää* Bayesin kaavalla

$$P(A_k|E) = \frac{P(A_k)P(E|A_k)}{P(E)}.$$

Todistus. Väite seuraa ehdollisen todennäköisyyden määritelmästä 1.2.10 ja tulokaavasta 1.2.11. Nimittäin

$$P(A_k|E) = \frac{P(A_k \cap E)}{P(E)} = \frac{P(A_k)P(E|A_k)}{P(E)}.$$

Edellä käytettiin havaintoa $A_k \cap E = E \cap A_k$. □

1.2.18 Huomautus. Kaavan 1.2.17 tekijöille on seuraavat tulkinnat

- $P(A_k)$ on vaihtoehdon A_k *prioritodennäköisyys*,
- $P(E|A_k)$ on tapahtuman E *uskottavuus* priorin $P(A_k)$ vallitessa,
- $P(A_k|E)$ on vaihtoehdon A_k *posterioritodennäköisyys* evidenssin E valossa.

Joskus todennäköisyys $P(E)$ on vaikea laskea suoraan. Tällöin voidaan käyttää *kokonaistodennäköisyyden kaavaa*:

1.2.19 Lause (Kokonaistodennäköisyyden kaava). *Olkoot* A_j, $j \in \mathbb{N}$, *sellaisia tapahtumia, että täsmälleen yksi niistä sattuu. Tällöin*

$$P(E) = \sum_{j \in \mathbb{N}} P(A_j)P(E|A_j),$$

Todistus. Kaikki seuraa olennaisesti todennäköisyyden täysadditiivisuudesta 1.2.1(iii):

$$\begin{aligned} \mathrm{P}(E) &= \mathrm{P}\left(E \cap \Omega\right) \\ &= \mathrm{P}\left(E \cap (\cup_{j\in\mathbb{N}} A_j)\right) \\ &= \mathrm{P}\left(\cup_{j\in\mathbb{N}} (E \cap A_j)\right) \\ &= \sum_{j\in\mathbb{N}} \mathrm{P}\left(E \cap A_j\right) \\ &= \sum_{j\in\mathbb{N}} \mathrm{P}(A_j)\mathrm{P}(E|A_j). \end{aligned}$$

Yllä oli olennaista, että $\{A_j; j \in \mathbb{N}\}$ on ositus, eli täsmälleen yksi tapahtumista A_j sattuu. □

Yhdistämällä kokonaistodennäköisyyden kaavan 1.2.19 Bayesin kaavaan 1.2.17 saamme kaavan, jota myös kutsutaan Bayesin kaavaksi:

1.2.20 Lause (Baysin kaava II)**.** *Olkoot* A_j*,* $j \in \mathbb{N}$*, sellaisia tapahtumia, että täsmälleen yksi niistä tapahtuu. Tällöin jokaiselle* A_k *ja jokaiselle tapahtumalle* E*, jolle* $\mathrm{P}(E) > 0$*, pätee*

$$\mathrm{P}(A_k|E) = \frac{\mathrm{P}(A_k)\mathrm{P}(E|A_k)}{\sum_{j\in\mathbb{N}} \mathrm{P}(A_j)\mathrm{P}(E|A_j)}.$$

1.2.21 Esimerkki. Työpaikalla järjestetään huumetesti. Testi on 99% varma. Toisin sanoen vain 1% huumeenkäyttäjistä jää paljastumatta ja vain 1% niistä, jotka eivät käytä huumeita antavat väärän positiivisen. Oletamme, että noin yksi kymmenestä tuhannesta ihmisestä käyttää huumeita.

Työntekijä Pekka menee huumetestiin ja saa positiivisen tuloksen (ja mitä todennäköisimmin potkut). Mikä on todennäköisyys, että Pekka on huumeidenkäyttäjä?

Olkoot

$$\begin{aligned} A &= \text{“Pekka on huumeidenkäyttäjä”}, \\ E &= \text{“huumetestistä tuli positiivinen tulos Pekalle”}. \end{aligned}$$

Kysytty todennäköisyys on $\mathrm{P}(A|E)$. Tyypillinen ajatusvirhe tässä ongelmassa on vastata todennäköisyydellä $\mathrm{P}(E|A) = 99\%$. Tätä ei kuitenkaan kysytty! Todennäköisyys $\mathrm{P}(A|E)$ voidaan laskea käyttämällä Bayesin kaavaa

$$\begin{aligned} \mathrm{P}(A|E) &= \frac{\mathrm{P}(A)\mathrm{P}(E|A)}{\mathrm{P}(E)} \\ &= \frac{0{,}0001 \cdot 0{,}99}{\mathrm{P}(E)} \\ &= \frac{0{,}000099}{\mathrm{P}(E)}. \end{aligned} \tag{1.2.22}$$

Jotta kaavaa (1.2.22) voitaisiin käyttää tulee meidän määrätä todennäköisyys $\mathrm{P}(E)$. Kokonaistodennäköisyyden kaavan nojalla

$$\begin{aligned}\mathrm{P}(E) &= \mathrm{P}(E\cap A)+\mathrm{P}(E\cap A^c)\\ &= \mathrm{P}(A)\mathrm{P}(E|A)+\mathrm{P}(A^c)\mathrm{P}(E|A^c)\\ &= 0{,}0001\cdot 0{,}99+0{,}9999\cdot 0{,}01\\ &= 0{,}010098.\end{aligned}$$

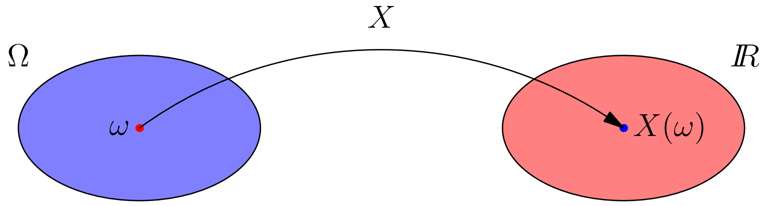
Sijoittamalla tämä kaavaan (1.2.22) saamme

$$\mathrm{P}(A|E) \quad = \quad 0{,}0098039 \quad = \quad 0{,}1\%.$$

Siis vaikka testi on 99% varma, niin Pekka mitä luultavimmin ei ole huumeidenkäyttäjä! Syy pieneen todennäköisyyteen olla huumeidenkäyttäjä positiivisella testituloksella on huumeidenkäyttäjien pieni osuus populaatiosta. Siten suurin osa positiivistista testituloksista (yli 99%) on vääriä positiivisia.

1.3 Satunnaismuuttujat

1.3.1 Määritelmä (Satunnaismuuttuja)**.** *Satunnaismuuttuja* X on satunnaiskokeen tulos $X:\Omega\to\mathbb{R}$. Satunnaismuuttuja on siis *funktio* otosavaruudelta Ω.



1.3.2 Huomautus. Tarkastelemme (lähes) ainoastaan *diskreettejä* satunnaismuuttujia. Satunnaismuuttuja on *diskreetti*, jos sen maalijoukko, eli mahdollisten arvojen joukko, on *numeroituva*[4] joukko:

$$\{X(\omega)\,;\,\omega\in\Omega\} \quad = \quad \{x_j\,;\,j\in\mathbb{N}\} \quad = \quad \{x_1,x_2,\ldots\}\,.$$

Toinen tapa määritellä diskreetti satunnaismuuttuja on vaatia, että se saa jokaisen mahdollisen arvonsa (aidosti) positiivisella todennäköisyydellä. Tätä kautta numeroituvuus liittyy läheisesti todennäköisyyden aksioomaan 1.2.1(iii).

[4]Joukko A on numeroituva, jos on olemassa sellainen kuvaus $f:\mathbb{N}\to A$, että jokaiselle $a\in A$ löytyy alkukuva $j=f^{-1}(a)\in\mathbb{N}$. Joukko on äärellinen, jos sen alkioiden lukumäärä on äärellinen. Erityisesti äärelliset joukot ovat numeroituvia. Myös esimerkiksi kokonaislukujen joukko $\mathbb{Z}$ on numeroituva, mutta reaalilukujen joukko $\mathbb{R}$ ei ole numeroituva.

Satunnaismuuttujan jakauma

1.3.3 Määritelmä (Jakauma ja fraktiilit)**.** Satunnaismuuttujan X *jakauma* P_X kertoo sen tulosmahdollisuuksien todennäköisyydet:

$$\mathrm{P}_X(x) = \mathrm{P}(X = x).$$

Fraktiilit puolestaan muodostavat satunnaismuuttujan todennäköisyyksien kertymäfunktion käänteisfunktion: satunnaismuuttujan X *q-fraktiili* x_q on

$$x_q = \max\{x\ ;\ \mathrm{P}(X < x) \leq q\}.$$

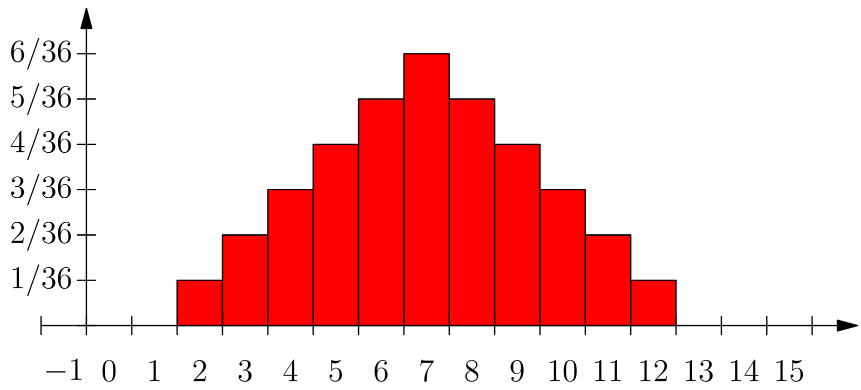
1.3.4 Esimerkki. Kahta (kuusisivuista ja symmetristä) noppaa heitetään. Olkoon X = "silmälukujen summa".

Nyt X on satunnaismuuttuja. Voimme esimerkiksi mallittaa $\Omega = \{(i,j); i,j = 1,2,3,4,5,6\}$ ja $X(i,j) = i + j$. Satunnaismuuttuja X jakauma on

$$\begin{aligned}
\mathrm{P}_X(2) &= 1/36, & \mathrm{P}_X(6) &= 5/36, & \mathrm{P}_X(10) &= 3/36,\\
\mathrm{P}_X(3) &= 2/36, & \mathrm{P}_X(7) &= 6/36, & \mathrm{P}_X(11) &= 2/36,\\
\mathrm{P}_X(4) &= 3/36, & \mathrm{P}_X(8) &= 5/36, & \mathrm{P}_X(12) &= 1/36,\\
\mathrm{P}_X(5) &= 4/36, & \mathrm{P}_X(9) &= 4/36, & &
\end{aligned}$$

ja $\mathrm{P}_X(x) = 0$ kaikilla muilla x.

Jakauma voidaan esittää myös kuvana



Satunnaismuuttujan X fraktiileja ovat esimerkiksi *kvartiilit*

$$x_{0,25} = 5, \quad x_{0,50} = 7 \quad \text{ja} \quad x_{0,75} = 9.$$

1.3.5 Huomautus. Usein satunnaismuuttuja X on, kuten esimerkin 1.3.4 tapauksessa, $\mathbb{N}$-arvoinen. Tällöin (ja usein muulloinkin) merkitsemme lyhyesti

$$p_n \quad = \quad \mathrm{P}_X(n) \quad = \quad \mathrm{P}(X = n).$$

Satunnaismuuttujien riippumattomuus

Intuitiivisesti satunnaismuuttujat $X_1, \ldots, X_n$ ovat riippumattomia, jos tieto joistain satunnaismuuttujien tuloksista ei voi koskaan muuttaa muiden satunnaismuuttujien tulosmahdollisuuksien todennäköisyyksiä. Formaali määritelmä on:

1.3.6 Määritelmä (Satunnaismuuttujien riippumattomuus)**.** Satunnaismuuttujat $X_1, \ldots, X_n$ ovat *riippumattomia*, jos

$$\mathrm{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) \quad = \quad \mathrm{P}(X_1 \leq x_1) \cdots \mathrm{P}(X_n \leq x_n)$$

kaikille $x_1, \ldots, x_n \in \mathbb{R}$.

1.3.7 Huomautus. Jos satunnaismuuttujat $X_1, \ldots, X_n$ ovat diskreettejä, niin määritelmä 1.3.6 voidaan antaa myös ehtona

$$\mathrm{P}(X_1 = x_1, \ldots, X_n = x_n) \quad = \quad \mathrm{P}(X_1 = x_1) \cdots \mathrm{P}(X_n = x_n)$$

kaikille $x_1, \ldots, x_n \in \mathbb{R}$.

Satunnaismuuttujan paikan mittarit

1.3.8 Määritelmä (Paikkalukuja)**.**

(i) Satunnaismuuttujan X *odotusarvo* $\mathrm{E}(X)$ on sen todennäköisyyksin painotettu keskiarvo
$$\mathrm{E}(X) = \sum_{x \in \mathbb{R}} x \, \mathrm{P}_X(x).$$

(ii) *Mediaani* jakaa jakauma kahteen yhtä suureen osaan siten, että puolet jakaumasta on mediaania pienempää ja puolet mediaania suurempaa. Satunnaismuuttujan X *mediaani*, eli keskiverto, m on
$$m = \max\{x \; ; \; \mathrm{P}(X < x) \leq 1/2\}.$$
Mediaani on siis $1/2$-fraktiili $x_{0{,}50}$.

(iii) Satunnaismuuttujan X *moodi*, eli tyyppiarvo, mikä tahansa luku M, jolle $\mathrm{P}_X(M)$ maksimoituu. Moodi on siis satunnaiskokeen X todennäköisin, eli tyypillisin, arvo.[5]

1.3.9 Huomautus. Esimerkissä 1.3.4 odotusarvo, mediaani ja moodi ovat samoja lukuja. Yleisesti ne voivat kaikki olla eri lukuja.

Odotusarvolle (toisin kuin esimerkiksi mediaanille tai moodille) pätee seuraavat näppärät laskusäännöt:

[5] Sivumennen sanoen, moodi lienee hyödyttömin kaikista jakaumien tunnusluvuista.

1.3.10 Lause (Odotusarvon laskusääntöjä)**.** *Olkoot $X_1, \ldots, X_n$ satunnaismuuttujia.*

(i) *Aina pätee*

$$\mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}\left(X_i\right).$$

(ii) *Jos satunnaismuuttujat ovat riippumattomia, niin pätee myös*

$$\mathrm{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathrm{E}\left(X_i\right).$$

(iii) *Jos $f : \mathbb{R}^n \to \mathbb{R}$, niin aina pätee*

$$\begin{aligned} &\mathrm{E}\left(f(X_1, \ldots, X_n)\right) \\ &= \sum_{x_1, \ldots, x_n \in \mathbb{R}} f(x_1, \ldots, x_n)\, \mathrm{P}\left(X_1 = x_1, \ldots, X_n = x_n\right) \end{aligned}$$

Todistus. Todistamme väitteet (i)–(iii) ainoastaan tilanteessa $n = 2$. Yleisen tapauksen voi todistaa samalla tavalla kohtalaisen helposti esimerkiksi käyttämällä induktiota.

(i) Tämä väite seuraa osan (iii) väitteestä. Nimittäin valitsemalla $f(x_1, x_2) = x_1 + x_2$ saamme

$$\begin{aligned} \mathrm{E}(X_1 + X_2) &= \sum_{x_1, x_2 \in \mathbb{R}} (x_1 + x_2)\mathrm{P}(X_1 = x_1, X_2 = x_2) \\ &= \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \mathbb{R}} x_1 \mathrm{P}(X_1 = x_1, X_2 = x_2) \\ &\quad + \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \mathbb{R}} x_2 \mathrm{P}(X_1 = x_1, X_2 = x_2) \\ &= \sum_{x_1 \in \mathbb{R}} x_1 \sum_{x_2 \in \mathbb{R}} \mathrm{P}(X_1 = x_1, X_2 = x_2) \\ &\quad + \sum_{x_2 \in \mathbb{R}} x_2 \sum_{x_1 \in \mathbb{R}} \mathrm{P}(X_1 = x_1, X_2 = x_2) \\ &= \sum_{x_1 \in \mathbb{R}} x_1 \mathrm{P}(X_1 = x_1) + \sum_{x_2 \in \mathbb{R}} x_2 \mathrm{P}(X_2 = x_2) \\ &= \mathrm{E}(X_1) + \mathrm{E}(X_2). \end{aligned}$$

Väitteen (i) voi todistaa myös suoraan käyttämällä kokonaistodennäköisyyden kaavaa 1.2.19 toistuvasti etu- ja takaperin:

$$
\begin{aligned}
&\mathrm{E}\,(X_1 + X_2) \\
&= \sum_{z\in\mathbb{R}} z\mathrm{P}(X_1 + X_2 = z) \\
&= \sum_{z\in\mathbb{R}} z \sum_{x_1\in\mathbb{R}} \mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = z - x_1|X_1 = x_1) \\
&= \sum_{z\in\mathbb{R}} \sum_{x_1\in\mathbb{R}} z\mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = z - x_1|X_1 = x_1) \\
&= \sum_{x_2\in\mathbb{R}} \sum_{x_1\in\mathbb{R}} (x_1 + x_2)\mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = x_2|X_1 = x_1) \\
&= \sum_{x_2\in\mathbb{R}} \sum_{x_1\in\mathbb{R}} x_1\mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = x_2|X_1 = x_1) \\
&\quad + \sum_{x_2\in\mathbb{R}} \sum_{x_1\in\mathbb{R}} x_2\mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = x_2|X_1 = x_1) \\
&= \sum_{x_1\in\mathbb{R}} x_1\mathrm{P}(X_1 = x_1) \sum_{x_2\in\mathbb{R}} \mathrm{P}(X_2 = x_2|X_1 = x_1) \\
&\quad + \sum_{x_2\in\mathbb{R}} x_2 \sum_{x_1\in\mathbb{R}} \mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = x_2|X_1 = x_1) \\
&= \sum_{x_1\in\mathbb{R}} x_1\mathrm{P}(X_1 = x_1) + \sum_{x_2\in\mathbb{R}} x_2\mathrm{P}(X_2 = x_2) \\
&= \mathrm{E}(X_1) + \mathrm{E}(X_2).
\end{aligned}
$$

(ii) Tämäkin väite seuraa kohdasta (iii). Nimittäin valitsemalla funktion $f(x_1, x_2) = x_1x_2$ saamme

$$
\begin{aligned}
\mathrm{E}(X_1X_2) &= \sum_{x_1,x_2\in\mathbb{R}} x_1x_2\mathrm{P}(X_1 = x_1, X_2 = x_2) \\
&= \sum_{x_1\in\mathbb{R}} \sum_{x_2\in\mathbb{R}} x_1x_2\mathrm{P}(X_1 = x_1)\mathrm{P}(X_2 = x_2) \\
&= \sum_{x_1\in\mathbb{R}} x_1\mathrm{P}(X_1 = x_1) \sum_{x_2\in\mathbb{R}} x_2\mathrm{P}(X_2 = x_2) \\
&= \mathrm{E}(X_1)\mathrm{E}(X_2).
\end{aligned}
$$

Tämänkin väitteen olisi voinut todistaa suoraan käyttämättä kohtaa (iii), mutta jätämme sen tekemättä.

(iii) Koska kohta (ii) todistettiin käyttämällä tätä kohtaa, pitää nyt olla tarkkana, ettemme käytä kohtaa (ii) perustelussa. Tällöin sortuisimme kehäpäättelyyn.

Väite seuraa onneksi suoraan "pakettisummaamalla":

$$
\begin{aligned}
&\mathrm{E}\left(f(x_1, x_2)\right) \\
&= \sum_{z\in\mathbb{R}} z\,\mathrm{P}\left(f(X_1, X_2) = z\right) \\
&= \sum_{z\in\mathbb{R}} \sum_{x_1, x_2\in\mathbb{R}\,;\,f(x_1,x_2)=z} f(x_1, x_2)\,\mathrm{P}(X_1 = x_1, X_2 = x_2) \\
&= \sum_{x_1,x_2\in\mathbb{R}} f(x_1, x_2)\mathrm{P}(X_1 = x_1, X_2 = x_2).
\end{aligned}
$$

□

Satunnaismuuttujan hajonnan mittarit

Varianssi kuvastaa jakauman hajontaa sen odotusarvo ympärillä: mitä enemmän varianssia sitä enemmän hajontaa, eli satunnaisuutta.

1.3.11 Määritelmä (Varianssi)**.** Satunnaismuuttujan X *varianssi* $\mathrm{Var}(X)$ on

$$
\mathrm{Var}(X) \quad = \quad \mathrm{E}\left(\left(X - \mathrm{E}(X)\right)^2\right)
$$

1.3.12 Huomautus (Steinerin siirtosääntö)**.** Varianssin voi laskea myös käyttämällä *Steinerin siirtosääntöä*:

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}\left(\left(X - \mathrm{E}(X)\right)^2\right) \\
&= \sum_{x\in\mathbb{R}} \left(x - \sum_{y\in\mathbb{R}} y\mathrm{P}_X(y)\right)^2 \mathrm{P}_X(x) \\
&= \sum_{x\in\mathbb{R}} x^2\,\mathrm{P}_X(x) - \left(\sum_{x\in\mathbb{R}} x - \mathrm{P}_X(x)\right)^2 \\
&= \mathrm{E}(X^2) - (\mathrm{E}(X))^2.
\end{aligned}
$$

1.3.13 Huomautus. Muitakin hajontalukuja kuin varianssi on olemassa. On esimerkiksi keskipoikkeama $\mathrm{E}(|X - \mathrm{E}(X)|)$, kvartiilivälin pituus $x_{0{,}75} - x_{0{,}25}$ ja kvartiilipoikkeama $(x_{0{,}75} - x_{0{,}25})/2$. Emme käsittele näitä hajontalukuja tällä kurssilla.

1.3.14 Esimerkki. Esimerkin 1.3.4 satunnaismuuttujan odotusarvo on

$$
\begin{aligned}
\mathrm{E}(X) &= 2\cdot\frac{1}{36} + 3\cdot\frac{2}{36} + 4\cdot\frac{3}{36} + 5\cdot\frac{4}{36} + 6\frac{5}{36} + 7\cdot\frac{6}{36} \\
&\quad + 8\frac{5}{36} + 9\cdot\frac{4}{36} + 10\cdot\frac{3}{36} + 11\cdot\frac{2}{36} + 12\cdot\frac{1}{36} \\
&= 7.
\end{aligned}
$$

Esimerkin 1.3.4 satunnaismuutujan X mediaani ja moodi ovat myös 7. Varianssiksi saamme

$$\begin{aligned}
\mathrm{Var}(X) &= (2-7)^2\cdot\frac{1}{36}+(3-7)^2\cdot\frac{2}{36}+(4-7)^2\cdot\frac{3}{36}\\
&\quad+(5-7)^2\cdot\frac{4}{36}+(6-7)^2\,\frac{5}{36}+(7-7)^2\cdot\frac{6}{36}\\
&\quad+(8-7)^8\,\frac{5}{36}+(9-7)^2\cdot\frac{4}{36}+(10-7)^2\cdot\frac{3}{36}\\
&\quad+(11-7)^2\cdot\frac{2}{36}+(12-7)^2\cdot\frac{1}{36}\\
&= 2{,}9167.
\end{aligned}$$

1.3.15 Esimerkki. 80% suomalaisista uskoo olevansa keskimääräistä suomalaista köyhempiä. Miten tämä on mahdollista? Suurin osa suomalaisista uskoo olevansa keskimääräistä parempia autoilijoita. Miten tämä on mahdollista?[6]

Esimerkin 1.3.15 keskeinen ajatus on *mediaanin* ja *keskiarvon* erot.

Tarkastelemme aluksi ensimmäistä uskomusta. Väite

> “80% suomalaisista uskoo olevansa keskimääräistä suomalaista köyhempiä.”

saattaa olla totta tai saattaa olla olematta: mikäänhän ei sinänsä rajoita ihmisten uskomuksia. Mielenkiintoisemmaksi tilanne muuttuu, kun tarkastelemme uskomusten sijasta väitettä

> “80% suomalaisista *on* keskimääräistä suomalaista köyhempiä.”

Aluksi tämä väite saattaa tuntua ristiriitaiselta, mutta se saattaa itse asiassa olla totta! Ainakaan siinä ei ole mitään loogista ristiriitaa. Todellakin, esimerkiksi Tilastokeskuksen mukaan vuonna 2006 kotitalouksien luokitellut keskimääräiset omaisuustulot[7] olivat

[6] Esimerkissä 1.3.15 ei esiinny satunnaismuuttujia. Esimerkki voidaan kuitenkin tulkita satunnaismuuttujan $X =$ “umpimähkään valittu suomalainen” avulla. Tällöin väitteet ovat: todennäköisyydellä 80% umpimähkään valittu suomalainen uskoo olevansa keskimääräistä umpimähkään valittua suomalaista köyhempi ja todennäköisyydellä, joka on suurempi kuin 50% umpimähkään valittu suomalainen uskoo olevansa keskimääräistä umpimähkään valittua suomalaista parempi autoilija.

[7] Omaisuustulojen samaistaminen rikkauteen tai henkilöiden samaistaminen kotitalouksiin ei ole täysin perusteltua, mutta tarkoitus on vain antaa esimerkki — ei tehdä luotettava katsaus Suomen varallisuuden jakautumasta.

Luokka	Omaisuustulot	Graafinen esitys
I	144	
II	214	
III	318	
IV	426	
V	508	
VI	611	
VII	779	
VIII	1.049	
IX	1.759	
X	17.283	

Eli arvio itse jakaumasta (luokiteltu jakauma) on

alle 1.000	
1.001........2.000	
2.001........3.000	
3.001........4.000	
4.001........5.000	
5.001........6.000	
6.001........7.000	
7.001........8.000	
8.001........9.000	
9.001......10.000	
10.001.....11.000	
11.001.....12.000	
12.001.....13.000	
13.001.....14.000	
14.001.....15.000	
15.001.....16.000	
16.001.....17.000	
yli 17.001	

Luokkataulukosta arvioimme, että suomalaisten kotitalouksien omaisuustulojen keskiarvo on

$$\frac{144 + 214 + 318 + 426 + 508 + 611 + 779 + 1.049 + 1.759 + 17.283}{10}$$
$$= \quad 2.309$$

Siten tämän aineiston valossa 80% (itse asiassa 90%) suomalaisista todellakin ovat keskimääräistä köyhempiä.

Mitä autoilijoihin tulee, niin ei ole mitään loogista ristiriitaa siinä, että suurin osa suomalaisista on keskimääräistä suomalaista parempia autoilijoita. Sen sijaan väite ei todennäköisesti pidä paikkaansa, sillä muutamat ammattiautoilijat ovat luultavasti niin paljon parempia kuin suurin osa autoilijoista, että tyypillinen suomalainen autoilija on luultavasti paljon keskimääräistä autoilijaa huonompi. Tässä siis tarkasteltiin väitettä

> "Suurin osa suomalaisista *autoilijoista* on keskimääräistä suomalaista autoilijaa parempi autoilija",

ja argumentoitiin, että ko. väite saattaa hyvinkin olla epätotta. Sen sijaan väite

> "Suurin osa suomalaisista on keskimääräistä suomalaista parempia autoilijoita"

saattaa hyvinkin olla totta. Nimittäin Suomessa on merkittävä — mutta silti riittävän pieni — joukko ihmisiä, jotka ovat todella älyttömän huonoja autoilijoita (lapset ja imeväiset).

Todennäköisyyttä älyköille*

Tämä osio on *nörteille*. Peruskylteri tai -teekkari, joka haluaa helppoja vastauksia pohtikoon esimerkkiä 1.3.16 omalla vastuullaan!

1.3.16 Esimerkki. Peluuttaja C panee pelaajien A ja B otsaan lapun seuraavalla tavalla: Peluuttaja C valitsee luvun n umpinäkään. Sitten peluuttaja C lätkäisee luvun n pelaajan A otsaan ja heittää kolikkoa. Jos tulee klaava, niin peluuttaja C lätkäisee pelaajan A otsaan luvun $n+1$; jos tulee kruuna, niin peluuttaja C lätkäisee pelaajan A otsaan luvun $n-1$.

Nyt pelaajat A ja B näkevät toistensa otsaluvut, ja on sovittu, että pienemmän otsaluvun saaja maksaa suuremman otsasumman saajalle suuremman otsasumman määrän euroja. Jos siis, esimerkin vuoksi, A:lla on otsassa 43 ja B:llä otsassa 42, niin B maksaa A:lle 1:n euron.

Tarinan pointti on: Kummallakin pelaajalla on veto-oikeus, jolla peli voidaan mitätöidä. Kumpikin pelaaja näkee toisen otsan. Nyt siis pelaaja (vaikkapa A) on nähnyt toisen pelaajan otsan (vaikkapa B:n otsa on n). Kannattaako pelaajan (tässä A) käyttää veto-oikeuttaan?

Vihje: σ-algebra ja σ-additiivisuus! Nörtit puhuvat Kolmogoroviaanisesta todennäköisyysavaruudesta $(\Omega, \mathcal{F}, \mathrm{P})$,

Suurten lukujen laki

Esitämme varsin erikoisen version heikosta suurten lukujen laista. Tämä versio on siten rajoitettu, että siinä esiintyvillä satunnaismuuttujilla pitää olla "ohuet hännät". Toisin sanoen suuret (ja pienet) arvot eivät saa tapahtua liian suurilla todennäköisyyksillä.

1.3.17 Määritelmä (Ohut häntä)**.** Satunnaismuuttujalla X on *ohuet hännät*, jos $\mathrm{E}\left(e^{\theta X}\right) < \infty$ kaikilla $\theta \in \mathbb{R}$.

1.3.18 Esimerkki (Ohuet ja paksut hännät)**.** Esitämme kaksi esimerkkiä ohuista hännistä ja yhden paksuista:

(a) Rajoitetulla satunnaismmuuttujalla X on ohuet hännät. Nimittäin, jos $|X| \leq M$ jollekin vakiolle M, niin

$$\mathrm{E}\left(e^{\theta X}\right) \quad \leq \quad \mathrm{E}\left(e^{|\theta M|}\right) \quad = \quad e^{|\theta M|} \quad < \quad \infty.$$

(b) Satunnaismuuttuja X on Poisson-jakautunut parametrilla $\lambda > 0$, jos

$$\mathrm{P}(X = x) \quad = \quad e^{\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Poisson(λ)-jakautuneella satunnaismuuttujalla X on ohuet hännät. Nimittäin *eksponenttifunktion sarjaesityksen*

$$e^y \quad = \quad \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

nojalla

$$\begin{aligned}
\mathrm{E}\left(e^{\theta X}\right) &= \sum_{x=0}^{\infty} e^{\theta x}\, e^{-\lambda}\frac{\lambda^x}{x!} \\
&= e^{-\lambda}\sum_{x=0}^{\infty} \frac{\left(\lambda e^{\theta}\right)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^{\theta}} \\
&= e^{\lambda(e^{\theta}-1)} \\
&< \infty.
\end{aligned}$$

(c) Sitten paksut hännät: Olkoon X satunnaismuuttuja, jolle

$$\mathrm{P}(X = x) \quad = \quad c\frac{1}{x^2}, \qquad \text{kun } x = 1, 2, \ldots .$$

(Jos kiinnostaa, niin c:n on oltava $6/\pi^2$.) Nyt,

$$\mathrm{E}\left(e^{\theta X}\right) \;=\; \sum_{x=1}^{\infty} e^{\theta x}\,\mathrm{P}(X=x) \;=\; c\sum_{x=1}^{\infty}\frac{e^{\theta x}}{x^2} \;=\; \infty,$$

koska $e^{\theta x}/x^2 \geq 1$, kunhan vain x on riittävän iso (eksponentti voittaa aina potenssin). Itse asiassa, tällä satunnaismuuttujalla ei edes ole odotusarvoa, jota kohti sen keskiarvo voisi supeta.

1.3.19 Lause (Heikko suurten lukujen laki ohuilla hännillä)**.** *Olkoot* X_i, $i \in \mathbb{N}$*,riippumattomia samoin jakautuneita ohuthäntäisiä satunnaismuuttujia. Merkitään*

$$\begin{aligned}
\mu &= \mathrm{E}(X_i),\\
\Lambda(\theta) &= \ln \mathrm{E}\left(e^{\theta X_i}\right),\\
\Lambda^*(x) &= \max_{\theta\in\mathbb{R}}\left\{\theta x - \Lambda(\theta)\right\},
\end{aligned}$$

ja olkoon

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

Tällöin $\bar{X}_n \to \mu$ *siinä mielessä, ja sitä vauhtia, että kaikille* $\delta > 0$ *pätee*

$$\mathrm{P}\left(\bar{X}_n \geq \mu+\delta\right) \;\leq\; e^{-n\Lambda^*(\mu+\delta)}, \tag{1.3.20}$$

$$\mathrm{P}\left(\bar{X}_n \leq \mu-\delta\right) \;\leq\; e^{-n\Lambda^*(\mu-\delta)}. \tag{1.3.21}$$

1.3.22 Huomautus. Funktiolle Λ^* pätee: $\Lambda^*(\mu)=0$ ja $\Lambda^*(x)>0$ kun $x\neq\mu$. Siten lause 1.3.19 sanoo, että suppeneminen $\bar{X}_n \to \mu$ tapahtuu eksponentiaalisen nopeasti. Funktiota Λ^* kutsutaan *suurten poikkeamien vauhtifunktioksi*. Funktio Λ^* on *konveksi*[8], itse asiassa se on funktion Λ, joka myös on konveksi, *konveksi duaali*. Puhumme konvekseista funktioista myöhemmin lisää.

1.3.23 Huomautus. Suurten lukujen laki $\bar{X}_n \to \mu$ pätee toki, vaikka satunnaismuuttujilla olisi paksutkin hännät. Itse asiassa se pätee jos ja vain jos $\mathrm{E}(|X_i|) < \infty$. Paksuhäntäisessä tilanteessa ei kuitenkaan päde arviot (1.3.20)–(1.3.21).

1.3.24 Huomautus (Frekvenssitulkinta ja Suurten lukujen laki)**.** Frekvenssitulkinta 1.1.7 seuraa suurten lukujen laista 1.3.19 tarkastelemalla binäärisiä satunnaismuuttujia

$$X_n \;=\; \begin{cases} 1, & \text{jos tapahtuma } A \text{ sattuu toistossa } n,\\ 0, & \text{muulloin.}\end{cases}$$

[8] Karkeasti ottaen funktio on konveksi, jos sen toinen derivaatta on positiivinen. Tyypillinen konveksi funktio on x^2.

Nimittäin tällöin *otoskeskiarvo*

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

on tapahtuman A suhteellinen frekvenssi $f_n(A)$ ja odotusarvo μ on tapahtuman A todennäköisyys.

Näin siis frekvenssitulkinta 1.1.7 seuraa loogisesti Kolmogorovin aksioomista 1.2.1! Lisäksi epäyhtälöt (1.3.20)–(1.3.21) antavat mahdollisuuden arvioida, kuinka lähellä olemme oikeaa todennäköisyyttä μ. Nimittäin binäärimuuttujille X_i

$$\Lambda(\theta) = \ln\left(\mu e^{\theta} + 1 - \mu\right).$$

Sitten pitää maksimoida θ:n suhteen lauseke

$$\theta x - \Lambda(\theta) = \theta x - \ln\left(\mu e^{\theta} + 1 - \mu\right).$$

Tämä tehdään perinteisellä koulupojan metodilla derivoimalla lauseke θ:n suhteen ja asettamalla derivoitu lauseke nollaksi:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\theta x - \ln\left(\mu e^{\theta} + 1 - \mu\right)\right) = x - \frac{\mu e^{\theta}}{\mu e^{\theta} - \mu + 1}.$$

Lauseke on nolla, kun $\theta = \ln x/\mu - \ln(1-x)/(1-\mu)$. Sijoittamalla tämän paikalleen saamme maksimipisteen

$$\Lambda^*(x) = x \ln\frac{x}{\mu} + (1-x)\ln\frac{1-x}{1-\mu}.$$

Siten, esimerkiksi todennäköisyys todellisuutta δ:n verran suuremmalle arviolle on korkeintaan

$$\begin{aligned}(1.3.25)\qquad & \mathrm{P}\left(\bar{X}_n \geq \mu + \delta\right) \\ & \leq \exp\left\{-n\left((\mu+\delta)\ln\frac{\mu+\delta}{\mu} + (1-\mu-\delta)\ln\frac{1-\mu-\delta}{1-\mu}\right)\right\}\end{aligned}$$

Ongelma arviossa (1.3.25) on tietysti se, että tietenkään μ:n oikeaa arvoa ei tiedetä. Karkea ratkaisu ongelmaan on sijoittaa μ:n paikalle sen saatu arvio $\bar{x}_n = \bar{X}_n$. Tällöin emme vain voi olla varmoja arviomme oikeudesta. Varma ratkaisu on korvata μ pahimmalla mahdollisella arvolla $1/2$. Nimittäin tällöin arvion (1.3.25) oikea puoli maksimoituu. Tämä havainto on hyödyllinen myöhemmin, kun estimoimme todennäköisyyksiä suhteellisten frekvenssien menetelmällä luvussa 4.

Ennen lauseen 1.3.19 todistamista esitämme vielä varsin käyttökelpoisen estimaatin:

1.3.26 Apulause (Tšebyševin epäyhtälö)**.** *Olkoon X positiivinen satunnaismuuttuja. Tällöin kaikilla $x > 0$ pätee*

$$\mathrm{P}\Big(X \geq x\Big) \quad \leq \quad \frac{\mathrm{E}(X)}{x}.$$

Todistus. Väite voidaan kirjoittaa myös muodossa

$$x\mathrm{P}(X \geq x) \quad \leq \quad \mathrm{E}(X).$$

Tämä epäyhtälö puolestaan seuraa seuraavasta pikku pyörittelystä:

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{\xi \geq 0} \xi \mathrm{P}(X = \xi) \\
&\geq \sum_{\xi \geq x} \xi \mathrm{P}(X = \xi) \\
&\geq x \sum_{\xi \geq x} \mathrm{P}(X = \xi) \\
&= x\mathrm{P}(X \geq x).
\end{aligned}$$

□

Suurten lukujen lain 1.3.19 todistus. Olkoon

$$\begin{aligned}
S_n &= n\bar{X}_n, \\
\Lambda_n(\theta) &= \ln \mathrm{E}\left(e^{\theta S_n}\right).
\end{aligned}$$

Olkoon $x > \mu$. Tällöin kaikilla $\theta > 0$ pätee

$$\left\{\bar{X}_n \geq x\right\} \quad = \quad \{S_n \geq nx\} \quad = \quad \left\{e^{\theta S_n} \geq e^{\theta nx}\right\}$$

Siten, Tšebyševin epäyhtälön 1.3.26 nojalla

$$\mathrm{P}\left(\bar{X}_n \geq x\right) \quad \leq \quad e^{-n\theta x + \Lambda_n(\theta)}.$$

Käyttämällä nyt tietoa siitä, että riippumattomien satunnaismuuttujien tulon odotusarvo on odotusarvojen tulo, eli lausetta 1.2.3(ii), huomaamme, että

$$\Lambda_n(\theta) \quad = \quad n\Lambda(\theta).$$

Siten,

$$\mathrm{P}\left(\bar{X}_n \geq x\right) \quad \leq \quad e^{-n\{\theta x - \Lambda(\theta)\}}$$

Optimoimalla yli kaikkien $\theta > 0$ tästä seuraa, että

$$\mathrm{P}\left(\bar{X}_n \geq x\right) \quad \leq \quad e^{-n\Lambda^*(x)},$$

sillä voidaan osoittaa, että jos $x > \mu$, niin

$$\Lambda^*(\theta) \quad = \quad \max_{\theta\in\mathbb{R}} \{\theta x - \Lambda(\theta)\} \quad = \quad \max_{\theta>0} \{\theta x - \Lambda(\theta)\}$$

Jos $x < \mu$, niin saamme samalla tavalla, mutta valitsemalla $\theta < 0$

$$\mathrm{P}\left(\bar{X}_n \geq x\right) \quad \leq \quad e^{-n\Lambda^*(x)}.$$

Lause on todistettu. □

1.4 Harjoitustehtäviä lukuun 1

1.1 Harjoitustehtävä. Mitkä todennäköisyystulkinnat sopivat seuraaviin väitteisiin?

(a) Todennäköisyys että kolikonheitossa saadaan lopulta klaava on 1.
(b) Pokerikäsi “3 ässää” on todennäköisempi kuin pokerikäsi “4 ässää”.
(c) Maija on todennäköisesti Mattia parempi tekemään päätöksiä epävarmuuden vallitessa.
(d) Todennäköisyys että isä on poikaa pidempi on $0{,}39$.
(e) Ruotsi on todennäköisesti Suomea parempi jääkiekossa.
(f) Todennäköisyys että Vaasan Sport voittaa SM-liigan vuonna 2025 on $0{,}08$.

1.2 Harjoitustehtävä. Bayesläinen todennäköisyystulkinta perustuu informaation käsitteelle. Jos informaatio on subjektiivista, niin silloin todennäköisyydetkin ovat subjektiivisia. Toisaalta voidaan toivoa, että mikäli informaatio on objektiivista, niin silloin todennäköisyyskin on objektiivista. Seuraava tarina kuitenkin asettaa objektiivisen informaation käsitteen ylle ikävän varjon:
Kylässä elää 2.000 avioparia, jotka ovat kaikki lainkuuliaisia. Kylässä on laki, jonka mukaan jokaisen vaimon on tapettava aviorikoksen tehnyt aviomiehensä seuraavana päivänä siitä, kun vaimo sai tiedon siitä että hänen aviomiehensä on avionrikkoja. Lisäksi kylän vaimot ovat sillä tavalla noitia, että jokainen vaimo tietää automaattisesti jos joku muu mies, kuin oma aviomies tekee aviorikoksen. Vaimot ovat myös sillä tavalla jaloja, etteivät juorua aviorikoksista kenellekään. Oletamme, että kylässä on tehty 100 aviorikosta. Tällöin kaikki vaimot tietävät, että kylässä on avionrikkoja. Tämä on siis kaikkien tiedossa. *Eräänä päivänä kylään tulee matriarkka, joka kuuluttaa kylän torilla kaikille, että kylässä on avionrikkoja. Nyt siis* kaikkien tiedossa *olevasta asiasta tulee* yleisesti tiedetty.

Mitä tapahtuu 101 päivän kuluttua matriarkan ilmoituksesta?

1.3 Harjoitustehtävä. Olkoot A ja B tapahtumia. Osoita, että

$$\begin{aligned}
\mathrm{P}(A \setminus B) &= \mathrm{P}(A) - \mathrm{P}(A \cap B), \\
\mathrm{P}(B \setminus A) &= \mathrm{P}(B) - \mathrm{P}(A \cap B).
\end{aligned}$$

1.4 Harjoitustehtävä. Todista lauseen 1.2.3 kohta (iv).

1.5 Harjoitustehtävä. Todista yleinen tulokaava 1.2.12.

1.6 Harjoitustehtävä. Sadistinen miljonääri tarjoaa arpalippuja (ilmaiseksi). Arpajaisista on neljä versiota:

(i) Arpalippuja on 1.000.000 kappaletta. Arpalipuista 999.999 on sellaisia, että sadistinen miljonääri antaa lipun haltijalle 1.000 euroa, mutta yksi lipuista on sellainen, että lipun haltija joutuu kokemaan tuskallisen kuoleman sadistisen miljonäärin käsissä (kidutus kestää kaksi tuntia).

(ii) Sadistinen miljonääri muuttaa panoksia: arpalippuja on 100.000 kappaletta, joista 1 johtaa tuskalliseen kidutuskuolemaan kuten kohdassa (i), mutta 99.999 arpalippua antaa haltijalleen 10.000 euroa.

(iii) Sadistisen miljonäärin panokset vaan kovenee: nyt on jaossa 10.000 arpalippua, joista 1 johtaa tuskalliseen kidutuskuolemaan kuten kohdassa (i), mutta 9.999 arpalippua antaa haltijalleen 100.000 euroa.

(iv) Nyt sadistisella miljonäärillä on tosi kovat panokset: jaossa on vain 1.000 arpalippua, joista 1 johtaa tuskalliseen kidutuskuolemaan kuten kohdassa (i), mutta loput 999 arpalippua antaa haltijalleen 1.000.000 euroa.

Kysymme:

(a) Antti Ahnas haluaa voittaa 1.000.000 euroa. Mihinkä sadistisen miljonäärin arpajaisista (i), (ii), (iii) tai (iv) kannattaa Antti Ahnaan osallistua?

(b) Mihin arpajaisiin itse osallistuisit ja kuinka monella lipulla? Jos et mihinkään, niin miksi et? Kuinka suurena pidät sadistisen miljonäärin tarjoamaa kidutuskuoleman riskiä verrattuna riskiin kuolla tuskallisesti tänä vuonna?

1.7 Harjoitustehtävä. [9] n avioparia osallistuu parinvaihtopippaloihin, jossa uudet parit arvotaan umpimähkään. Mikä on todennäköisyys, että vähintään yksi arvottu pari on aviopari, kun

(a) $n = 2$,

(b) $n = 3$,

(c) $n = 4$,

(d) $n = 40$,

(e) $n = 400$,

(f) $n = 400.000$?

1.8 Harjoitustehtävä. Anna esimerkki satunnaismuuttujasta, jolla

[9] Tämä on kuuluisa *Recontre-ongelma*.

(a) mediaani on pienempi kuin odotusarvo,
(b) mediaani on suurempi kuin odotusarvo,
(c) on useita moodeja,
(d) sekä moodi, mediaani että odotusarvo ovat eri kohdissa,
(e) varianssi on nolla.

Voit antaa esimerkit puhtaasti matemaattisina, mutta mieti myös luonnollisia esimerkkejä.

1.9 Harjoitustehtävä. Todista lause 1.3.10 yleisessä n:n satunnaismuuttujan tilanteessa.

1.10 Harjoitustehtävä. Todista, että $\mathrm{P}(X = \mathrm{E}(X)) = 1$ jos ja vain jos $\mathrm{Var}(X) = 0$.

1.11 Harjoitustehtävä. Pörssissä on käsitteet *karhu* (bear) ja *sonni* (bull). Sanotaan, että on karhumarkkinat, jos uskotaan osakkeiden hintojen laskevan. Jos taas uskotaan, että osakkeiden hinnat nousevat, sanotaan että on sonnimarkkinat.

Pörssisijoittaja N. N. Taleb on juuri ostanut suuren määrän osakkeita. Hänen kollegansa uskovat, että N. N. Taleb uskoo markkinoiden olevan sonnivaiheessa. Kysyttäessä N. N. Taleb kuitenkin vastaa, että hän uskoo että on 80% todennäköisyys, että osakkeiden hinnat laskevat tulevaisuudessa. "Olet siis muuttanut mielesi!", kommentoivat kollegat. "Uskot siis karhumarkkinoihin. Nyt kai myyt osakkeesi nopeasti". "En suinkaan", sanoo N. N. Taleb, "Uskon karhumarkkinoihin, mutta silti katson että osakkeita kannattaa ostaa." N. N. Taleb katsoo olevansa rationaalinen. Kuinka tämä on mahdollista?

1.12 Harjoitustehtävä. Kurssin *ORMS2020: Päätöksenteko epävarmuuden vallitessa* loppukoe koostuu monivalintatehtävistä. Kokeen suunnitelija, professori S., ei halua päästää kurssista läpi hannuhanhia, jotka eivät oikeasti osaa päätöksentekoa epävarmuuden vallitessa. Kuinka professori S.:n tulee laatia monivalintakoe?

1.13 Harjoitustehtävä. Dosentti K.:lle on tarjottu Nikolaigradin ylpistöstä kahta professuuria: talousmatematiikan proessuuria sekä estetiikan ja vertailevan erotiikan professuuria. Dosentti K. haluaa jossain vaiheessa uraansa päätyä Nikolaigradin ylpistön rehtoriksi. Nikolaigradin ylpistössä rehtori valitaan professorien keskuudesta. Ylpistön historian aikana on ollut yhdeksän rehtoria, joista kaksi on ollut talousmatematiikan professoreja. Estetiikan ja vertailevan erotiikan professoria ole koskaan valittu rehtoriksi. Professoreja Nikolaigradin ylpistössä on viisikymmentä. Jos urakehitystä ei oteta huomioon, niin Dosentti K. olisi mieluummin estetiikan ja vertailevan erotiikan professori kuin talousmatematiikan professori. Kumman tarjotuista professuureista dosentti K.:n kannattaa ottaa vastaan?

1.14 Harjoitustehtävä. Laske satunnaismuuttujalle X funktiot Λ ja Λ^*, kun

(a) satunnaismuuttuja X on binomijakautunut parametrein n ja p,
(b) kun satunnaismuuttuja X on Poisson-jakautunut parametrilla λ.

1.15 Harjoitustehtävä. Osoita, että arvion (1.3.25) yläraja maksimoituu, kun $\mu = 1/2$.

Luku 2

Päätösmatriisit

> If you cannot calculate you cannot speculate on future pleasure and your life will not be that of a human, but that of an oyster or a jellyfish. – Plato

> A little simplification would be the first step toward rational living, I think.
> – Eleanor Roosevelt

> But it is strange how many rational beings believe the ultimate truths of the universe to be reducible to patterns on a blackboard.
> – Frederick Pollock

2.1 Staattinen asetelma

Tarkastelemme päätösongelmia, jotka ovat *staattisia*: päätös tehdään vain kerran ja sen jälkeen katsotaan seuraukset. Aikaulottuvuutta eli *dynamiikkaa* ei ole.

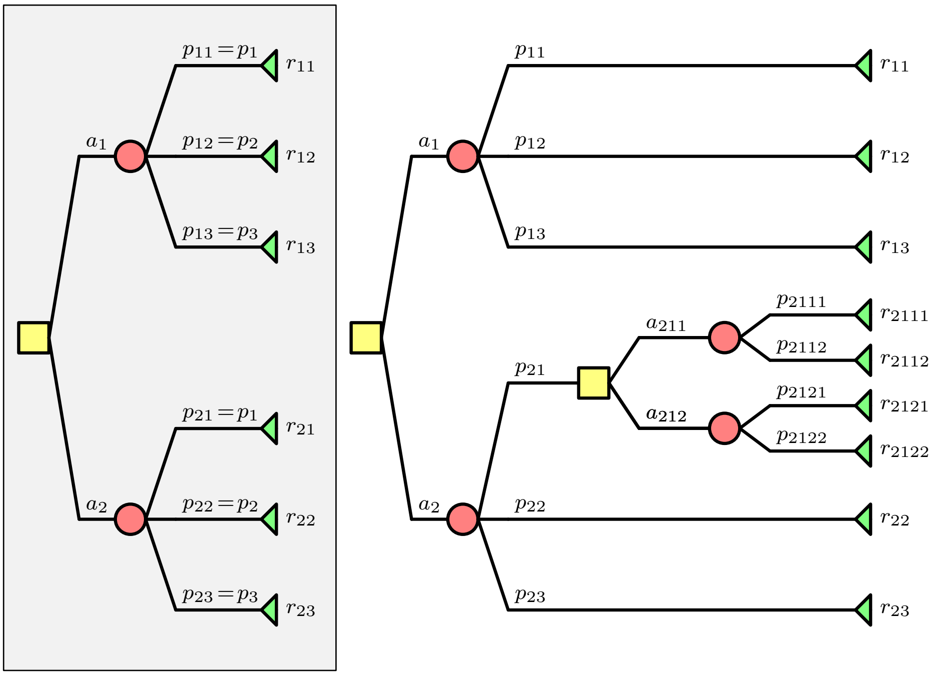
Päätöksentekijä valitsee päätöksen joukosta $A = \{a_i\,;\, i \in I\}$. Sitten kohtalon jumalatar Lady Fortuna arpoo jonkin maailmantilan joukosta $\Omega = \{\omega_j\,;\, j \in J\}$. Päätöksen a_i seuraus, jota kutsumme tässä myös a_i:ksi, on siis satunnaismuuttuja. Jos päätettiin a_i ja Lady Fortunan valinta oli ω_j, niin saadaan palkkio r_{ij}. Toisin sanoen $r_{ij} = a_i(\omega_j)$. Palkkiot r_{ij} voivat olla periaatteessa melkein mitä tahansa reaaliarvoisia vähintään välimatka-asteikkoisia suureita, mutta usein käytännössä palkkiot kannattaa tulkita rahaksi.

2.1.1 Määritelmä (Palkkiomatriisi, päätösmatriisi)**.** Olkoon I päätösindeksit ja J sattumaindeksit. Matriisi $R = [r_{ij}]_{i\in I, j\in J}$ on *palkkio-* eli *päätösmatriisi.* Maailmantila eli alkeistapaus ω_j tapahtuu todennäköisyydellä p_j, eli $p_j = \mathrm{P}(\omega_j) = \mathrm{P}(a_i = r_{ij})$.

2.1.2 Huomautus. Indeksijoukot I ja J voivat olla periaatteessa mitä tahansa, mutta käytännön mallinnuksessa ne ovat tyypillisesti äärellisiä: päätöksentekijällä on siis tällöin vain äärellinen määrä valintoja, samoin kuin kohtalon jumalattarellakin.

Jos I ja J eivät ole äärellisiä, niin määritelmän 2.1.1 palkkiomariisi R ei ole matriisi sanan perinteisessä mielessä, mikä aiheuttaa ongelmia käytännön laskuissa.

2.1.3 Huomautus (Staattinen vs. dynaaminen päätöstilanne)**.** Kuvallisesti määritelmän 2.1.1 mukainen staattinen päätösasetelma tarkoittaa seuraavan kuvan vasemman puolen puun kaltaista tilannetta (neliö tarkoittaa päätöstä ja ympyrä tarkoittaa sattumaa; kylkikolmio tarkoittaa tarkastelun loppua).



Vasemmanpuoleinen puu kuvastaa staattista tilannetta, jossa tehdään yksi päätös, a_1 tai a_2, ja sitten katsotaan, mikä palkkio saadaan. Palkkiovaihtoehdot valinnan a_1 jälkeen ovat r_{11}, r_{12} ja r_{13}, ja niiden todennäköisyydet ovat p_1, p_2 ja p_3. Valinnan a_2 jälkeen palkkiovaihtoehdot ovat r_{21}, r_{22} ja r_{23}, ja niiden todennäköisyydet ovat p_1, p_2 ja p_3. Huomattavaa siis on, että tapahtumien eli sattumien todennäköisyydet p_j ovat riippumattomia valinnoista a_i.

Oikeanpuoleinen puu taas kuvastaa dynaamista tilannetta, jossa esimerkiksi päätös a_2 johtaa uuteen päätöstilanteeseen, jossa on uudet aikaisemmista päätöksistä riippuvat todennäköisyydet. Tällaisia dynaamisia päätöstilanteita käsittelemme myöhemmin luvussa 3.

2.1.4 Määritelmä (Päätössääntö ja arvofunktio)**.** Olkoon $\mathcal{A}$ kaikkien staattisten päätöstilanteiden joukko. Toisin sanoen $\mathcal{A}$ koostuu kolmikoista (R, p, a), missä R on palkkiomatriisi, p on palkkiomatriisia R vastaava todennäköisyysvektori ja a on mahdolliset päätökset (eli olennaisesti R:n rivit). *Arvofunktio* $V : \mathcal{A} \to \bar{\mathbb{R}}$

antaa jokaiselle päätöstilanteen (R, p, a) päätökselle a_i numeerisen arvon $V(a_i)$.[1] Arvofunktiota V vastaava *päätössääntö* on valita sellainen päätös a_{i^*}, jolla arvofunktio V maksimoituu:

$$V(a_{i^*}) = \max_{i \in I} V(a_i).$$

Toisin sanoen

$$i^* = \operatorname*{argmax}_{i \in I} V(a_i).$$

2.1.5 Huomautus. Määritelmässä 2.1.4 $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. Hyväksymme siis myös arvot $-\infty$ ja ∞, ja sovellamme järjestystä $-\infty < a < \infty$ kaikille $a \in \mathbb{R}$ ja laskusääntöjä

$$\begin{aligned} +\infty + \infty &= +\infty, \\ -\infty - \infty &= -\infty, \\ a \cdot \infty &= +\infty, \text{ jos } a > 0, \\ a \cdot \infty &= -\infty, \text{ jos } a < 0, \\ a/\infty &= 0. \end{aligned}$$

Sen sijaan $\infty - \infty$, ∞/∞, $0 \cdot \infty$ eivät ole määriteltyjä.

2.1.6 Huomautus. Määritelmässä 2.1.4 mikään ei takaa, että optimaalinen päätös a_{i^*} eli

$$i^* = \operatorname*{argmax}_{i \in I} V(a_i)$$

on yksikäsitteinen. Jos optimaalisia päätöksiä on useita, niin silloin joko mikä tahansa niistä valitaan tai keksitään jokin lisäsääntö.

Eri tavat muodostaa arvofunktio V vastaavat erilaisia päätössääntöjä. Tätä ei pidä kuitenkaan ymmärtää niin, että jos päätössääntö V_1 (funktiona) on eri kuin päätössääntö V_2, niin V_1 ja V_2 johtavat välttämättä eri päätöksiin. Esimerkiksi seuraava tulos valaissee tätä ilmiötä:

2.1.7 Apulause (Päätösfunktioiden invarianssi)**.** *Olkoon $V : \mathcal{A} \to \bar{\mathbb{R}}$ päätösfunktio ja $f : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ aidosti kasvava funktio. Tällöin päätösfunktio*[2] *$f \circ V$ johtaa täsmälleen samoihin päätöksiin kuin päätösfunktio V.*

Todistus. Meidän tulee osoittaa, että jos a_{i^*} on optimaalinen päätössäännön V vallitessa, niin se on optimaalinen myös päätössäännön $f \circ V$ vallitessa, ja päin vastoin.

[1] Tässä raiskaamme notaatiota aika paljon. Pienempi raiskaus oli kirjoittaa $V(a)$:n sijaan $V(R, p, a)$.

[2] Merkintä $f \circ V$ tarkoittaa *yhdistettyä funktiota*: $(f \circ V)(x) = f(V(x))$. Jos esimerkiksi $f(x) = e^{2x}$ ja $V(x) = 3x$, niin $(f \circ V)(x) = e^{2(3x)} = e^{6x}$.

(i) Olkoon a_{i^*} jokin päätösfunktion V antama valinta. Tällöin siis $V(a_{i^*}) \geq V(a_i)$ kaikilla $i \in I$. Koska f on kasvava, niin myöskin $f(V(a_{i^*})) \geq f(V(a_i))$ kaikilla $i \in I$. Siten a_{i^*} on myös päätösfunktion $f \circ V$ antama optimaalinen päätös.
(ii) Kääntäen, olkoon a_{i^*} päätösfunktion $f \circ V$ antama optimaalinen valinta. Tällöin siis $f(V(a_{i^*})) \geq f(V(a_i))$ kaikilla $i \in I$. Koska f on aidosti kasvava, on se kääntyvä ja sen käänteisfunktio f^{-1} on myöskin aidosti kasvava. Siten kaikille $i \in I$ pätee

$$\begin{aligned} V(a_{i^*}) &= f^{-1}\left(f\left(V(a_{i^*})\right)\right) \\ &\geq f^{-1}\left(f\left(V(a_i)\right)\right) \\ &= V(a_i), \end{aligned}$$

eli a_{i^*} on optimaalinen valinta myös päätössäännön V vallitessa. □

2.2 Dominanssi ja pöhköt säännöt

Periaatteessa mikä tahansa funktio $V : \mathcal{A} \to \bar{\mathbb{R}}$ voi määrittää päätössäännön, mutta kaikki päätössäännöt eivät välttämättä ole järkeviä. Tarkastelemme järkevää eli rationaalista päätöksentekoa myöhemmin luvussa 5. Tässä luvussa tyydymme tarkistamaan, että päätössäännöt eivät ole aivan pöhkojä.

2.2.1 Määritelmä (Pöhkö päätössääntö)**.** Arvofunktiota $V : \mathcal{A} \to \bar{\mathbb{R}}$ vastaava päätössääntö on *pöhkö*, jos se (oi johtaa päätökseen a_{i^*}, vaikka olisi olemassa päätös $a_{i^\dagger}$, jolle

$$\forall j \in J : r_{i^\dagger,j} \geq r_{i^*,j} \quad \text{ja} \quad \exists j \in J : r_{i^\dagger,j} > r_{i^*,j}. \tag{2.2.2}$$

2.2.3 Huomautus (Dominanssi ja pöhköys)**.** (i) Kaava (2.2.2) sanoo, että päätös $a_{i^\dagger}$ *dominoi* päätöstä a_{i^*}. Siispä päätössääntö on pöhkö jos ja vain jos se johtaa päätökseen, joka on dominoitu.

(ii) Pöhköyden välttämisestä seuraa *dominanssiperiaate*: "jos valinta a_{i_1} dominoi valintaa a_{i_2}, älä tee valintaa a_{i_2}".

(iii) Dominanssiperiaate ei sano "jos valinta a_{i_1} dominoi valintaa a_{i_2}, tee valinta a_{i_1}". Voi nimittäin olla esimerkiksi jokin kolmas valinta a_{i_0}, joka dominoi valintaa a_{i_1} (ja siten automaattisesti myös valintaa a_{i_1}). Tai sitten valinta a_{i_0} on sellainen, että a_{i_1} ei dominoi sitä, eikä päinvastoin, mutta valinta a_{i_0} voi jostain muusta syystä olla valintaa a_{i_1} parempi.

2.2.4 Esimerkki (Dominanssin riittävyys ja riittämättömyys)**.** Joissain, muttei kaikissa, päätöstilanteissa pelkästään pöhköyden välttäminen riittää päätöksen tekemiseen.

(i) Olkoon

$$R = \begin{bmatrix} 1 & 9 & 0 & 0 & 8 \\ 1 & 8 & 0 & 0 & 0 \end{bmatrix}.$$

Tällöin mikä tahansa ei-pöhkö päätössääntö antaa valinnan a_1, sillä valinta a_1 antaa aina vähintään yhtä hyvän palkkion, kuin valinta a_2, ja joissain skenaarioissa a_1 antaa aidosti paremman palkkion, kuin valinta a_2.

(ii) Olkoon

$$R = \begin{bmatrix} 1 & 9 & 0 & 0 & 8 \\ 1 & 8 & 0 & 0 & 0 \\ 1 & 8 & 0 & 0 & 9 \end{bmatrix}.$$

Edelleenkin päätös a_1 dominoi päätöstä a_2, mutta a_1 ei dominoi päätöstä a_3. Siten kumpikaan valinnoista a_1 tai a_3 ei ole välttämättä pöhkö, mutta valinta a_2 on edelleen pöhkö, koska se on (edelleen) dominoitu päätösellä a_1 (ja päätöksellä a_3).

2.2.5 Huomautus (Sääntö vs. tilanne)**.** Pöhkökin päätössääntö voi tilanteesta riippuen johtaa järkeviin päätöksiin. Olkoon esimerkiksi päätössääntö

$$V(a_i) = \prod_{j \in J} r_{ij}. \tag{2.2.6}$$

Tämä on pöhkö sääntö, kuten tulemme myöhemmin huomaamaan.

Tarkastelemme päätöstilannetta

$$R = \begin{bmatrix} 1 & 9 & 1 & 1 & 8 \\ 1 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.2.7}$$

Tällöin a_1 on selvästi ainoa oikea päätös, koska se dominoi kaikkia muita päätöksiä. Toisaalta

$$\begin{aligned} V(a_1) &= 1 \cdot 9 \cdot 1 \cdot 1 \cdot 8 &= 72, \\ V(a_2) &= 1 \cdot 8 \cdot 0 \cdot 0 \cdot 0 &= 0, \\ V(a_3) &= 0 \cdot 0 \cdot 0 \cdot 0 \cdot 0 &= 0. \end{aligned}$$

Siten päätössääntö (2.2.6) antaa järkevän tuloksen tilanteessa (2.2.7).

Tarkastelemme sitten päätöstilannetta

$$R = \begin{bmatrix} 1 & 9 & 1 & -1 & 8 \\ 1 & 8 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}. \tag{2.2.8}$$

Edelleen a_1 on selvästi ainoa oikea päätös. Toisaalta

$$\begin{aligned} V(a_1) &= 1 \cdot 9 \cdot 1 \cdot (-1) \cdot 8 &= -72, \\ V(a_2) &= 1 \cdot 8 \cdot 0 \cdot (-2) \cdot 0 &= 0, \\ V(a_3) &= 0 \cdot 0 \cdot 0 \cdot (-2) \cdot 0 &= 0. \end{aligned}$$

Nyt päätössääntö (2.2.6), joka vielä tilanteessa (2.2.7) antoi järkevän tuloksen, antaa pöhkön tuloksen.

Päätössääntö on siis pöhkö, jos on olemassa päätöstilanne, jossa se antaa päätöksen, joka on dominoitu. Yllä päätössäännölle (2.2.6) tilanne (2.2.8) oli tällainen.

2.3 Suosittuja päätössääntöjä

Ei-stokastinen päätössääntö ei välitä millä tavalla Lady Fortuna valitsee toteutuvat maailmantilat:

2.3.1 Määritelmä (Ei-stokastinen päätössääntö)**.** Jos arvofunktion V määrittelyssä käytetään ainoastaan palkkioita R, mutta ei todennäköisyyksiä p, niin arvofunktiota V vastaava päätössääntö on *ei-stokastinen.*

Ei-stokastisen säännön vastakohta (tai pikemminkin yleistys) on tietenkin stokastinen sääntö:

2.3.2 Määritelmä (Stokastinen päätössääntö)**.** Mikäli arvofunktion $V : \mathcal{A} \to \bar{\mathbb{R}}$ määrittelyssä käytetään myös todennäköisyyksiä p, niin arvofunktiota V vastaava päätössääntö on *stokastinen.*

Esitämme nyt muutamia suosittuja päätössääntöjä ja niitä vastaavia päätöksiä seuraavan esimerkin tilanteessa. Lisäksi analysoimme hieman näiden sääntöjen mahdollista pöhköyttä.

2.3.3 Esimerkki (R. investoi)**.** Sijoittaja R., joka ei usko hajauttamiseen, vaan haluaa munansa yhteen koriin, haluaa sijoittaa tasan yhteen neljästä kohteesta:

$$\begin{aligned} a_1 &= \text{pankkitalletus}, \\ a_2 &= \text{valtion obligaatio}, \\ a_3 &= \text{vanhan vakavaraisen yrityksen osake}, \\ a_4 &= \text{uuden riskialttiin, mutta lupaavan yrityksen osake.} \end{aligned}$$

Sijoittaja R. arvelee, että jokin skenaarioista

$$\begin{aligned} \omega_1 &= \text{karhumarkkinat}, \\ \omega_2 &= \text{tasainen nykymeno}, \\ \omega_3 &= \text{sonnimarkkinat}, \end{aligned}$$

toteutuu ja että jokainen skenaario on yhtä todennäköinen. Eri sijoitusten vuotuiset tuottoprosentit eri skenaarioissa ovat

$$R = \begin{bmatrix} -1 & -1 & -1 \\ -5 & 0 & 5 \\ -10 & 2 & 8 \\ -50 & 3 & 30 \end{bmatrix}.$$

2.3.4 Huomautus. Esimerkin 2.3.3 päätöstilanteessa ei ole dominoituja valintoja. Toisin sanoen mikä tahansa päätös a_1, a_2, a_3 tai a_4 on periaatteessa järkevä. Siten periaattessa pöhkö sääntö voi näyttää ihan järkevältä esimerkin 2.3.3 valossa.

Ei-stokastisia sääntöjä

Aloitamme päätössäännöillä, jotka ovat riippumattomia päätöstilanteiden (R, p, a) todennäköisyyksistä p; siis ei-stokastisia.

2.3.5 Määritelmä (Pessimisti)**.** *Pessimisti* haluaa tehdä parhaan mahdollisen päätöksen huonointa mahdollista tilannetta varten:

$$V(a_i) = \min_{j \in J} r_{ij}.$$

2.3.6 Huomautus (Pessimisti on "maximin")**.** Pessimisti ajattelee, että Lady Fortuna on häntä vastaan ja varautuu siihen. Pessimistin valinnalle a_{i^*} pätee

$$V(a_{i^*}) = \max_{i \in I} \min_{j \in J} r_{ij}.$$

Tästä syystä pessimistin sääntöä kutsutaan myös *maximin-säännöksi.*

2.3.7 Esimerkki (Pessimisti R. investoi)**.** Esimerkissä 2.3.3 pessimistinen R. arvottaa

$$\begin{aligned} V(a_1) &= \min(-1, -1, -1) &= -1, \\ V(a_2) &= \min(-5, 0, 5) &= -5, \\ V(a_3) &= \min(-10, 2, 8) &= -10 \\ V(a_4) &= \min(-50, 3, 30) &= -50. \end{aligned}$$

Siten pessimistisen R.:n valinta on a_1, koska

$$V(a_1) = -1 = \max(-1, -5, -10, -50).$$

Pessimistien päätössääntö on varsin suosittu ja vaikuttaa järkeenkäyvältä. Valitettavasti pessimistin valinta voi olla pöhkö, kuten seuraava esimerkki todistaa:

2.3.8 Esimerkki (Pöhkö pessimisti)**.** Olkoon päätöstilanne

$$R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Tällöin pessimistille $V(a_i) = 0$ kaikilla $i \in I = \{1, 2, 3\}$. Siten pessimisti voi valita esim a_3. Mutta a_3 on dominoitu (sekä a_1:llä että a_2:lla). Siten pessimisti teki pöhkön valinnan.

Pöhköt valinnat voidaan estää *hienostuneella pessimismillä*:

2.3.9 Määritelmä (Hienostunut pessimisti)**.** Hienostunut pessimisti on pessimisti, joka on ensin poissulkenut dominoidut vaihtoehdot päätösmatriisistaan:

$$V(a_i) = \begin{cases} -\infty, & \text{jos } a_i \text{ on dominoitu,} \\ \min_{j \in J} r_{ij}, & \text{muulloin.} \end{cases}$$

2.3.10 Lause. *Hienostunut pessimisti ei ole pöhkö.*

Todistus. Väite on triviaali: pöhköjä valintoja ei tehdä, koska niille on annettu arvo $-\infty$. □

Lauseen 2.3.10 "todistuksessa" ei viitattu mitenkään itse pessimistin arvofunktioon. Siten samalla "todistuksella" saamme seuraavan lauseen:

2.3.11 Lause (Päätössäännön hienostaminen)**.** *Olkoon V_0 mikä tahansa päätössääntö. Jos on olemassa sellainen päätös a_i, että $V_0(a_i) > -\infty$, niin tällöin* hienostettu päätössääntö

$$V(a_i) = \begin{cases} -\infty, & \textit{jos } a_i \textit{ on dominoitu,} \\ V_0(a_i), & \textit{muulloin.} \end{cases}$$

on ei-pöhkö.

2.3.12 Määritelmä (Optimisti)**.** *Optimisti* haluaa tehdä parhaan mahdollisen päätöksen parhainta mahdollista tilannetta varten:

$$V(a_i) = \max_{j \in J} r_{ij}.$$

2.3.13 Huomautus (Optimisti on "maximax")**.** Optimisti ajattelee, että Lady Fortuna on hänen puolellaan ja varautuu siihen. Optimistin valinnalle a_{i^*} pätee

$$V(a_{i^*}) = \max_{i \in I} \max_{j \in J} r_{ij}.$$

Tästä syystä optimistin sääntöä kutsutaan myös *maximax-säännöksi.*

2.3.14 Esimerkki (Optimisti R. investoi)**.** Esimerkissä 2.3.3 optimistinen R. arvottaa

$$V(a_1) = -1,\ V(a_2) = 5,\ V(a_3) = 8 \ \text{ ja } V(a_4) = 30.$$

Siten optimistinen R. valitsee a_4:n.

2.3.15 Huomautus (Pöhkö optimisti)**.** Kuten pessimisti, myös optimisti voi olla pöhkö. Tämän näyttäminen on harjoitustehtävä 2.2. Luonnollisesti optimismi voidaan hienostaa pessimismin tavoin ei-pöhköksi eliminoimalla ensiksi dominoidut päätökset.

Seuraavassa säännössä ei poikkeuksellisesti maksimoidakaan arvofunktiota, vaan minimoidaan katumusta.

2.3.16 Määritelmä (Katumuksen kaihtaja)**.** *Katumuksen kaihtaja* haluaa *minimoida* katumuksen

$$k_{ij} \quad = \quad \max_{i\in I} r_{ij} - r_{ij},$$

kun on valittu a_i, mutta sattuukin ω_j. Katumuksen kaihtaja tekee valinnan a_{i^*}, jolle suurin mahdollinen katumus

$$K(a_i) \quad = \quad \max_{j\in J} k_{ij}.$$

minimoituu:

$$K(a_{i^*}) \quad = \quad \min_{i\in I} K(a_i).$$

2.3.17 Huomautus (Katumuksen kaihtaja on "minimax-katumus")**.** Merkitään

$$r_j^* \quad = \quad \max_{i\in I} r_{ij}$$

Tällön r_j^* on se paras mahdollinen palkkio, joka oltaisiin saatu, jos oltaisiin tiedetty, että Lady fortuna valitsee maailmantilan ω_j. Tällöin oltaisiin tietysti valittu parasta mahdollista palkkiota r_j^* vastaava päätös $a_{i^*(j)}$. Koska Lady Fortunan valintaa $j \in J$ ei tiedetä, on pakko tehdä jokin valinta a_i. Tästä valinnasta seuraa kadutus k_{ij}, jos Lady Fortuna tekee valinnan ω_j. Pahimmillaan tämä kadutus voi olla

$$K(a_i) \quad = \quad \max_{j\in J} k_{ij}$$

Valitaan siis päätös a_{i^*}, jolla pahin mahdollinen katumus $K(a_i)$ minimoituu:

$$K(a_{i^*}) \quad = \quad \min_{i\in I} \max_{j\in J} k_{ij}.$$

Tästä johtuu nimi *minimax-katumus-sääntö.*

2.3.18 Huomautus (Katumuksen kaihtaja arvofunktion maksimoijana)**.** Katumuksen kaihtaminen voidaan esittää myös arvofunktion maksimoimisena. Nimittäin katumuksen karttaminen on riemastuksen rakastamisena. Jos k_{ij} on katumus, niin $\ell_{ij} = -k_{ij}$ on riemastus. Päätöksentekijä varautuu pienimpään mahdolliseen riemastukseen

$$V(a_i) \quad = \quad \min_{j \in J} \ell_{ij}.$$

Valinnalle a_{i^*}, joka maksimoi pienimmän mahdollisen riemastuksen pätee

$$\begin{aligned} V(a_{i^*}) &= \max_{i \in I} \min_{j \in J} \ell_{ij} \quad = \quad \max_{i \in I} \Big(- \max_{j \in J} \big(- \ell_{ij} \big) \Big) \\ &= \max_{i \in I} \Big(- \max_{j \in J} k_{ij} \Big) \quad = \quad - \min_{i \in I} \max_{j \in J} k_{ij} \\ &= -K(a_{i^*}). \end{aligned}$$

2.3.19 Esimerkki (Katumusta kaihtava R. investoi)**.** Esimerkissä 2.3.3 sijoittaja R.:n katumusmatriisi on

$$K \quad = \quad \begin{bmatrix} -1-(-1) & 3-(-1) & 30-(-1) \\ -1-(-5) & 3-0 & 30-5 \\ -1-(-10) & 3-2 & 30-8 \\ -1-(-50) & 3-3 & 30-30 \end{bmatrix} \quad = \quad \begin{bmatrix} 0 & 4 & 31 \\ 4 & 3 & 25 \\ 9 & 2 & 22 \\ 49 & 0 & 0 \end{bmatrix}.$$

Suurimmat mahdolliset katumukset ovat

$$K(a_1) = 31, \ K(a_2) = 25, \ K(a_3) = 22 \ \text{ ja } K(a_4) = 49.$$

Katumusta kaihtavan R.:n valinta: a_3.

2.3.20 Huomautus (Pöhkö katumuksen kaihtaja)**.** Kuten pessimisti, myös katumuksen kaihtaja voi olla pöhkö. Tämän näyttäminen on harjoitustehtävä 2.3. Luonnollisesti katumuksen kaihtaja voidaan hienostaa pessimismin tavoin ei-pöhköksi eliminoimalla ensiksi dominoidut päätökset.

Seuraava ns. Hurwiczin sääntö mahdollistaa keskitien (tai minkä kohdan tahansa tieltä) optimismin ja pessimismin väliltä.

2.3.21 Määritelmä (Hurwiczin sääntö)**.** Olkoon $0 \leq w \leq 1$. *Hurwiczin päätössääntö* on $w\cdot$maximax + $(1-w)\cdot$maximin eli

$$V(a_i) \quad = \quad w \max_{j \in J} r_{ij} \; + \; (1-w) \min_{j \in J} r_{ij}.$$

Painoa w kutsutaan päätöksentekijän *optimismin asteeksi.*

2.3.22 Huomautus. Jos $w = 1$, on Hurwiczin sääntö maximax-sääntö. Jos taas $w = 0$ on Hurwiczin sääntö maximin-sääntö. Siten on selvää, että myös Hurwiczin sääntö voi johtaa pöhköihin valintoihin, jos sitä ei hienonneta normaaliin tapaan.

2.3.23 Esimerkki (Realisti R. investoi)**.** Esimerkissä 2.3.3 “realistisella” sijoittaja R.:llä on $w = 0{,}5$. Hän arvottaa

$$
\begin{aligned}
V(a_1) &= 0{,}5\cdot(-1) + 0{,}5\cdot(-1) &= -1,\\
V(a_2) &= 0{,}5\cdot(-5) + 0{,}5\cdot 5 &= 0,\\
V(a_3) &= 0{,}5\cdot(-10) + 0{,}5\cdot 8 &= -1,\\
V(a_1) &= 0{,}5\cdot(-50) + 0{,}5\cdot 30 &= -10.
\end{aligned}
$$

“Realistin” valinta: a_2.

2.3.24 Huomautus (Optimismin aste päätöksistä)**.** Harva meistä tietää optimismimme asteen w. Sen takia Hurwiczin sääntö onkin mielenkiintoinen käänteisesti: w voidaan tietyissä rajoissa päätellä tehdyistä päätöksistä. Jätämme harjoitustehtäväksi 2.4 keksiä koejärjestelyn, jolla optimismin aste voidaan määritellä.

Stokastisia sääntöjä

Tarkastelemme nyt päätössääntöjä, joissa todennäköisyydet p on otettava huomioon; siis stokastisia sääntöjä.

2.3.25 Määritelmä (Odotusarvosääntö)**.** *Odotusarvosääntö* tarkoittaa sitä, että arvotetaan odotusarvojen mukaan:

$$
V(a_i) \quad = \quad \mathrm{E}(a_i) \quad = \quad \sum_{j\in J} r_{ij}\, p_j .
$$

2.3.26 Huomautus (Odotusarvosääntö on “maxiE”)**.** Odotusarvoilla arvottavan valinnalle a_{i^*} pätee

$$
V(a_{i^*}) \quad = \quad \max_{i\in I} \mathrm{E}(a_i).
$$

Tämän vuoksi odotusarvosääntöä kutsutaan myös nimellä *maxiE-sääntö*.

2.3.27 Huomautus (Odotusarvosäännön luonnollisuus)**.** Odotusarvosääntö on monessa mielessä luonnollinen:

(i) Odotusarvo on painotettu keskiarvo, mikä on *intuitiivisesti tasapainoinen valinta*.

(ii) Odotusarvo on *varianssivirheen minimoija*: Jos on etsittävä luku m, jolle satunnaismuuttujan $X-m$ varianssi $\mathrm{Var}(X-m)$ on mahdollisimman pieni, on tuo luku odotusarvo: $m = \mathrm{E}(X)$

(iii) *Suurten lukujen laki* 1.3.19 sanoo, että jos pelataan toistuvasti peliä X, siis joka kierroksella saadaan satunnaismuuttujan X verran, niin “pitkässä juoksussa” saadaan keskimäärin $\mathrm{E}(X)$ verran.

(iv) Odotusarvo on myös *riskineutraali valinta*: jos päätöksentekijä on indifferentti valintojen

(a) varma voitto 1,

(b) voitto 2 todennäköisyydella 1/2 ja voitto 0 todennäköisyydellä 1/2,

on hän riskineutraali.

2.3.28 Esimerkki (Riskineutraali R. investoi)**.** Oletamme, että esimerkissä 2.3.3 jokainen skenaario ω_1, ω_2 ja ω_3 ovat yhtä todennäköisiä. Tällöin odotusarvon mukaan sijoittaja R.:n arvotukset ovat

$$\begin{aligned}
V(a_1) &= (-1)\cdot 0{,}333 + (-1)\cdot 0{,}333 + (-1)\cdot 0{,}333 &&= -1, \\
V(a_2) &= (-5)\cdot 0{,}333 + 0\cdot 0{,}333 + 5\cdot 0{,}333 &&= 0, \\
V(a_3) &= (-10)\cdot 0{,}333 + 2\cdot 0{,}333 + 8\cdot 0{,}333 &&= 0, \\
V(a_4) &= (-50)\cdot 0{,}333 + 3\cdot 0{,}333 + 30\cdot 0{,}333 &&= -5{,}667.
\end{aligned}$$

Sijoittaja R.:n valinta: joko a_2 tai a_3.

Toisin kuin aikaisemmissa säännöissä odotusarvosäännössä ei ole pöhköyden vaaraa.

2.3.29 Lause (Riskneutraali ei ole koskaan pöhkö)**.** *Odotusarvosääntöä noudattava päätöksentekijä ei koskaan valitse dominoitua päätöstä.*

Todistus. Oletetaan, että päätös a_{i_1} on dominoitu päätöksellä a_{i_2}. Tällön $r_{i_1 j} \leq r_{i_1 j}$ kaikilla $j \in J$ ja on olemassa jokin $j^\dagger \in J$, jolle $r_{i_1 j^\dagger} < r_{i_2 j^\dagger}$. Tällöin

$$\begin{aligned}
\mathrm{E}\,(a_{i_1}) &= \sum_{j\in J} r_{i_1 j} p_j \\
&= \sum_{j\in J\setminus\{j^\dagger\}} r_{i_1 j} p_j + r_{i_1 j^\dagger} p_{j^\dagger} \\
&\leq \sum_{j\in J\setminus\{j^\dagger\}} r_{i_2 j} p_j + r_{i_1 j^\dagger} p_{j^\dagger} \\
&< \sum_{j\in J\setminus\{j^\dagger\}} r_{i_2 j} p_j + r_{i_2 j^\dagger} p_{j^\dagger} \\
&= \sum_{j\in J} r_{i_2 j} p_j \\
&= \mathrm{E}\,(a_{i_2}).
\end{aligned}$$

Siispä odotusarvoarvoittaja ei valitse dominoitua päätöstä a_{i_1}. $\square$

Odotusarvosääntö on siis riskineutraali. Kaikki eivät kuitenkaan ole riskineutraaleja: toisen karttavat riskiä (ja ottavat esimerkiksi vakuutuksia) ja toiset rakastavat riskiä (ja esimerkiksi lottoavat). Suhtautumista riskiin mitataan hyötyjen avulla. Idea on muuttaa palkkiot hyödyiksi epälineaarisella tavalla. Tällöin päätöksentekijä voi suhtautua palkkioiden muutoksiin eri tavalla riippuen siitä kuinka suuri palkkio on tai mikä on päätöksentekijän alkuvarallisuus. Esimerkiksi kymmenen euron muutos kymmenen euron palkkiossa merkitsee monille (riskiä kaihtaville) päätöksentekijöille enemmän kuin kymmenen euron muutos miljoonan euron palkkiossa. Toisin sanoen rajahyöty ei ole vakio.

2.3.30 Määritelmä (Hyötyfunktio)**.** Höytyfunktio on mikä tahansa jatkuva ja kasvava funktio $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$.

2.3.31 Määritelmä (Odotetun hyödyn sääntö)**.** *Odotetun hyödyn sääntö* on kuten odotusarvosääntö, mutta nyt ei tarkastella palkkioita r_{ij}, vaan niiden hyötyjä $u(r_{ij})$. Odotetun hyödyn maksimoijan arvofunktio on siis

$$V(a_i) \quad = \quad \mathrm{E}\,(u(a_i)) \quad = \quad \sum_{j \in J} u\,(r_{ij})\; p_j,$$

missä $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ on päätöksentekijän *hyötyfunktio*.

2.3.32 Huomautus (Odotetun hyödyn sääntö on “maxiEu”)**.** Odotetun hyödyn maksimoijan valinta on se a_{i^*}, jolle

$$V(a_{i^*}) \quad = \quad \max_{i \in I} \mathrm{E}\,(u(a_i))\,.$$

Tästä syystä odotetun hyödyn sääntöä kutsutaan *maxiEu-säännöksi*.

2.3.33 Määritelmä (Rajahyöty)**.** Hyötyfunktion $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ derivaattafunktio $\frac{\mathrm{d}u}{\mathrm{d}r} : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ on *rajahyöty*.

2.3.34 Huomautus (Muuttuva rajahyöty)**.** Koska hyötysäännön 2.3.31 keskeinen idea on muuttuva rajahyöty, hyötylähestymistavassa on aina katsottava kokonaistilannetta, eikä vain muutoksia. Riskineutraalissa lähestymistavassa muutosten tarkastelu riittää. Tai pikemminkin riskineutraalissa tilanteessa ei ole väliä katsotaanko muutoksia vai lopputulemaa.

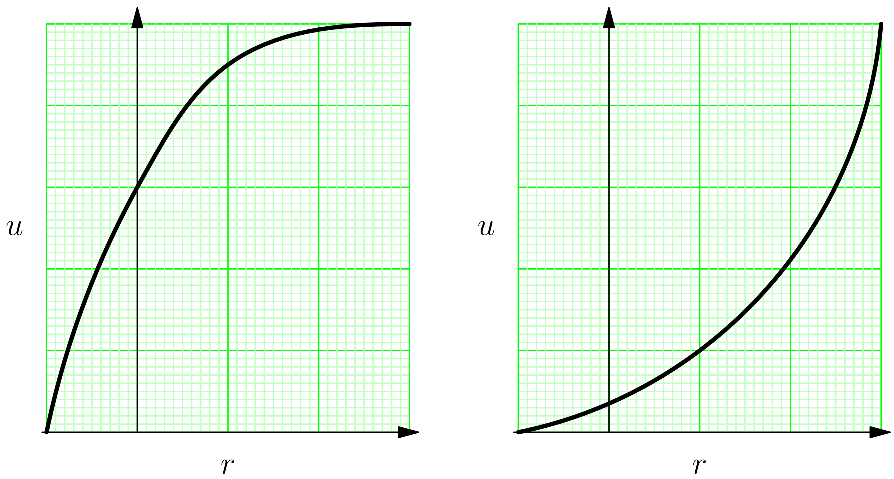
2.3.35 Määritelmä (Riskinkaihtaja ja -rakastaja sekä riskineutraali)**.** Jos hyötyfunktion u derivaatta $\frac{\mathrm{d}u}{\mathrm{d}r}$, eli *rajahyöty*, on laskeva, niin hyötyfunktio u on *konkaavi*. Päätöksentekijä, jolla on konkaavi hyötyfunktio on *riskinkaihtaja*.

Vastaavasti, jos rajahyöyty $\frac{\mathrm{d}u}{\mathrm{d}r}$ on kasvava, on hyöty *konveksi*, ja päätöksentekijä, jolla on konveksi hyötyfunktio on *riskinrakastaja*.

Jos rajahyöty $\frac{\mathrm{d}u}{\mathrm{d}r}$ ei kasva eikä laske, eli se on vakio, niin päätöksentekijä on *riskineutraali*.

Hyödyistä, riskinkaihtamisesta ja -rakastamisesta puhumme myöhemmin lisää paljonkin luvussa 5.

Seuraavassa kuvassa on hahmoteltu riskinkaihtajan (vasen puoli, konkaavi funktio) ja riskinrakastajan (oikea puoli, konveksi funktio) hyötyfunktioita.



2.3.36 Lause (Hyötysäännön pöhköys ja ei-pöhköys)**.** *Hyötysääntö ei ole pöhkö jos ja vain jos hyötyfunktio* $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ *on* aidosti *kasvava.*

Todistus. Lauseessa 2.3.36 on kaksi puolta: ensimmäinen puoli sanoo, että aidosti kasvavaa hyötyfunktiota vastaava päätössääntö ei voi olla pöhkö ja toinen puoli sanoo, että jos hyötyfunktio ei ole aidosti kasvava, niin löytyy päätöstilanne, jossa se on pöhkö.

Tarkastelemme aluksi ensimmäistä puolta. Oletamme, että päätös a_{i_1} on dominoitu päätöksellä a_{i_2}. Tällöin $r_{i_1 j} \leq r_{i_1 j}$ kaikilla $j \in J$ ja on olemassa jokin $j^\dagger \in J$, jolle $r_{i_1 j^\dagger} < r_{i_2 j^\dagger}$. Tällöin, koska u on aidosti kasvava, niin perustelu menee täsmälleen samalla tavalla kuin riskineutraalissa tapauksessa:

$$
\begin{aligned}
\mathrm{E}\left(u(a_{i_1})\right) &= \sum_{j \in J} u\left(r_{i_1 j}\right) p_j \\
&= \sum_{j \in J \setminus \{j^\dagger\}} u\left(r_{i_1 j}\right) p_j + u\left(r_{i_1 j^\dagger}\right) p_{j^\dagger} \\
&\leq \sum_{j \in J \setminus \{j^\dagger\}} u\left(r_{i_2 j}\right) p_j + u\left(r_{i_1 j^\dagger}\right) p_{j^\dagger} \\
&< \sum_{j \in J \setminus \{j^\dagger\}} u\left(r_{i_2 j}\right) p_j + u\left(r_{i_2 j^\dagger}\right) p_{j^\dagger} \\
&= \sum_{j \in J} u\left(r_{i_2 j}\right) p_j \\
&= \mathrm{E}\left(u(a_{i_2})\right).
\end{aligned}
$$

Siispä dominoitua päätöstä a_{i_1} ei valita.

Tarkastelemme sitten väitteen toista puolta. Riittää siis osoittaa, että jos $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ joka ei ole *aidosti* kasvava, niin löytyy päätöstilanne (R, p, a), jossa tehdään pöhkö päätös. Koska u:n on joka tapauksessa oltava kasvava (muuten pöhköys on ilmiselvää) voimme olettaa, että u on vakio jollakin välillä $[r_1, r_2]$. Tällöin deterministinen päätöstilanne

$$R = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

altistaa pöhköydelle, sillä

$$V(a_1) = u(r_1) = u(r_2) = V(a_2),$$

mutta a_1 on selvästi dominoitu. □

2.3.37 Esimerkki (Riskiä kaihtava R. investoi)**.** Olkoon esimerkissä 2.3.3 sijoittaja R.:n hyöty logaritminen tuottoprosenttiin nähden:

$$u(r) = \ln\big(r + 100\%\big).$$

Tällöin hänen rajahyötynsä on kääntäen verrannollinen tuottoon nähden, eli

$$\frac{\mathrm{d}}{\mathrm{d}r}u(r) = \frac{1}{r}.$$

Luku 100% sijoittaja R.:n hyötyfunktiossa viittaa R.:n koko sijoitettavaan varallisuuteen. Oikeastaan olisi parempi käyttää rahamääriä tuoton sijasta, koska sijoittaja R.:n hyötyfunktio voi riippua hänen (sijoitettavasta) varallisuudestaan. Sijoittaja R.:n arvotukset ovat (kun edelleen oletamme, että skenaariot ω_1, ω_2 ja ω_3 ovat yhtä todennäköisiä)

$$\begin{aligned}
V(a_1) &= \ln(0{,}99)\cdot 0{,}333 + \ln(0{,}99)\cdot 0{,}333 + \ln(0{,}99)\cdot 0{,}333 \\
&= -0{,}003, \\
V(a_2) &= \ln(0{,}95)\cdot 0{,}333 + \ln(1)\cdot 0{,}333 + \ln(1{,}05)\cdot 0{,}333 \\
&= -0{,}001, \\
V(a_3) &= \ln(0{,}90)\cdot 0{,}333 + \ln(1{,}02)\cdot 0{,}333 + \ln(1{,}08)\cdot 0{,}333 \\
&= -0{,}003, \\
V(a_4) &= \ln(0{,}50)\cdot 0{,}333 + \ln(1{,}03)\cdot 0{,}333 + \ln(1{,}30)\cdot 0{,}333 \\
&= -0{,}133.
\end{aligned}$$

Sijoittaja R.:n valinta: a_2.

2.4 Yhdistettyjä sääntöjä

2.4.1 Määritelmä (Yhdistetty sääntö)**.** Päätössäännöistä $V_1, V_2, \ldots, V_m$ voidaan muodostaa uusi sääntö V painottamalla:

$$V(a_i) \;=\; w_1V_1(a_i) + w_2V_2(a_i) + \cdots + w_mV_m(a_i),$$

missä painot $w_1, w_2, \ldots, w_m$ ovat positiivisia.

2.4.2 Määritelmä (Koveksi sääntöjen yhdistäminen)**.** Mikäli määritelmässä 2.4.1 yhdistämispainot $w_1, \ldots, w_m$ toteuttavat

$$w_1 + w_2 + \cdots + w_m \;=\; 1$$

on kyseessä *konveksi yhdistäminen*.

2.4.3 Esimerkki. Hurwiczin sääntö on konveksi yhdistelmä maximax- ja maximin-säännöistä.

2.4.4 Esimerkki (Sijoitta R. yhdistelee sääntöjä)**.** Esimerkin 2.3.3 sijoittaja R. painottaa optimistista sääntöä (maximax) painolla 0,75 ja katumussääntöä (minimax-katumus) painolla 0,25. Yhdistämisessä käytetään riemastustulkintaa katumukselle. Arvotukset ovat:

$$\begin{aligned}
V(a_1) &= 0{,}75\cdot(-1) \,+\, 0{,}25\cdot(-31) &&= -8{,}50,\\
V(a_2) &= 0{,}75\cdot 5 \,+\, 0{,}25\cdot(-25) &&= -2{,}50,\\
V(a_3) &= 0{,}75\cdot 8 \,+\, 0{,}25\cdot(-22) &&= 0{,}50,\\
V(a_4) &= 0{,}75\cdot 30 \,+\, 0{,}25\cdot(-49) &&= 10{,}25.
\end{aligned}$$

R.:n valinta: a_4.

2.4.5 Esimerkki (Eri arvojen tasapainottaminen)**.** Yhtiön pitää valita mikä päätös a_i, $i \in I$, tulee valita, kun palkkiomatriisi on $R = [r_{ij}]_{i\in I, j\in J}$. Yhtiöllä on kolme osakkeenomistajaa, joiden osuudet koko osakekannasta ovat w_1, w_2 ja $w_3 = 1 - w_1 - w_2$. Osakkeenomistajilla on arvofunktiot V_1, V_2 ja V_3. Koska esimerkiksi arvofuntkio $10V_1$ on käytännössä sama kuin arvofuntkio V_1, on osakkeenomistajien arvofunktiot ensiksi laitettava samalle skaalalle. Tämä tapahtuu normeerauksella

$$\tilde{V}_k(a_i) \;=\; \frac{V_k(a_i) - V_k(a_{i(k,-)})}{V_k(a_{i(k,+)}) - V_k(a_{i(k,-)})}, \qquad k = 1, 2, 3,$$

missä $a_{i(k,+)}$ ja $a_{i(k,-)}$ ovat osakkeenomistaja k:n eniten ja väiten arvottamat päätökset. Nyt kaikilla osakkeenomistajilla k ja kaikilla päätöksillä a_i, $0 \leq V_k(a_i) \leq 1$ ja osakeomistuksilla painotettu yhteinen arvofunktio on

$$V(a_i) \;=\; w_1\tilde{V}_1(a_i) + w_2\tilde{V}_2(a_i) + w_3\tilde{V}_3(a_i).$$

2.4.6 Määritelmä (Reunaehdon yhdistäminen)**.** *Reunaehto* V' on mikä tahansa sääntö, joka sulkee pois jotkin valinnat a_i, mutta ei erottele muita valintoja: $V'(a_i) = -1$, jos valinta a_i ei toteuta reunaehtoa ja $V'(a_i) = 0$, jos valinta a_i toteuttaa reunaehdon. Sääntöön V'' voidaan yhdistää reunaehto V' asettamalla[3]

$$V(a_i) \quad = \quad V''(a_i) + \infty \cdot V'(a_i).$$

2.4.7 Esimerkki (Sijoittaja R. asettaa reunaehtoja)**.** Esimerkissä 2.3.3 sijoittaja R. asettaa reunaehdon "yli 15% sijoituksesta ei saa missään tapauksessa hävitä". Siis $V'(a_i) = -1$, kun $i = 4$ ja $V'(a_i) = 0$ muulloin. Jos perussääntö V'' on optimistinen maximax, niin

$$\begin{aligned}
V(a_1) &= -1 + \infty \cdot 0 = -1, \\
V(a_2) &= 5 + \infty \cdot 0 = 5, \\
V(a_3) &= 8 + \infty \cdot 0 = 8, \\
V(a_4) &= 30 + \infty \cdot (-1) = -\infty
\end{aligned}$$

R.:n valinta: a_3.

Lopuksi esitämme laajan esimerkin, jossa kertaamme tämän luvun asioita:

2.4.8 Esimerkki. Leipuri Pulla myy pullia Kumputien Leipomossa. Hän paistaa pullat aamulla ja myy ne lounastauolla viereisen Ministeriön Erikoisosaston virkamiehille. Eilisiä pullia ei voi tänään enää myydä.

Pullan paistaminen maksaa leipuri Pullalle 0,20€ pullalta, ja hän myy niitä 1,00€ kappalehintaan. Leipuri Pulla tietää, että pullia myydään 0:sta 10:een lounastaukoa kohti. Itse asiassa hän arvioi, että

$$\begin{aligned}
p_0 &= 0{,}01, & p_3 &= 0{,}04, & p_6 &= 0{,}10, & p_9 &= 0{,}02, \\
p_1 &= 0{,}02, & p_4 &= 0{,}10, & p_7 &= 0{,}04, & p_{10} &= 0{,}01, \\
p_2 &= 0{,}03, & p_5 &= 0{,}60, & p_8 &= 0{,}03,
\end{aligned}$$

missä p_j on todennäköisyys myydä j pullaa.

Kuinka monta pullaa leipuri Pulla paistaa aamulla, jos hän on

(a) optimisti (Maximax),
(b) pessimisti (Maximin),
(c) realisti (1/2·Maximax + 1/2·Maximin),
(d) katumuksen kaihtaja,
(e) riskineutraali odotusarvoon uskoja,
(f) riskineutraali odotusarvoon uskoja, jolla kuitenkin on seuraavat reunaehdot:

[3] Tunnetusti $\infty \cdot 0 = 0$.

– pitää olla mahdollista saada voittoa vähintää 1,00 euroa ja
– missään tapauksessa ei saa tulla tappiota yli 1,00 euroa?

Aluksi on hyvä huomata, että säännöt (a)–(d) ovat ei-stokastisia. Siten todennäköisyydet $p = (p_j)$ ovat niissä irrelevantteja.

Seuraavaksi määräämme palkkio- eli päätösmatriisin $R = [r_{ij}]$. Leipuri Pullan järkevät pullanpaistomäärät ovat $A = \{0, 1, \ldots, 10\}$. Jos kaikki pullat myydään $(i \le j)$, niin leipuri Pullan voitto on

$$1{,}00 \cdot i - 0{,}20 \cdot i \quad = \quad 0{,}80 \cdot i.$$

Jos taas pullia jää myymättä $(i > j)$, niin leipuri Pullan voitto on

$$1{,}00 \cdot j - 0{,}20 \cdot i.$$

Palkkiomatriisi R, eli voitot eri pullanpaisto- ja -myyntimäärillä, on

$$R = \begin{bmatrix}
0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 0{,}00 \\
-0{,}20 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 & 0{,}80 \\
-0{,}40 & 0{,}60 & 1{,}60 & 1{,}60 & 1{,}60 & 1{,}60 & 1{,}60 & 1{,}60 & 1{,}60 & 1{,}60 & 1{,}60 \\
-0{,}60 & 0{,}40 & 1{,}40 & 2{,}40 & 2{,}40 & 2{,}40 & 2{,}40 & 2{,}40 & 2{,}40 & 2{,}40 & 2{,}40 \\
-0{,}80 & 0{,}20 & 1{,}20 & 2{,}20 & 3{,}20 & 3{,}20 & 3{,}20 & 3{,}20 & 3{,}20 & 3{,}20 & 3{,}20 \\
-1{,}00 & 0{,}00 & 1{,}00 & 2{,}00 & 3{,}00 & 4{,}00 & 4{,}00 & 4{,}00 & 4{,}00 & 4{,}00 & 4{,}00 \\
-1{,}20 & -0{,}20 & 0{,}80 & 1{,}80 & 2{,}80 & 3{,}80 & 4{,}80 & 4{,}80 & 4{,}80 & 4{,}80 & 4{,}80 \\
-1{,}40 & -0{,}40 & 0{,}60 & 1{,}60 & 2{,}60 & 3{,}60 & 4{,}60 & 5{,}60 & 5{,}60 & 5{,}60 & 5{,}60 \\
-1{,}60 & -0{,}60 & 0{,}40 & 1{,}40 & 2{,}40 & 3{,}40 & 4{,}40 & 5{,}40 & 6{,}40 & 6{,}40 & 6{,}40 \\
-1{,}80 & -0{,}80 & 0{,}20 & 1{,}20 & 2{,}20 & 3{,}20 & 4{,}20 & 5{,}20 & 6{,}20 & 7{,}20 & 7{,}20 \\
-2{,}00 & -1{,}00 & 0{,}00 & 1{,}00 & 2{,}00 & 3{,}00 & 4{,}00 & 5{,}00 & 6{,}00 & 7{,}00 & 8{,}00
\end{bmatrix}.$$

(a) Optimisti Pulla toivoo parasta ja arvottaa tilat seuraavasti:

$$\begin{aligned}
V(a_0) &= \max_{j\in J} r_{0j} = 0{,}00, & V(a_6) &= \max_{j\in J} r_{6j} = 4{,}80,\\
V(a_1) &= \max_{j\in J} r_{1j} = 0{,}80, & V(a_7) &= \max_{j\in J} r_{7j} = 5{,}60,\\
V(a_2) &= \max_{j\in J} r_{2j} = 1{,}60, & V(a_8) &= \max_{j\in J} r_{8j} = 6{,}40,\\
V(a_3) &= \max_{j\in J} r_{3j} = 2{,}40, & V(a_9) &= \max_{j\in J} r_{9j} = 7{,}20,\\
V(a_4) &= \max_{j\in J} r_{4j} = 3{,}20, & V(a_{10}) &= \max_{j\in J} r_{10,j} = 8{,}00.\\
V(a_5) &= \max_{j\in J} r_{5j} = 4{,}00, & &
\end{aligned}$$

Optimistin valinta: kaikki pullat uuniin.

(b) Pessimisti Pulla pelkää pahinta ja arvottaa tilat seuraavasti:

$$
\begin{aligned}
V(a_0) &= \min_{j\in J} r_{0j} = 0{,}00, & V(a_6) &= \min_{j\in J} r_{6j} = -1{,}20,\\
V(a_1) &= \min_{j\in J} r_{1j} = -0{,}20, & V(a_7) &= \min_{j\in J} r_{7j} = -1{,}40,\\
V(a_2) &= \min_{j\in J} r_{2j} = -1{,}40, & V(a_8) &= \min_{j\in J} r_{8j} = -1{,}60,\\
V(a_3) &= \min_{j\in J} r_{3j} = -1{,}60, & V(a_9) &= \min_{j\in J} r_{9j} = -1{,}80,\\
V(a_4) &= \min_{j\in J} r_{4j} = -0{,}80, & V(a_{10}) &= \min_{j\in J} r_{10,j} = -2{,}00.\\
V(a_5) &= \min_{j\in J} r_{5j} = -1{,}00, & &
\end{aligned}
$$

Pessimistin valinta: ei paisteta.

(c) Realisti Pullan arvotukset ovat

$$
\begin{aligned}
V(a_0) &= 0{,}5\cdot 0{,}00 + 0{,}5\cdot 0{,}00 = 0{,}00,\\
V(a_1) &= 0{,}5\cdot 0{,}80 + 0{,}5\cdot(-0{,}20) = 0{,}30,\\
V(a_2) &= 0{,}5\cdot 1{,}60 + 0{,}5\cdot(-0{,}40) = 0{,}60,\\
V(a_3) &= 0{,}5\cdot 2{,}40 + 0{,}5\cdot(-0{,}60) = 0{,}90,\\
V(a_4) &= 0{,}5\cdot 3{,}20 + 0{,}5\cdot(-0{,}80) = 1{,}20,\\
V(a_5) &= 0{,}5\cdot 4{,}00 + 0{,}5\cdot(-1{,}00) = 1{,}50,\\
V(a_6) &= 0{,}5\cdot 4{,}80 + 0{,}5\cdot(-1{,}20) = 1{,}80,\\
V(a_7) &= 0{,}5\cdot 5{,}60 + 0{,}5\cdot(-1{,}40) = 2{,}10,\\
V(a_8) &= 0{,}5\cdot 6{,}40 + 0{,}5\cdot(-1{,}60) = 2{,}40,\\
V(a_9) &= 0{,}5\cdot 7{,}20 + 0{,}5\cdot(-1{,}80) = 2{,}70,\\
V(a_{10}) &= 0{,}5\cdot 8{,}00 + 0{,}5\cdot(-2{,}00) = 3{,}00.
\end{aligned}
$$

Realistin valinta on paistaa kymmenen pullaa.

(d) Leipuri Pullan katumusmatriisi on

$$
K = \begin{bmatrix}
0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20 & 4{,}00 & 4{,}80 & 5{,}60 & 6{,}40 & 7{,}20 & 8{,}00\\
0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20 & 4{,}00 & 4{,}80 & 5{,}60 & 6{,}40 & 7{,}20\\
0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20 & 4{,}00 & 4{,}80 & 5{,}60 & 6{,}40\\
0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20 & 4{,}00 & 4{,}80 & 5{,}60\\
0{,}80 & 0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20 & 4{,}00 & 4{,}80\\
1{,}00 & 0{,}80 & 0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20 & 4{,}00\\
1{,}20 & 1{,}00 & 0{,}80 & 0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40 & 3{,}20\\
1{,}40 & 1{,}20 & 1{,}00 & 0{,}80 & 0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60 & 2{,}40\\
1{,}60 & 1{,}40 & 1{,}20 & 1{,}00 & 0{,}80 & 0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80 & 1{,}60\\
1{,}80 & 1{,}60 & 1{,}40 & 1{,}20 & 1{,}00 & 0{,}80 & 0{,}60 & 0{,}40 & 0{,}20 & 0{,}00 & 0{,}80\\
2{,}00 & 1{,}80 & 1{,}60 & 1{,}40 & 1{,}20 & 1{,}00 & 0{,}80 & 0{,}60 & 0{,}40 & 0{,}30 & 0{,}00
\end{bmatrix}.
$$

Suurimmat mahdolliset katumukset eri valinnoilla ovat siis

$$
\begin{aligned}
K(a_0) &= \max_{j\in J} k_{0j} = 8{,}00, & K(a_6) &= \max_{j\in J} k_{6j} = 3{,}20,\\
K(a_1) &= \max_{j\in J} k_{1j} = 7{,}20, & K(a_7) &= \max_{j\in J} k_{7j} = 2{,}40,\\
K(a_2) &= \max_{j\in J} k_{2j} = 6{,}40, & K(a_8) &= \max_{j\in J} k_{8j} = 1{,}60,\\
K(a_3) &= \max_{j\in J} k_{3j} = 5{,}60, & K(a_9) &= \max_{j\in J} k_{9j} = 1{,}80,\\
K(a_4) &= \max_{j\in J} k_{4j} = 4{,}80, & K(a_{10}) &= \max_{j\in J} k_{10,j} = 2{,}00.\\
K(a_5) &= \max_{j\in J} k_{5j} = 4{,}00, & &
\end{aligned}
$$

Katumuksen kaihtaja paistaa 8 pullaa.

(e) Riskineutraalin Pullan arvotukset ovat

$$
\begin{aligned}
V(a_0) = {} & 0{,}00\cdot 0{,}01 + 0{,}00\cdot 0{,}02 + 0{,}00\cdot 0{,}03 + 0{,}00\cdot 0{,}04 + \\
& 0{,}00\cdot 0{,}10 + 0{,}00\cdot 0{,}60 + 0{,}00\cdot 0{,}10 + 0{,}00\cdot 0{,}04 + \\
& 0{,}00\cdot 0{,}03 + 0{,}00\cdot 0{,}02 + 0{,}00\cdot 0{,}01 = 0,
\end{aligned}
$$

$$
\begin{aligned}
V(a_1) = {} & -0{,}20\cdot 0{,}01 + 0{,}80\cdot 0{,}02 + 0{,}80\cdot 0{,}03 + 0{,}80\cdot 0{,}04 + \\
& 0{,}80\cdot 0{,}10 + 0{,}80\cdot 0{,}60 + 0{,}80\cdot 0{,}10 + 0{,}80\cdot 0{,}04 + \\
& 0{,}80\cdot 0{,}03 + 0{,}80\cdot 0{,}02 + 0{,}80\cdot 0{,}01 = 0{,}79,
\end{aligned}
$$

$$
\begin{aligned}
V(a_2) = {} & -0{,}40\cdot 0{,}01 + 0{,}60\cdot 0{,}02 + 1{,}60\cdot 0{,}03 + 1{,}60\cdot 0{,}04 + \\
& 1{,}60\cdot 0{,}10 + 1{,}60\cdot 0{,}60 + 1{,}60\cdot 0{,}10 + 1{,}60\cdot 0{,}04 + \\
& 1{,}60\cdot 0{,}03 + 1{,}60\cdot 0{,}02 + 1{,}60\cdot 0{,}01 = 1{,}56,
\end{aligned}
$$

$$
\begin{aligned}
V(a_3) = {} & -0{,}60\cdot 0{,}01 + 0{,}40\cdot 0{,}02 + 1{,}40\cdot 0{,}03 + 2{,}40\cdot 0{,}04 + \\
& 2{,}40\cdot 0{,}10 + 2{,}40\cdot 0{,}60 + 2{,}40\cdot 0{,}10 + 2{,}40\cdot 0{,}04 + \\
& 2{,}40\cdot 0{,}03 + 2{,}40\cdot 0{,}02 + 2{,}40\cdot 0{,}01 = 2{,}30,
\end{aligned}
$$

$$
\begin{aligned}
V(a_4) = {} & -0{,}80\cdot 0{,}01 + 0{,}20\cdot 0{,}02 + 1{,}20\cdot 0{,}03 + 2{,}20\cdot 0{,}04 + \\
& 3{,}20\cdot 0{,}10 + 3{,}20\cdot 0{,}60 + 3{,}20\cdot 0{,}10 + 3{,}20\cdot 0{,}04 + \\
& 3{,}20\cdot 0{,}03 + 3{,}20\cdot 0{,}02 + 3{,}20\cdot 0{,}01 = 3{,}09,
\end{aligned}
$$

$$
\begin{aligned}
V(a_5) = {} & -1{,}00\cdot 0{,}01 + 0{,}00\cdot 0{,}02 + 1{,}00\cdot 0{,}03 + 2{,}00\cdot 0{,}04 + \\
& 3{,}00\cdot 0{,}10 + 4{,}00\cdot 0{,}60 + 4{,}00\cdot 0{,}10 + 4{,}00\cdot 0{,}04 + \\
& 4{,}00\cdot 0{,}03 + 4{,}00\cdot 0{,}02 + 4{,}00\cdot 0{,}01 = 3{,}60,
\end{aligned}
$$

$$\begin{aligned} V(a_6) &= -1{,}20 \cdot 0{,}01 - 0{,}20 \cdot 0{,}02 + 0{,}80 \cdot 0{,}03 + 1{,}80 \cdot 0{,}04 + \\ &\quad 2{,}80 \cdot 0{,}10 + 3{,}80 \cdot 0{,}60 + 4{,}80 \cdot 0{,}10 + 4{,}80 \cdot 0{,}04 + \\ &\quad 4{,}80 \cdot 0{,}03 + 4{,}80 \cdot 0{,}02 + 4{,}80 \cdot 0{,}01 \quad = \quad 4{,}08, \end{aligned}$$

$$\begin{aligned} V(a_7) &= -1{,}40 \cdot 0{,}01 - 0{,}40 \cdot 0{,}02 + 0{,}60 \cdot 0{,}03 + 1{,}60 \cdot 0{,}04 + \\ &\quad 2{,}60 \cdot 0{,}10 + 3{,}60 \cdot 0{,}60 + 4{,}60 \cdot 0{,}10 + 5{,}60 \cdot 0{,}04 + \\ &\quad 5{,}60 \cdot 0{,}03 + 5{,}60 \cdot 0{,}02 + 5{,}60 \cdot 0{,}01 \quad = \quad 4{,}06, \end{aligned}$$

$$\begin{aligned} V(a_8) &= -1{,}60 \cdot 0{,}01 - 0{,}60 \cdot 0{,}02 + 0{,}40 \cdot 0{,}03 + 1{,}40 \cdot 0{,}04 + \\ &\quad 2{,}40 \cdot 0{,}10 + 3{,}40 \cdot 0{,}60 + 4{,}40 \cdot 0{,}10 + 5{,}40 \cdot 0{,}04 + \\ &\quad 6{,}40 \cdot 0{,}03 + 6{,}40 \cdot 0{,}02 + 6{,}40 \cdot 0{,}01 \quad = \quad 4{,}00, \end{aligned}$$

$$\begin{aligned} V(a_9) &= -1{,}80 \cdot 0{,}01 - 0{,}80 \cdot 0{,}02 + 0{,}20 \cdot 0{,}03 + 1{,}20 \cdot 0{,}04 + \\ &\quad 2{,}20 \cdot 0{,}10 + 3{,}20 \cdot 0{,}60 + 4{,}20 \cdot 0{,}10 + 5{,}20 \cdot 0{,}04 + \\ &\quad 6{,}20 \cdot 0{,}03 + 7{,}20 \cdot 0{,}02 + 7{,}20 \cdot 0{,}01 \quad = \quad 3{,}19, \end{aligned}$$

$$\begin{aligned} V(a_{10}) &= -2{,}00 \cdot 0{,}01 - 1{,}00 \cdot 0{,}02 + 0{,}00 \cdot 0{,}03 + 1{,}00 \cdot 0{,}04 + \\ &\quad 2{,}00 \cdot 0{,}10 + 3{,}00 \cdot 0{,}60 + 4{,}00 \cdot 0{,}10 + 5{,}00 \cdot 0{,}04 + \\ &\quad 6{,}00 \cdot 0{,}03 + 7{,}00 \cdot 0{,}02 + 8{,}00 \cdot 0{,}01 \quad = \quad 3{,}02. \end{aligned}$$

Riskineutraali valinta on siis paistaa 6 pullaa.

(f) Nyt paistetaan 5 pullaa, sillä a_5:llä on paras odotusarvo valintojen a_2, a_3, a_4, a_5 joukossa, jotka toteuttavat asetetut reunaehdot.

2.5 Harjoitustehtäviä lukuun 2

2.1 Harjoitustehtävä. Mitkä seuraavia arvofunktioita vastaavat päätössäännöt ovat pöhköjä?

(a) $V(a_i) = \sum_{j \in J} r_{ij}$.
(b) $V(a_i) = \prod_{j \in J} r_{ij}$.
(c) $V(a_i) = 0$.

2.2 Harjoitustehtävä. Osoita, että optimistin sääntö on pöhkö joissakin päätöstilanteissa.

2.3 Harjoitustehtävä. Osoita, että katumuksen kaihtajan sääntö on pöhkö joissakin päätöstilanteissa.

2.4 Harjoitustehtävä. Keksi koejärjestely, jolla päätöksentekijän optimismin aste w voidaan päätellä yhden desimaalin tarkkuudella.

2.5 Harjoitustehtävä. Penan Grilli ja Jaskan Grilli ovat kilpailijoita. Molempien täytyy päättää samanaikaisesti ja toisistaan tietämättä mainostaako

- ei ollenkaan,
- hieman,
- kohtalaisesti,
- paljon.

Pena uskoo, että Jaska mainostaa ei ollenkaan, hieman, kohtalaisesti tai paljon yhtä suurin todennäköisyyksin. Penan Grillin tuotot (euroa/vuosi) eri tilanteissa ovat:

	Jaskan valinta			
Penan valinta	Ei ollenkaan	Hieman	Kohtalaisesti	Paljon
Ei ollenkaan	90.000	10.000	10.000	10.000
Hieman	60.000	60.000	50.000	20.000
Kohtalaisesti	50.000	50.000	60.000	10.000
Paljon	90.000	90.000	50.000	0

Miten Penan tulisi mainostaa, jos hän on

(a) optimisti,
(b) pessimisti,
(c) realisti?

2.6 Harjoitustehtävä. Miten tehtävän 2.5 Penan tulisi mainostaa, jos hän on

(a) katumuksen kaihtaja,
(b) riskineutraali,
(c) hyödyn maksimoija hyötyfunktiolla $u(r) = \ln r$?

2.7 Harjoitustehtävä. Miten Penan ratkaisut tehtävissä 2.5 ja 2.6 muuttuvat, jos hän uskoo, että Jaska mainostaa paljon todennäköisyydellä 90%, kohtalaisesti todennäköisyydella 5% ja hieman todennäköisyydellä 5%?

2.8 Harjoitustehtävä. Ovatko tehtävän 2.5 Pena ja Jaska symmetrisessä asemassa päätöksenteon suhteen?

2.9 Harjoitustehtävä. Sijoittaja R.:llä on 1.000 euroa löysää rahaa. Hän aikoo sijoittaa ne vuodeksi joko valtion obligaatioiin tai kultaan (mutta ei molempiin). Jos sijoittaja R. sijoittaa 1.000 euroa valtion obligaatioihin saa hän varmasti vuoden kuluttua 1.296 euroa. Jos hän taas sijoittaa kultaan saa hän vuoden kuluttua 400 euroa todennäköisyydellä 0,75 ja 10.000 euroa todennäköisyydellä 0,25. Sijoittaja R.:n hyötyfunktio on $u(r) = \sqrt{r}$.

Sijoittaja R. haluaa optimoida odotetun (eli keskimääräisen) hyödyn. Tuleeko hänen sijoittaa valtion obligaatioihin vai kultaan?

2.10 Harjoitustehtävä. (a) Kutsuisitko tehtävän 2.9 sijoittajaa riskineutraaliksi, riskinrakastajaksi vai riskinhaihtajaksi?

(b) Miten tehtävän 2.9 päätösongelma muuttuu, jos sijoittaja R. voi myös hajauttaa sijoituksensa?

2.11 Harjoitustehtävä. Yhtyneet insinörtit Oy ja Lieto Oy kilpailevat samasta Nikolaigradin kaupungin rakennusprojektista. Lieto uskoo, että Yhtyneiden insinörttien tarjous on satunnaismuuttuja X, jonka jakauma on $\mathrm{P}_X(6.000) = 0{,}40$, $\mathrm{P}_X(8.000) = 0{,}30$ ja $\mathrm{P}_X(11.000) = 0{,}30$. Mikäli tarjoukset menevät tasan, Lieto voittaa tarjouskilpailun, koska Liedon hallituksen puheenjohtaja on myös Nikolaigradin kaupunginhallituksen puheenjohtaja.

Projektin tekeminen maksaa Liedolle 6.000 euroa.

Mikä on Liedon tarjous, kun se on

(a) optimisti,
(b) pessimisti,
(c) katumuksen kaihtaja,
(d) riskineutraali?

2.12 Harjoitustehtävä. Keksi koejärjestely, jolla voit selvittää, onko Pekka Päättäjä riskinkaihtaja, riskinrakastaja, riskineutraali vai “riskinvaihtaja”, eli ei mitään edellisistä.

2.13 Harjoitustehtävä. Haluat selvittää, onko Pekka Päättäja katumuksen kaihtaja, optimisti vai pessimisti. Keksi koejärjestely, jolla saat tämän selville.

Luku 3

Päätöspuut

> Be able to defend your arguments in a rational way. Otherwise, all you have is an opinion. – Marilyn vos Savant

> Opinion is like an asshole. Everybody has one. – Dirty Harry

> The attention span of an average executive is one slide. – Trad

Esitämme tässä luvussa niin sanotun päätöspuumenetelmän päätöksenteon avuksi. Menetelmä on siinä mielessä hyvä, että se on graafinen. Se siis tuottaa kauniita kuvia johtajien, joilla ei ole kärsivällisyyttä laskemiseen tai ymmärtämiseen, päätöksenteon avuksi.

Käsittelemme päätöspuita varsin kepeästi piirtelemällä niitä ja tarkastelemalla esimerkkejä. Todistuksia ja teoriaa on siis hyvin vähän.

3.1 Dynaaminen asetelma

Luvussa 2 tarkastelimme *staattista* päätöksentekoa, eli päätös tehtiin vain kerran, ja sitten katsottiin seuraukset. Nyt tarkastelemme *dynaamista* päätöksentekoa, jossa päätökset ja sattumat seuraavat toisiaan, ja menneisyys vaikuttaa tulevaisuuteen. Päätöksentekoa jatketaan kunnes saavutetaan *suunnitteluhorisontti*.

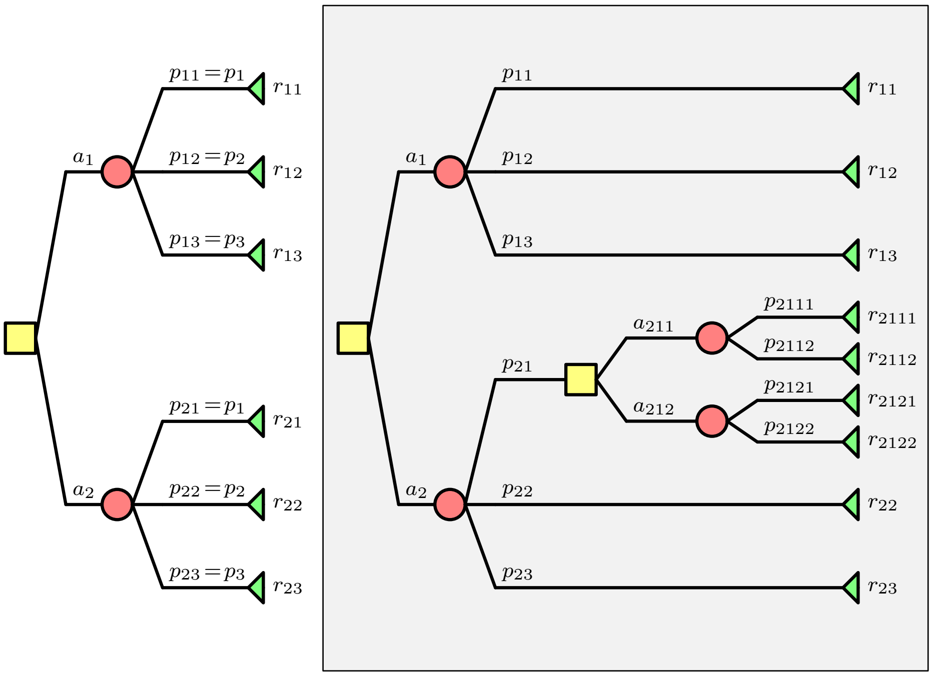
3.1.1 Määritelmä (Päätöspuu)**.** *Päätöspuu* on puu, joka koostuu kolmenlaisista solmuista:

(i) *päätössolmu* (kuvataan usein neliöllä), josta haarautuvista oksista päättäjä voi valita mieleisensä,

(ii) *sattumasolmu* (kuvataan usein ympyrällä), josta haarautuvista oksista sattuma valitsee omansa annettujen todennäkoisyyksien mukaan,

(iii) *lehdistä* (kuvataan usein kylkikolmiolla), eli palkkiosolmuista, joihin puu päättyy, ja joissa saadaan lehden osoittama palkkio.

3.1.2 Huomautus (Dynaaminen vs. staattinen asetelma)**.** Määritelmän 3.1.1 mukainen dynaaminen päätösasetelma tarkoittaa seuraavan kuvan oikean puolen puun kaltaista tilannetta (neliö tarkoittaa päätöstä ja ympyrä tarkoittaa sattumaa; kylkikolmio tarkoittaa tarkastelun loppua). Vasemmanpuoleinen puu kuvastaa staattista tilannetta (joka on tässä toki piirretty myös päätöspuuna).



Vasemmanpuoleinen puu siis kuvastaa edellisessä luvussa 2 käsiteltyä staattista tilannetta, jossa tehdään yksi tehdään yksi päätös, joko a_1 tai a_2, ja sitten katsotaan seuraukset.

Oikeanpuoleinen puu kuvastaa tässä luvussa tarkasteltavaa dynaamista tilannetta. Nytkin tehdään aluksi yksi päätös, joko a_1 tai a_2, mutta tämän päätöksen jälkeen voi seurata uusia päätöksiä uudessa tilanteessa. Kuvan puussa päätöksestä a_1 ei seuraa uusia päätöksiä, mutta päätöksestä a_2 voi seurata uusi päätöstilanne: valinta päätösten a_{211} ja a_{212} välillä, jos kohtalon jumalatar Lady Fortuna niin suo. (Hän suo sen todennäköisyydellä p_{21}, jos valinta a_2 tehtiin.)

3.2 Päätöspuun rakentaminen

Tässä osiossa esitämme, miten päätöspuun voi rakentaa seuraavan esimerkin 3.2.1 "Öljy-yhtiö poraa" avulla. Seuraavassa osiossa 3.3 esitämme miten päätöspuut voi "ratkaista" riskineutraalisti, eli kuinka optimaaliset päätökset lasketaan puista. Lopuksi, osiossa 3.6, esitämme, miten päätöspuut "ratkaistaan" hyötysäännön vallitessa.

3.2.1 Esimerkki (Öljy-yhtiö poraa)**.** Öljy-yhtiö Oy:n pitää päättää poratako öljyä paikasta P. Poraaminen maksaa 100.000€. Jos öljyä löytyy, niin siitä saadaan 6.000.000€. Öljy-yhtiö Oy arvelee, että öljyn löytymisen todennäköisyys on 45%. Ennen varsinaista poraamista Öljy-yhtiö Oy voi palkata geologin arvioimaan paikkaa P. Geologin palkkaaminen maksaa 10.000€. Geologi antaa todennäköisyydellä 50% suotuisan raportin. Mikäli raportti on suotuisa, löytyy öljyä paikasta P todennäköisyydellä 80%. Mikäli raportti on epäsuotuisa, löytyy öljyä paikasta P todennäköisyydellä 10%.

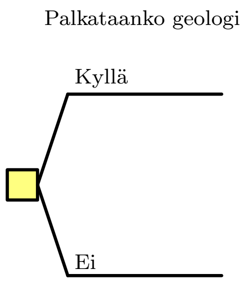
Mitä Öljy-yhtiö Oy:n kannattaa tehdä?

Rakennamme binäärisen puun, jossa peräkkäin kysytään neljä kyllä/ei -kysymystä. Jokaisen kysymyksen kohdalla puu (mahdollisesti) haarautuu. Kysymykset ovat:

1. Palkataanko geologi (päätös)?
2. Onko geologin raportti suotuisa (sattuma)?
3. Porataanko (päätös)?
4. Onko öljyä (sattuma)?

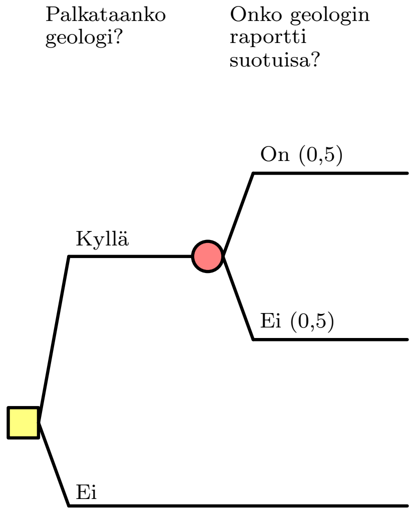
Kysymykset on järjestetty niin, että päätöksentekoa viivytetään mahdollisimman pitkälle, jotta kaikki mahdollinen (satunnainen) informaatio saadaan päätöksenteon tueksi.

Ensimmäiseen kysymykseen "Palkataanko geologi?" joudumme vastaamaan ilman mitään lisäinformaatiota. Tämä on puun *juuri*. Puumme siis alkaa *päätössolmusta* (neliö)

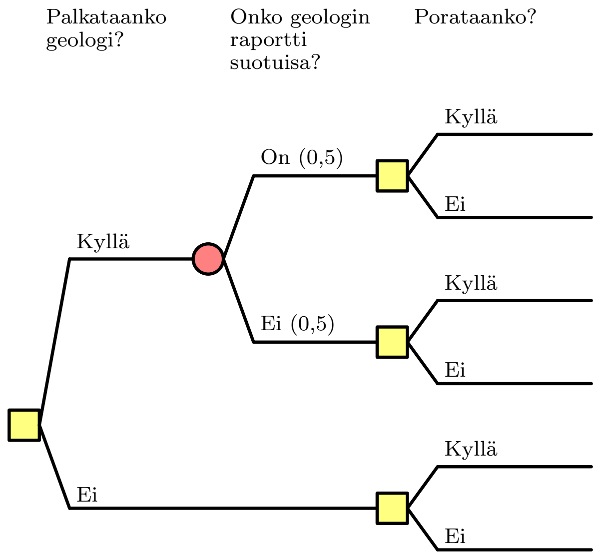


Juuressa puumme siis haarautuu kahdeksi *alipuuksi*: ylemmässä alipuussa geologi palkataan ja alemmassa alipuussa geologia ei palkata.

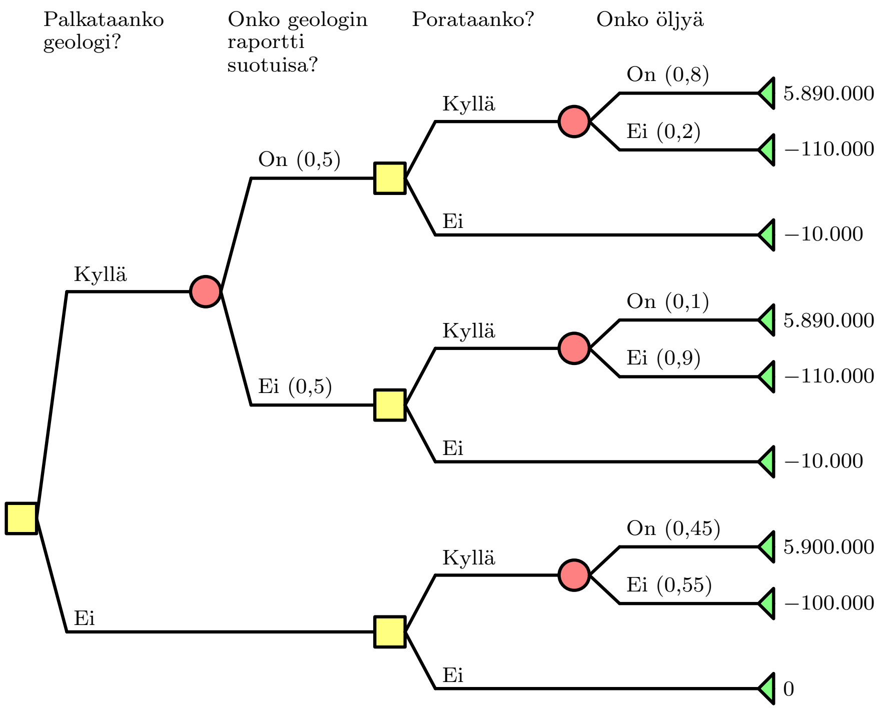
Toinen kysymys on "Onko geologin raportti suotuisa". Jos geologia ei palkattu, on tämä kysymys tietysti merkityksetön, ja siirrymme suoraan kolmanteen kysymykseen. Jos taas geologi palkattiin, niin geologin raportti aiheuttaa *sattumasolmun* (ympyrä) päätöshaaraan, jossa geologi palkattiin. Sattumasolmun haaroihin on merkitty haaran tapahtumisen todennäköisyydet (sulkuihin):



Kolmas kysymys "Porataanko?" koskee kaikkia puun haaroja. Lisäämme siis puuhun kolme päätössolmua:



Viimeinen, neljäs kysymys on "Onko öljyä?". Tämä kysymys koskee vain niitä puun haaroja, joissa porataan. Jokaiseen poraushaaraan tulee siis sattumasolmu. Eri sattumasolmuissa voi olla eri todennäköisyydet. Koska tämä oli viimeinen kysymys, pitää kaikki haarat nyt päättää *lehtiin* (kärkikolmio). Lehden päähän merkitsemme kyseistä puun haaraa vastaavan palkkion. Olemme siis saaneet puun



3.3 Riskineutraali sääntö päätöspuissa

Riskineutraali päätössääntö päätöspuissa perustuu laskemalla sattumasolmuille ja niistä haarautuville alipuille arvot odotusarvoperiaatteen mukaan ja valitsemalla päätössolmuissa se haara, jota vastaava alipuu (eli ko. alipuun juurisolmu) saa suurimman arvon. Tämä arvo tulee myös päätössolmun arvoksi.

Käytännössä arvojen määräminen menee seuraavasti

3.3.1 Määritelmä (Päätöspuiden riskineutraalit arvot)**.**

(i) Lehden arvo on se mikä se on (eli sitä vastaava palkkio).
(ii) Päätössolmun arvo on siitä haarautuvien alipuiden arvojen maksimi, eli parhaimman päätöksen arvo.
(iii) Sattumasolmun arvo on siitä haarautuvien alipuiden odotusarvo, eli todennäköisyyksin painotettu keskiarvo.

3.3.2 Huomautus. Koska lehtien arvot tiedetään ja kaikki päättyy lopulta lehtiin, voidaan kohtien (i)–(iii) avulla laskea kaikkien puiden solmujen arvot käymällä puuhun latvasta (tai pikemminkin latvoista) ja etenemällä takaperin.

3.3.3 Esimerkki (Öljy-yhtiön poraaminen jatkuu)**.** Tarkastelemme edelleen öljy-yhtiöesimerkkiä 3.2.1.

Määräämme puun solmujen arvot lähtien liikkeelle lehdistä. Mutta lehtien arvot on jo annettu. Siten ei tarvitse tehdä mitään. (Tässä kohtaa myöhemmin esitettävä hyötysääntö poikkeaa himema. Muuten hyötysääntö toimii täsmälleen samalla.)

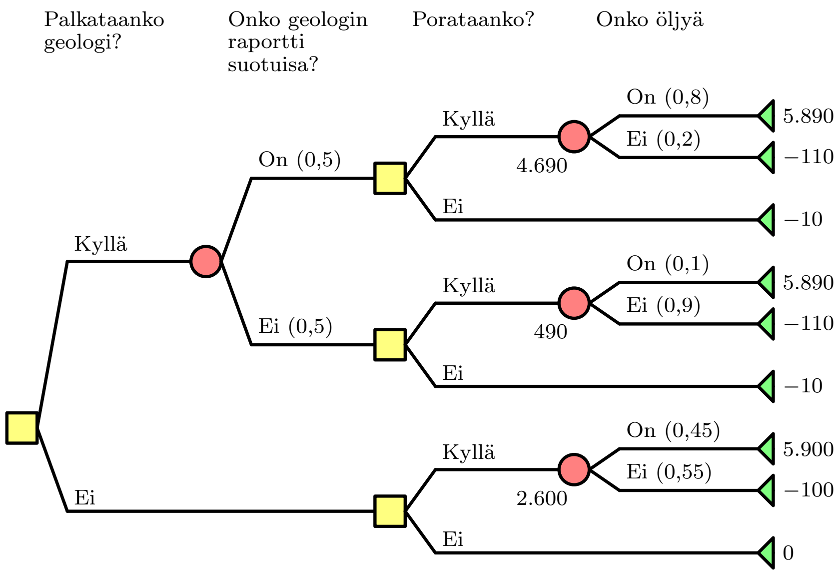
Lehtiä edeltävien sattumasolmujen “Onko öljyä?” arvot lasketaan odotusarvoperiaatteella. Siis, jos esimerkiksi sattuma haarautuu kahdeksi lehdeksi todennäköisyyksin p_1 ja p_2, ja lehtien arvot ovat r_1 ja r_2, niin kyseisen sattumasolmun arvo on odotusarvo

$$r_1 p_1 + r_2 p_2.$$

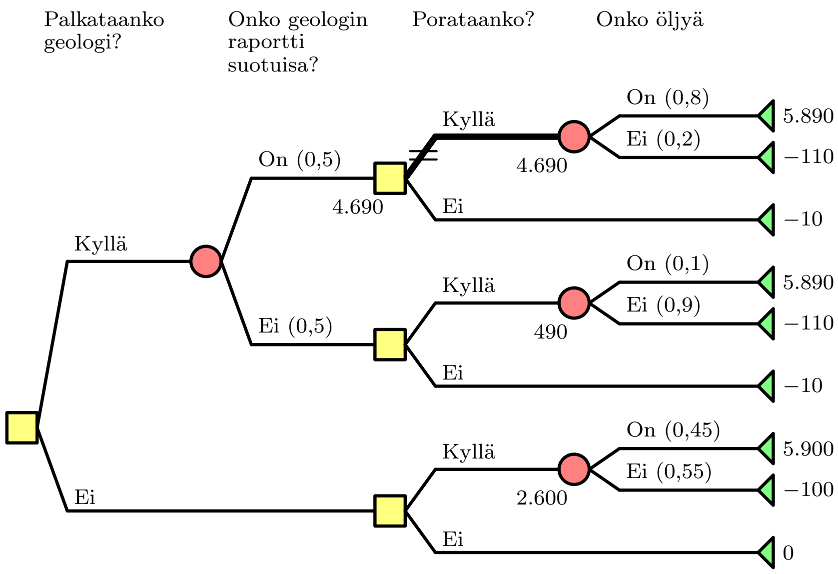
Öljy-yhtiön tapauksessa saamme siis solmujen arvoksi (solmut on lueteltu ylhäältä alas, ja olemme siirtyneet euroista kiloeuroihin)

$$\begin{aligned} 5.890 \cdot 0{,}80 + (-110) \cdot 0{,}20 &= 4.690, \\ 5.890 \cdot 0{,}10 + (-110) \cdot 0{,}90 &= 490, \\ 5.900 \cdot 0{,}45 + (-100) \cdot 0{,}55 &= 2.600. \end{aligned}$$

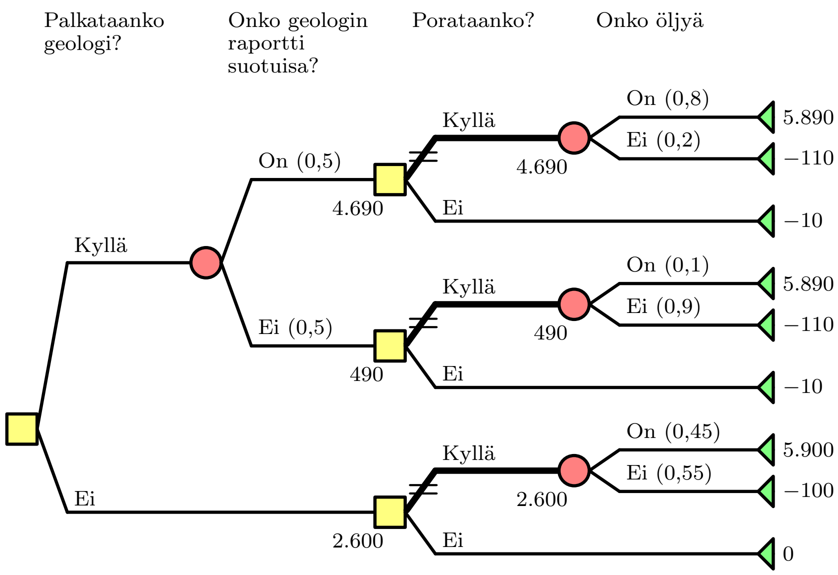
Päätöspuussa merkitsemme solmun arvon sen vasempaan alanurkkaan. Täydennämme päätöspuun siis muotoon



Seuraavaksi astumme puussa askeleen taaksepäin ja päädymme päätössolmuihin “Porataanko?”. Ylimmässä päätössolmussa alipuun “Kyllä” arvo laskettiin juuri. Se oli 4.690. Alipuun “Ei” arvo on lehtiarvo −10. Siis paras päätös tässä päätössolmussa on “Kyllä” ja sitä vastaava arvo 4.690 on tämän päätössolmun arvo. Merkitsemme päätössolmun arvon sen vasempaan alanurkkaan. Merkitsemme myös parhaimman päätöksen symbolilla $=$ ja piirrämme vielä selvyyden vuoksi parhaimman haaran leveällä tussilla. Olemme siis tässä vaiheessa saaneet puun



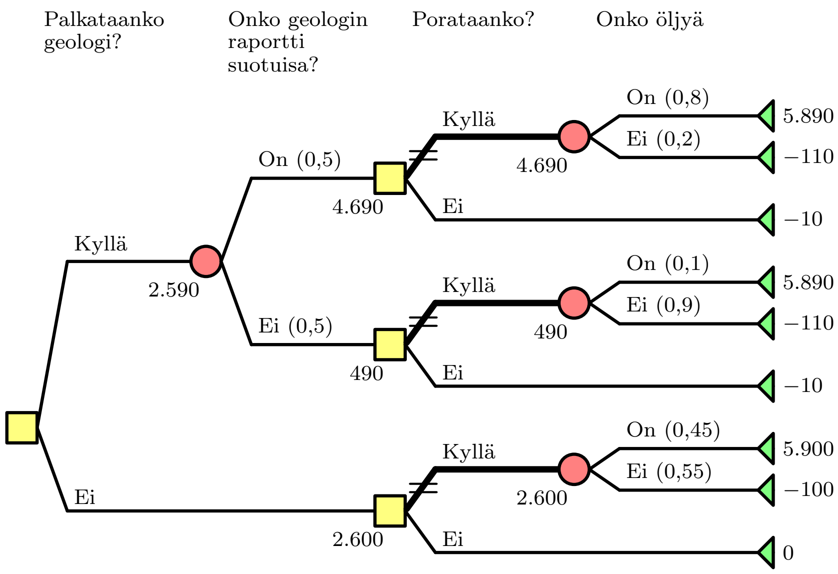
Täyttämällä muut päätössolmut “Porataanko” samalla tavalla saamme puun tämän tason täytetyksi:



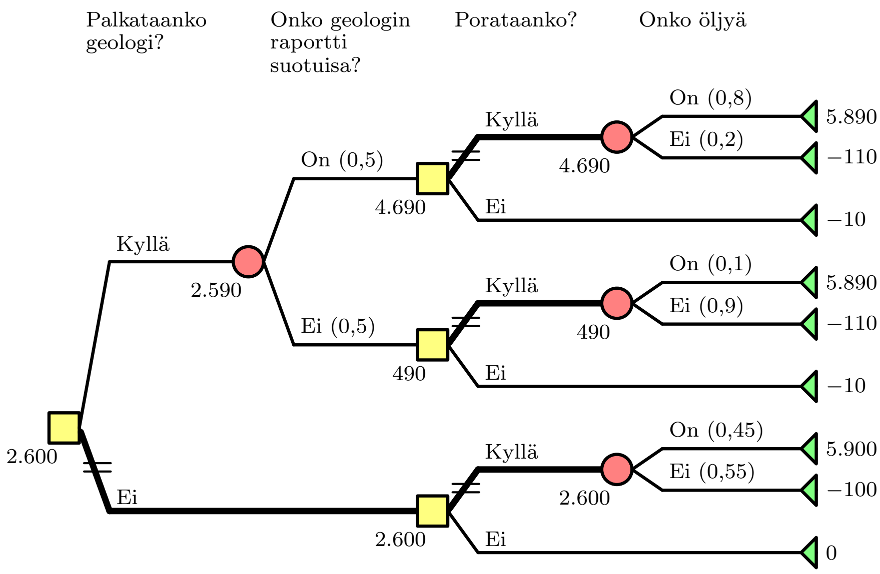
Astumme taas puussa taaksepäin. Päädymme sattumasolmuun “Onko geologin raportti suotuisa?”. Tämän sattumasolmun arvo on sen alipuiden arvojen todennäköisyyksin painotettu keskiarvo, eli odotusarvo. Alipuiden arvo on jo laskettu edellisessä vaiheessa. Saamme siis arvon

$$4.690 \cdot 0{,}50 + 490 \cdot 0{,}50 \quad = \quad 2.590.$$

Täytämme tämän päätöspuuhun ja saamme



Lopulta laskeudumme juureen, eli solmuun "Palkataanko geologi?". Näemme että alipuun "Kyllä" arvo on 2.590 ja alipuun "Ei" arvo on 2.600. Siten päätössolmun "Palkataanko geologi" arvo on 2.600 ja paras päätös tässä kohtaa on "Ei". Olemme siis saaneet lopulta täysin täytetyn päätöspuun:



Johtopäätös on, että ei kannata palkata geologia, ja että poraaminen kannattaa aina. Lisäksi tämän päätöstilanteen, eli päätöspuun, arvo on (öljy-yhtiölle) 2.600 euroa.

3.4 Bayesin kaava päätöspuissa

Usein päätösongelman kuvauksessa on ehdolliset todennäköisyydet annettu "väärin päin". Toisin sanoen on annettu todennäköisyys $\mathrm{P}(A|B)$, kun tarvitsemme puussa todennäköisyyden $\mathrm{P}(B|A)$. Tarvittavat todennäköisyydet saadaan usein käännettyä Bayesin kaavalla 1.2.17 tai 1.2.20.

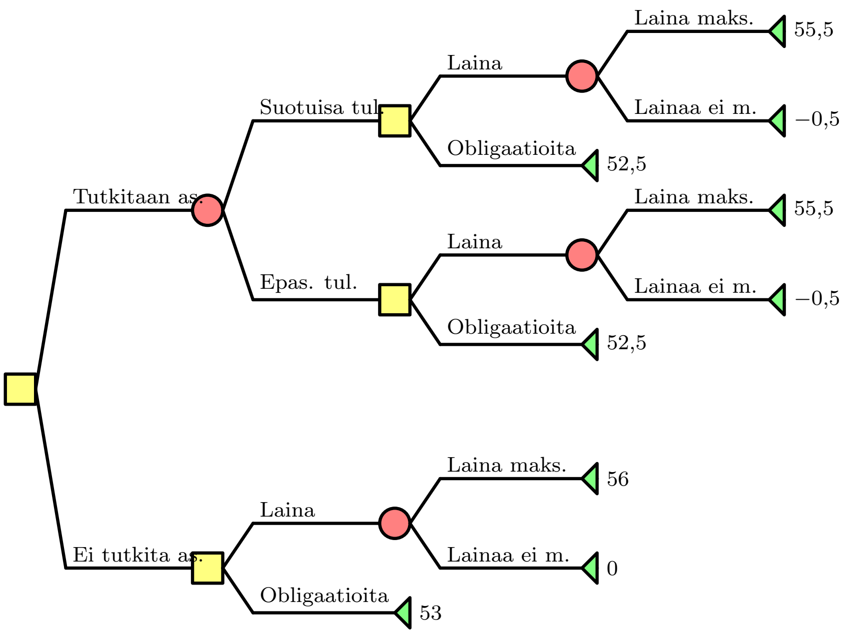
Tarkastelemme tätä todennäköisyyksien kääntämisen ongelmaa seuraavan esimerkin 3.4.1 kautta:

3.4.1 Esimerkki (Pankki lainaa, mutta kenelle)**.** Pankki harkitsee myöntääkö 50.000 euron lainan asiakkaalle 12% korolla, vai sijoittaako kyseiset 50.000 euroa valtion obligaatioihin 6% korolla. Pankki arvioi, että 4%:n todennäköisyydellä asiakas ei pysty maksamaan lainaansa takaisin. Pankki voi teettää tutkimuksen asiakkaan luotettavuudesta 500 eurolla. Todennäköisyys, että tutkimus antaa suotuisan tuloksen asiakkaalle, joka pystyy maksamaan lainansa on $77/96$. Todennäköisyys, että tutkimus antaa suotuisan tuloksen asiakkaalle, joka ei pysty maksamaan lainaansa on $1/4$.

Pankin kannalta tilanteessa on kaksi päätösongelmaa:

- Kannattaako pankin teettää tutkimus asiakkaasta?
- Kannattaako pankin antaa laina?

Pankin päätöstilannetta vastaava puu on (ilman todennäköisyyksiä, palkkiot kiloeuroissa):

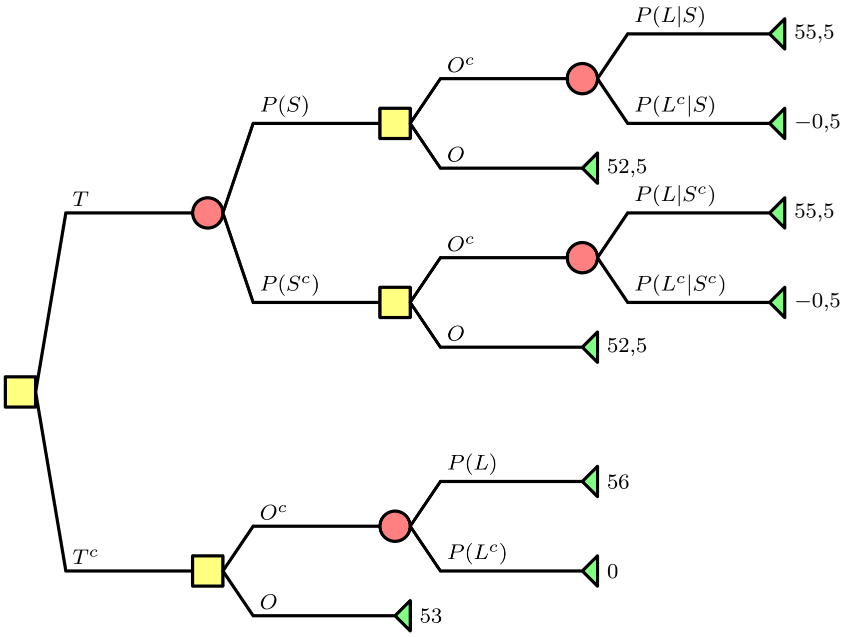


(Tää päätöspuu ei ole "synkronoitu" toisin kuin esimerkin 3.2.1 Öljy-yhtiön päätöspuu.)

Tutkimme mitkä todennäköisyydet tarvitsemme. Merkitsemme

$$\begin{aligned} L &= \text{“Laina maksetaan”}, \\ L^c &= \text{“Lainaa ei makseta”}, \\ S &= \text{“Raportti on suotuisa”}, \\ S^c &= \text{“Raportti ei ole suotuisa”}, \\ T &= \text{“Asiakas tutkitaan”}, \\ O &= \text{“Ostetaan obligaatioita”}, \\ O^c &= \text{“Annetaan laina”}. \end{aligned}$$

Näillä merkinnöillä abstraktein todennäköisyyksin täytetty päätöspuu on



Prioritodennäkösyydet L:lle on annettu:

$$\mathrm{P}(L^c) = 0{,}04 \quad \text{ja} \quad \mathrm{P}(L) = 0{,}96.$$

Lisäksi on annettu uskottavuudet

$$\mathrm{P}(S|L) = 77/96 \quad \text{ja} \quad \mathrm{P}(S^c|L) = 19/96$$

sekä

$$\mathrm{P}(S|L^c) = 1/4 \quad \text{ja} \quad \mathrm{P}(S^c|L^c) = 3/4.$$

Sen sijaan tarvittavia posterioritodennäköisyyksiä $\mathrm{P}(L|S)$, $\mathrm{P}(L^c|S)$, $\mathrm{P}(L|S^c)$ ja $\mathrm{P}(L^c|S^c)$ ei ole annettu. Voimme kuitenkin laskea ne annetuista todennäköisyyksistä Bayesin kaavalla 1.2.20.

Ensiksi posterioritodennäköisyydet $\mathrm{P}(L|S)$ ja $\mathrm{P}(L^c|S)$:

$$\begin{aligned} \mathrm{P}(L|S) &= \frac{\mathrm{P}(L)\mathrm{P}(S|L)}{\mathrm{P}(L)\mathrm{P}(S|L) + \mathrm{P}(L^c)\mathrm{P}(S|L^c)} \\ &= \frac{0{,}96 \cdot 77/96}{0{,}96 \cdot 77/96 + 0{,}04 \cdot 1/4} \\ &= 0{,}987. \end{aligned}$$

Koska ehdollinen todennäköisyys on (ehdon ollessa kiinnitetty) todennäköisyys, saamme käyttämällä komplementtikaavaa 1.2.3(i) edelliseen, että

$$\mathrm{P}(L^c|S) = 1 - \mathrm{P}(L|S) = 0{,}013.$$

Samalla tavalla Bayesin kaavalla 1.2.20 voimme laskea komplementaariset posterioritodennäköisyydet $\mathrm{P}(L|S^c)$ ja $\mathrm{P}(L^c|S^c)$:

$$\begin{aligned} \mathrm{P}(L|S^c) &= \frac{\mathrm{P}(L)\mathrm{P}(S^c|L)}{\mathrm{P}(L)\mathrm{P}(S^c|L) + \mathrm{P}(L^c)\mathrm{P}(S^c|L^c)} \\ &= \frac{0{,}96 \cdot 19/96}{0{,}96 \cdot 19/96 + 0{,}04 \cdot 3/4} \\ &= 0{,}864, \end{aligned}$$

ja soveltamalla komplementtikaavaa 1.2.3(i) edelliseen saamme

$$\mathrm{P}(L^c|S^c) = 1 - \mathrm{P}(L|S^c) = 0{,}136.$$

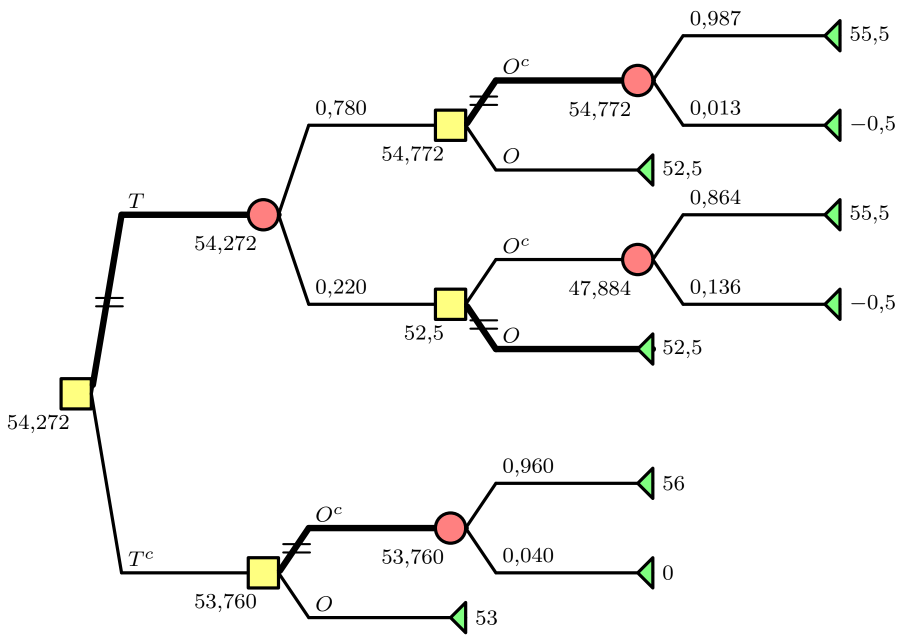
Lopulta huomaamme, että olemme jo itse asiassa laskeneet todennäköisyyden

$$\mathrm{P}(S) = \mathrm{P}(L)\mathrm{P}(S|L) + \mathrm{P}(L^c)\mathrm{P}(S|L^c) = 0{,}780.$$

Siten, komplementtikaavan 1.2.3(i) nojalla

$$\mathrm{P}(S^c) = 1 - \mathrm{P}(S) = 0{,}220.$$

Sijoittamalla saadut luvut aiemmin rakennettuun päätöspuuhun ja laskemalla odotusarvot normaaliin tapaan lähtien liikkeelle lehdistä saamme



Näemme, että pankin kannattaa teettää tutkimus asiakkaasta, ja jos tulos on suotuisa, antaa laina. Jos taas tulos on epäsuotuisa, ei lainaa kannata antaa.

3.5 Informaation arvo

Tässä osiossa tarkestelemme kysymystä, kuinka paljon kannattaa maksaa informaatiosta? Tai pikemminkin, kuinka paljon kannattaa maksaa epävarmuuden pienentymisestä.

Epätäydellisen informaation arvo

3.5.1 Määritelmä (Asiantuntijainformaatio)**.** *Asiantuntijainformaatio* eli *epätäydellinen informaatio* tarkoittaa informaatiota, joka muuttaa näkemystämme epävarmojen tapahtumien todennäköisyyksistä.

3.5.2 Huomautus (Informaatio ja entropia)**.** *Entropia* on satunnaisuuteen, eli epävarmuuteen, liittyvä satunnaisuuden mitta: satunnaismuuttujan jakauman p_j, $j \in J$, entropia on luku[1]

$$H(p) \quad = \quad -\sum_{j\in J} p_j \ln p_j.$$

[1] Tunnetusti $0 \cdot \infty = 0$.

Jos siis esimerkiksi tarkastelemme binääristä tilannetta, eli klaavan tapahtumista painotetun kolikon heitossa, on entropia

$$H(p) \quad = \quad -p \ln p - (1-p)\ln(1-p).$$

Kaikista entrooppisin tilanne on silloin, kun klaavan todennäköisyys $p = 1/2$. Tällöin entropia on $0{,}69315$. Vähiten entrooppinen tilanne on silloin kun klaavan todennäköisyys $p = 0$ tai $p = 1$.

Jos asiantuntija oikeasti tuntee asiansa on perusteltua olettaa, että entropia pienenee. Esimerkiksi Öljy-yhtiön tapauksessa 3.2.1 näkemys öljyn löytymisestä ilman asiantuntijan raporttia on $0{,}45$. Tämä tarkoittaa entropiaa

$$-0{,}45 \ln 0{,}45 - 0{,}55 \ln 0{,}55 \quad = \quad 0{,}68814.$$

Asiantuntijan entropiat taas ovat: (1) suotuisan raportin tapauksessa

$$-0{,}8 \ln 0{,}8 - 0{,}2 \ln 0{,}2 \quad = \quad 0{,}50040$$

ja (2) epäsuotuisan raportin tapauksessa

$$-0{,}1 \ln 0{,}1 - 0{,}9 \ln 0{,}9 \quad = \quad 0{,}32508.$$

Asiantuntija näyttää siis tuntevan asian. (Suotuisan raportin todennäköisyys 50% ei liity asiantuntijan ammattitaitoon, vaan siihen, miten tosiasiat ovat.)

3.5.3 Määritelmä (Asiantuntijainformaation riskineutraali arvo)**.** Olkoon päätöspuun DT (juuren) arvo $v_0 = v(\mathrm{DT})$. Asiantuntija antaa raportit $e_1, e_2, \ldots, e_n$ todennäköisyyksin $q_1, q_2, \ldots, q_n$. Jokaista raporttia e_k, $k = 1, 2, \ldots, n$, vastaa päätöspuu DT_{e_k}, joka saadaan päätöspuusta DT korvaamalla kaikki esiintyvät “raa’at” todennäköisyydet raporttia e_k vastaavilla asiantuntijatodennäköisyyksillä. Päätöspuulle DT_{e_k} voidaan laskea arvo $v_{e_k} = v(\mathrm{DT}_{e_k})$ normaaliin tapaan. Koska asiantuntija antaa puun DT_{e_k} todennäköisyydellä q_k, niin hän antaa itse asiassa puun DT_e, missä on juuressa sattumasolmu, joka haarautuu puihin $\mathrm{DT}_{e_1}, \mathrm{DT}_{e_2}, \ldots, \mathrm{DT}_{e_n}$ todennäköisyyksin $q_1, q_2, \ldots, q_n$. Siten asiantuntijan antaman puun riskineutraali arvo on

$$v_e \quad = \quad v(\mathrm{DT}_e) \quad = \quad \sum_{k=1}^{n} q_k \, v_{e_k}.$$

Nyt siis päätöksentekijällä on kaksi vaihtoehtoa: (1) joko pelata alkuperäisellä puullaan, jonka arvo on v_0 tai (2) maksaa x ja pelata asiantuntijapuulla, jonka arvo on v_e. *Asiantuntijainformaation riskineutraali arvo* on suurin x, jolle $v_e - x \geq v_0$, eli

$$x \quad = \quad v_e - v_0.$$

3.5.4 Esimerkki (Sijoittaja K. sijoittaa)**.** Sijoittaja K.:lla on 10.000 €. Hän voi sijoittaa sen joko Riskiin, Varmaan, tai Säästöön. Jos hän sijoittaa Riskiin tai Varmaan, joutuu hän maksamaan 200 € erinäisiä käsittelykuluja. Jos tulee karhumarkkinat Riski antaa −800 € voittoa ja Varma 100 € voittoa. Jos tulee sonnimarkkinat Riski antaa 1.700 € voittoa ja Varma antaa 1.200 € voittoa. Jos nykymeno jatkuu, niin Riski antaa 300 € voittoa ja Varma antaa 400 € voittoa. Säästö antaa aina 500 € voittoa. Sonnimarkkinoiden todennäköisyys on 0,50, nykymenon todennäköisyys on 0,30 ja karhumarkkinoiden todennäköisyys on 0,20.

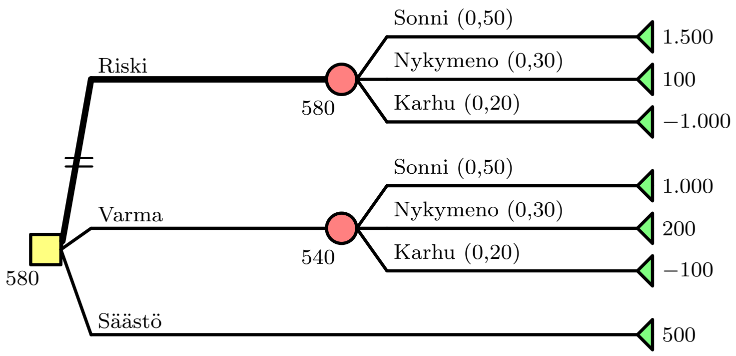
(a) Mihin kohteeseen sijoittaja K.:n tulee sijoittaa?

(b) Sijoittaja K. voi konsultoida markkina-asiantuntija R.:ää, jonka arviot ovat osoittautuneet vastaamaan todellisuutta seuraavan kaavion tavoin:

	Todellisuus		
Ennustus	Sonni	Nykymeno	Karhu
Sonni	0,80	0,15	0,20
Nykymeno	0,10	0,70	0,20
Karhu	0,10	0,15	0,60

Toisin sanoen esimerkiksi 15% tilanteista, joissa nykymeno on jatkunut markkina-asiantuntija onkin ennustanut sonnimarkkinoita. (Huomaa, että tässä taulokossa sarakkeet summautuvat ykkösiksi, eivät rivit. Tosin sanoen sarakkeet vastaavat todennäköisyysjakaumia, mutta rivit eivät.)
Kuinka paljon sijoittaja K.:n kannattaa korkeintaan maksaa markkina-asiantuntija R.:n ennusteesta?

Kohdassa (a) sijoittaja K.:n tilannetta kuvaava riskineutraalisti täytetty päätöspuu voittoineen ja tappioineen on:



Sijoittaja K.:n kannattaa siis sijoittaa Riskiin.

Rakennamme seuraavaksi kohtaan (b) asiantuntijapuun. Puuta varten tarvitsemme asiantuntijainformaation posterioritodennäköisyydet ja todennäköisyydet asiantuntija R.:n raporteille. Nämä pitää laskea, sillä esimerkin 3.5.4 taulukossa on annettu ehdolliset todennäköisyydet puun kannalta “väärin päin”.

Tarvitsemme siis Bayesin kaavaa ja kokonaistodennäköisyyden kaavaa taulukon "kääntämiseen".

Tonennäköisyydet asiantuntija R.:n raporteille saadaan kokonaistodennäköisyyden kaavalla 1.2.19:

$$\begin{aligned}
\mathrm{P}(\text{R.: ``Sonni''}) &= \mathrm{P}(\text{Sonni})\mathrm{P}(\text{R.: ``Sonni''}|\text{Sonni}) + \\
&\quad \mathrm{P}(\text{Nykymeno})\mathrm{P}(\text{R.: ``Sonni''}|\text{Nykymeno}) + \\
&\quad \mathrm{P}(\text{Karhu})\mathrm{P}(\text{R.: ``Sonni''}|\text{Karhu}) \\
&= 0{,}50 \cdot 0{,}80 + 0{,}30 \cdot 0{,}15 + 0{,}20 \cdot 0{,}20 \\
&= 0{,}485.
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(\text{R.: ``Nykymeno''}) &= \mathrm{P}(\text{Sonni})\mathrm{P}(\text{R.: ``Nykymeno''}|\text{Sonni}) + \\
&\quad \mathrm{P}(\text{Nykymeno})\mathrm{P}(\text{R.: ``Nykymeno''}|\text{Nykymeno}) + \\
&\quad \mathrm{P}(\text{Karhu})\mathrm{P}(\text{R.: ``Nykymeno''}|\text{Karhu}) \\
&= 0{,}50 \cdot 0{,}10 + 0{,}30 \cdot 0{,}70 + 0{,}20 \cdot 0{,}20 \\
&= 0{,}300.
\end{aligned}$$

Lopuksi voimme laskea

$$\begin{aligned}
\mathrm{P}(\text{R.: ``Karhu''}) &= 1 - \Big(\mathrm{P}(\text{R.: ``Sonni''}) + \mathrm{P}(\text{R.: ``Nykymeno''})\Big) \\
&= 1 - (0{,}485 + 0{,}300) \\
&= 0{,}215.
\end{aligned}$$

Posterioritodennäköisyydet saadaan nyt Bayesin kaavasta 1.2.17:

	Posteriori		
Ennustus	Sonni	Nykymeno	Karhu
Sonni	0,8247	0,0928	0,0825
Nykymeno	0,1667	0,7000	0,1333
Karhu	0,2325	0,2093	0,5581

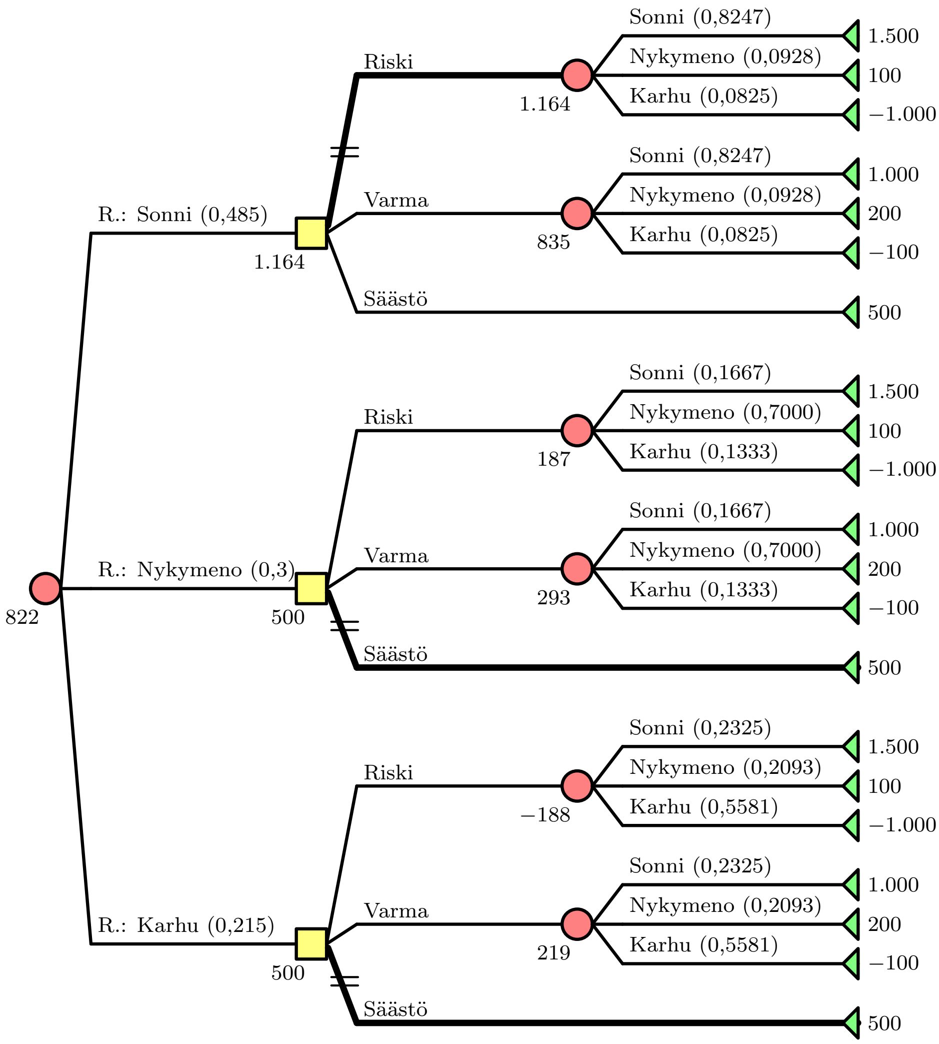
(Huomaa, että tässä taulokossa rivit summautuvat ykkösiksi, eivät sarakkeet. Tosin sanoen rivit vastaavat todennäköisyysjakaumia, mutta sarakkeet eivät. Tämä on siis "käännetty" versio aikaisemmasta asiantuntijataulukosta.)

Yllä olevassa taulukossa esimerkiksi posterioritodennäköisyydet 0,8247 ja 0,1667 on saatu laskemalla kaavalla 1.2.17

$$\begin{aligned}
&\mathrm{P}(\text{Sonni}|\text{R.: ``Sonni''}) \\
&= \frac{\mathrm{P}(\text{Sonni})\mathrm{P}(\text{R.: ``Sonni''}|\text{Sonni})}{\mathrm{P}(\text{R: ``Sonni''})} \\
&= \frac{0{,}50 \cdot 0{,}80}{0{,}485} \\
&= 0{,}8247.
\end{aligned}$$

$$
\begin{aligned}
&\mathrm{P}(\text{Sonni}|\text{R.: "Nykymeno"}) \\
&= \frac{\mathrm{P}(\text{Sonni})\mathrm{P}(\text{R.: "Nykymeno"}|\text{Sonni})}{\mathrm{P}(\text{R: "Nykymeno"})} \\
&= \frac{0{,}50 \cdot 0{,}10}{0{,}300} \\
&= 0{,}1667.
\end{aligned}
$$

Saamme asiantuntijapuun voittoineen ja tappioineen



Johtopäätös on, että markkina-asiantuntijan informaation arvo on sijoittaja K.:lle $822 - 580 = 242$ euroa.

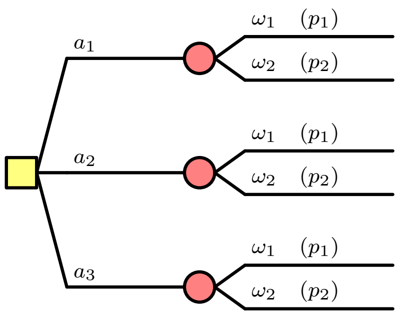
Oraakkeli-informaation arvo

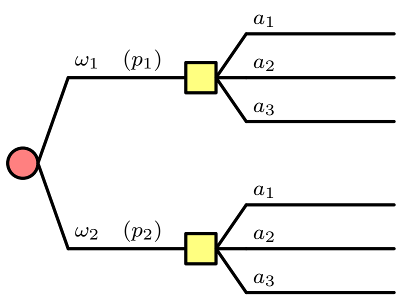
Oraakkelilla emme tarkoita Delfoin oraakkelia, joka antaa hyödytöntä informaatiota tyyliin "suuri valtakunta tuhoutuu, jos aloitat sodan", vaan täydellisen luotettavaa asiantuntijaa:

3.5.5 Määritelmä (Oraakkeli-informaatio)**.** *Oraakkeli* on asiantuntija, jonka raporteissa esiintyy ainoastaan todennäköisyyksiä 0 ja 1.

Oraakkeli-informaation arvo voidaan siis laskea samalla tavalla kuin asiantuntijainformaation arvo.

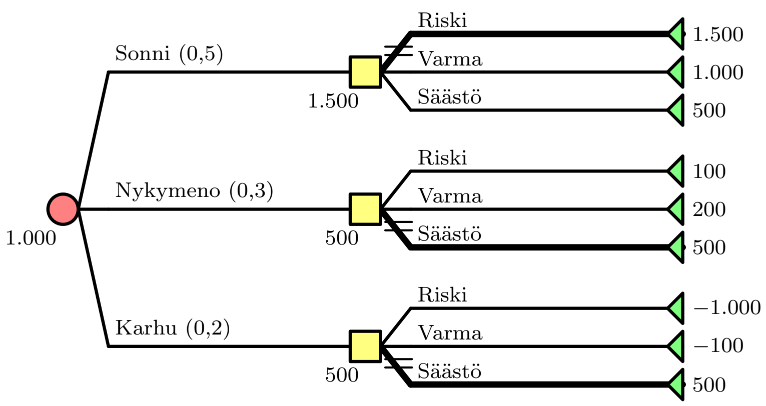
3.5.6 Huomautus (Oraakkeli-informaation arvo puita kääntämällä)**.** Mikäli oraakkeli antaa raporttinsa tunnettujen "raakojen" todennäköisyyksien mukaan (kuten on luonnollista olettaa) voidaan oraakkelipuun arvo laskea kääntämällä alkuperäisestä puusta kaikki sattuma ennen päätöksiä seuraavaan tyyliin (vasemmalla kääntämätön kohta, oikealla käännetty kohta):





Kääntämistä jatketaan, kunnes kaikki sattuma on ennen päätöksiä. Tällöin saadaan oraakkelipuu, joka vastaa päätöstilannetta "kysytään oraakkelilta".

3.5.7 Esimerkki (Sijoittaja K. konsultoi oraakkelia)**.** Esimerkissä 3.5.4 oraakkelipuu on



Johtopäätös on, että sijoittaja K.:n kannattaa maksaa oraakkelille korkeintaan $1000 - 580 = 420$ euroa.

3.6 Hyötysääntö päätöspuissa

Hyötyyn perustuva päätössääntö on helppo sovittaa päätöspuihin. Menetelmä on täsmälleen sama kuin riskineutraalissa tapauksessa paitsi, että (lehdissä olevat) palkkiot pitää aluksi muuttaa hyödyiksi.

Esitämme formaalin määritelmän kertauksen vuoksi:

3.6.1 Määritelmä (Päätöspuiden hyötyarvot)**.**

(i) Lehden arvo on sitä vastaavan palkkion hyötyarvo.

(ii) Päätössolmun arvo on siitä haarautuvien alipuiden arvojen maksimi, eli parhaimman päätöksen arvo.

(iii) Sattumasolmun arvo on siitä haarautuvien alipuiden odotusarvo, eli todennäköisyyksin painotettu keskiarvo.

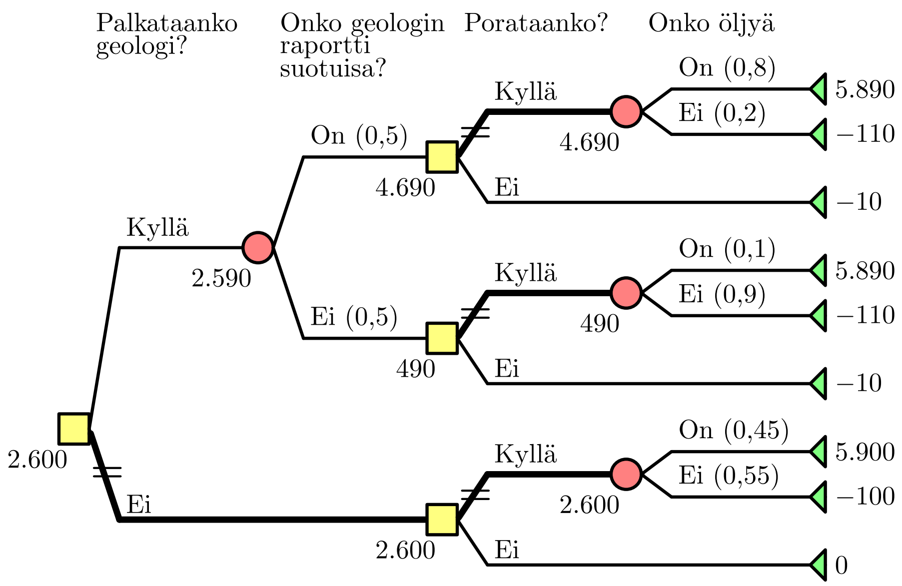
Palaamme vanhaan öljy-yhtiöesimerkkiin 3.2.1, mutta lisäämme hyödyt.

3.6.2 Esimerkki (Öljy-yhtiön hyöty)**.** Öljy-yhtiö Oy:n pitää päättää poratako öljyä paikasta P. Poraaminen maksaa 100.000 €. Jos öljyä löytyy, niin siitä saadaan 6.000.000 €. Öljy-yhtiö Oy arvelee, että öljyn löytymisen todennäköisyys on 45%. Ennen varsinaista poraamista Öljy-yhtiö Oy voi palkata geologin arvioimaan paikkaa P. Geologin palkkaaminen maksaa 10.000 €. Geologi antaa todennäköisyydellä 50% suotuisan raportin. Mikäli raportti on suotuisa, löytyy öljyä paikasta P todennäköisyydellä 80%. Mikäli raportti on epäsuotuisa, löytyy öljyä paikasta P todennäköisyydellä 10%.

Öljy-yhtiön tulos ilman porausta on 10.000.000 € ja Öljy-yhtiön hyötyfunktio on $u(r) = \sqrt{r/1.000}$.

Mitä Öljy-yhtiö Oy:n kannattaa tehdä?

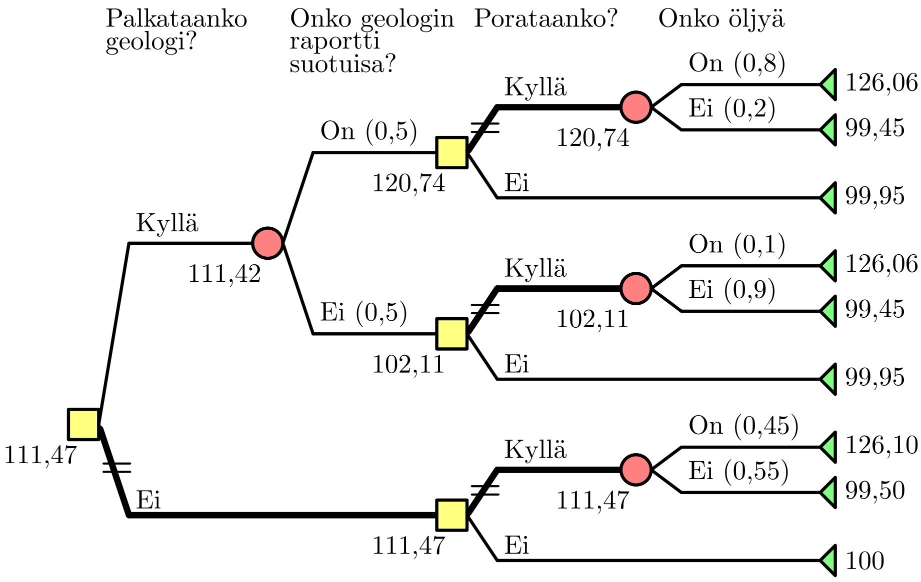
Aluksi palautamme mieliin, mikä tilanne oli riskineutraalissa päätöksenteossa:



Laskemme sitten tarvittavat hyödyt

$$
\begin{aligned}
u(10.000.000 + 5.900.000) &= \sqrt{15.900} &= 126{,}10, \\
u(10.000.000 + 5.890.000) &= \sqrt{15.890} &= 126{,}06, \\
u(10.000.000 + 0) &= \sqrt{10.000} &= 100{,}00, \\
u(10.000.000 - 10.000) &= \sqrt{9.990} &= 99{,}95, \\
u(10.000.000 - 100.000) &= \sqrt{9.900} &= 99{,}50, \\
u(10.000.000 - 110.000) &= \sqrt{9.890} &= 99{,}45,
\end{aligned}
$$

Sijoittamalla hyödyt puun lehtiin ja laskemalla solmujen arvot normaaliin tapaan takaperin saamme lopulta



Johtopäätös on, että geologia ei kannata palkita, ja poraaminen kannattaa aina. Tämä on muuten, sattumoisin, sama tulos kuin riskineutraalissakin tapauksessa.

3.6.3 Huomautus (Informaation hyötyarvo)**.** Seuraava luonnollinen kysymys on, kuinka paljon geologille kannattaisi korkeintaan maksaa. Tähän kysymykseen ei voi valitettavasti suoraan vastata edellä olevan päätöspuun avulla, sillä hyödyt toimivat epälineaarisesti palkkioihin nähden. Se on koko hyötykäsitteen idea! Jos haluamme selvittää, kuinka suuri asiantuntijainformaation arvo on päätöspuussa joudumme käytännössä tyypillisesti kokeilemalla haarukoimaan arvon. Toisin sanoen kasvatamme asiantuntijan kustannusta nollasta ylöspäin kunnes saavutamme pisteen, jossa asiantuntijapuun hyötyarvo on sama kuin alkuperäisen puun hyötyarvo. Sama ongelma koskee tietysti myös oraakkeli-informaation arvoa.

3.7 Päätöspuut vs. päätösmatriisit

Seuraava lause sanoo, että itse asiassa päätösmatriisit ja päätöspuut ovat matemaattiselta kannalta olennaisesti samoja asioita. Valinta näiden menetelmien välillä on siis puhtaasti mukavuuskysymys.

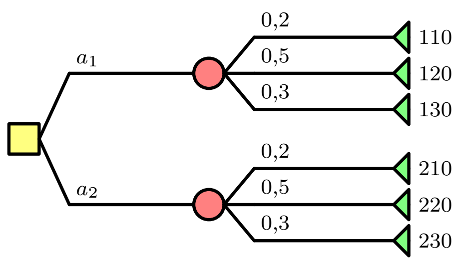
3.7.1 Lause (Matriisi–puu -ekvivalenssi)**.** *Riskineutraalin ja hyötysäännön vallitessa jokaista päätöspuuta vastaa päätösmatriisi ja päin vastoin*

Lausen 3.7.1 todistus on harjoitustehtävä 3.10. Havainnolistamme sen väittämää ekvivalenssia seuraavalla esimerkillä:

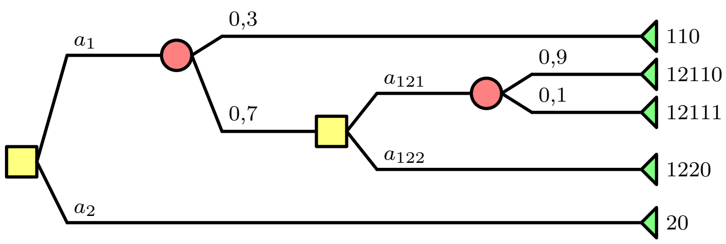
3.7.2 Esimerkki. (i) Päätöspuu päätösmatriisista: Olkoon

$$R = \begin{bmatrix} 110 & 120 & 130 \\ 210 & 220 & 230 \end{bmatrix}$$

ja $p = (0{,}2;\ 0{,}5;\ 0{,}3)$. Tätä (staattista) päätöstilannetta vastaa (ei niin kauhean dynaaminen) päätöspuu



(ii) Päätösmatriisi päätöspuusta: Olkoon päätöspuu



Yllä olevaa päätöspuuta vastaa (esimerkiksi) päätösmatriisi

$$R = \begin{bmatrix} 110 & 12110 & 12111 \\ 110 & 1220 & 1220 \\ 20 & 20 & 20 \end{bmatrix}$$

todennäköisyyksin $p = (0{,}3;\ 0{,}63 = 0{,}7 \cdot 0{,}9;\ 0{,}07 = 0{,}7 \cdot 0{,}1)$. Matriisissa R ensimmäinen päätös (eli rivi) vastaa puun päätöstä $a_1 \wedge a_{121}$, toinen päätös (eli rivi) vastaa puun päätöstä $a_1 \wedge a_{122}$, ja kolmas päätös (eli rivi) vastaa puun päätöstä a_2.

Lauseen 3.7.1 todistamisen (mahdolliseksi) helpottamiseksi esitämme seuraavan määritelmän 3.7.3 ja siihen liittyvä apulauseen 3.7.4.

3.7.3 Määritelmä (Binäärinen vaihteleva päätöspuu)**.** Päätöspuu on *binäärinen*, jos jokainen solmu, joka ei ole lehti, haarautuu täsmälleen kahteen haaraan.

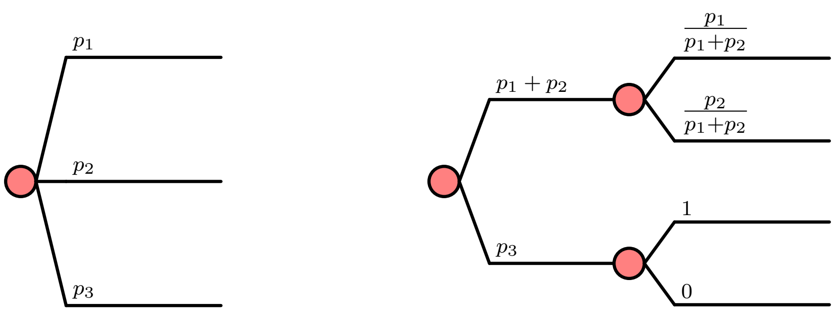
Puu on *vaihteleva*, jos jokaista päätössolmua seuraa sattumasolmu (tai lehti) ja jokaista sattumasolmua seuraa päätössolmu (tai lehti).

3.7.4 Apulause. *Jokainen päätöspuu on ekvivalentti jonkin binäärisen vaihtelevan päätöspuun kanssa.*

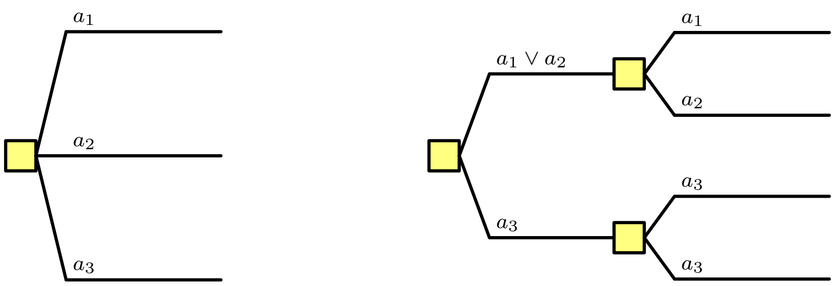
Todistus. Tämä todistus on harjoitustehtävä 3.8(c).

Vihjeenä todistuksen binääriosuuteen mainittakoon seuraavat päätöspuiden haarautumisten väliset ekvivalenssit (esitetyt päätöspuiden haarat ovat riveittäin ekvivalentteja).

Sattumasolmuille pätee ekvivalenssi



Päätössolmuille pätee ekvivalenssi



Edellä oikealla olevassa puussa siis toistetaan päätöstä a_3 vastaava alipuu. Tämä on käytännössä typerää, mutta voi helpottaa teoreettista tarkastelua. $\square$

3.8 Harjoitustehtäviä lukuun 3

3.1 Harjoitustehtävä. Colaco Oy valmistaa kasvisuutejuomia. Colaco harkitsee uuden suklaalla maustetun kasvisuutejuoman, Chocolan, markkinointia. Colacolla on kolme vaihtoehtoa:

1. Teettää markkinointitutkimus Chocolasta, ja sitten joko markkinoida tai olla markkinoimatta Chocolaa, riippuen tutkimuksen tuloksesta.
2. Markkinoida Chocolaa.
3. Olla markkinoimatta Chocolaa.

Ilman markkinointitutkimusta Colaco arvioi, että Chocolalla on 55% todennäköisyys menestyä ja 45% todennäköisyys flopata. Jos Chocola menestyy, Colaco saa 300.000€ voitot. Jos taas Chocola floppaa, Colaco kärsii 100.000€ tappiot. Jos markkinointitukimus antaa suotuisan tuloksen, niin Chocola menestyy markkinoilla 85% todennäköisyydellä. Markkinointitutkimus antaa suotuisan tuloksen 60% todennäköisyydellä. Jos taas markkinointitutkimus antaa epäsuotuisan tuloksen, niin Chocola menestyy 10% todennäköisyydellä. Markkinointitutkimus maksaa 30.000€.

Mitä Colacon tulisi tehdä, kun se on riskineutraali päätöksenteossaan?

3.2 Harjoitustehtävä. Douppausvalvontakomitean tulee päättää testataanko Narkomaan joukkue. Douppausvalvontakomitealla on käytössä testi, joka on 90% luotettava. Douppausvalvontakomitea arvioi, että keskimäärin 5% kaikista urheilijoista douppaa. Tilanteeseen liittyy kolmenlaisia kuluja:

$$\begin{aligned} c_1 &= \text{kulu, jos joukkuetta syytetään väärin douppauksesta,} \\ c_2 &= \text{kulu, jos joukkueen douppausta ei huomata,} \\ c_3 &= \text{kulu yksityisyyden loukkauksesta.} \end{aligned}$$

(a) Olkoon kulut $c_1 = 10$, $c_2 = 5$ ja $c_3 = 1$. Tuleekon douppausvalvontakomiten testata Narkomaan joukkue?
(b) Osoita että jos $c_1 > c_2 > c_3$, niin Narkomaan joukkuetta ei tule testata.

3.3 Harjoitustehtävä. Olkoon epävarman tilanteen todennäköisyysjakauma $p = (p_1, p_2, \ldots)$. Epävarmuuden, eli satunnaisuuden, suuruutta mitataan *entropialla*:

$$H(p) = -\sum_{j\in J} p_j \ln p_j.$$

(a) Millainen jakauma maksimoi epävarmuuden eli entropian?
(b) Millainen jakauma minimoi epävarmuuden eli entropian?
(c) Kasvattaako vai vähentääkö esimerkissä 3.5.4 markkina-asiantuntijan käyttö entropiaa?

3.4 Harjoitustehtävä. Kuinka paljon esimerkin 3.2.1 öljy-yhtiön kannattaisi korkeintaan maksaa oraakkelille, joka kertoo onko paikassa P öljyä vai ei?

3.5 Harjoitustehtävä. (a) Kuinka paljon esimerkin 3.4.1 pankin kannattaisi korkeintaan maksaa asiakastutkimuksesta?

(b) Kuinka paljon esimerkin 3.4.1 pankin kannattaisi korkeintaan maksaa oraakkelille, joka kertoo pystyykö asiakas maksamaan lainansa takaisin vai ei?

3.6 Harjoitustehtävä. Omppukone Oy valmistaa liukuhihnalla muistipiirejä kymmenen piirin sarjoissa. Omppukone arvioi, että keskimäärin 80% sarjoista sisältää 10% viallisia piirejä ja 20% sarjoista sisältää 50% viallisia piirejä. Jos "hyvä sarja" (siis sellainen, jossa on vain 10% viallisia piirejä) lähetetään liukuhihnalla eteenpäin se tulee maksamaan prosessointikuluina 1.000€. Jos "huono sarja" (siis sellainen, jossa on 50% viallisia piirejä) lähetetään liukuhihnalla eteenpäin se tulee maksamaan prosessointikuluina 4.000€. Omppukone voi vaihtoehtoisesti korjata sarjan hinnalla 1.000€. Korjattu sarja on automaattisesti "hyvä sarja". Lisäksi Omppukone voi halutessaan 100€ hinnalla testata täsmälleen yhden piirin sarjasta.

(a) Mitä Omppukoneen tulee tehdä?
(b) Kuinka paljon Omppukoneen kannattaa korkeintaan maksaa piirin testaamisesta?
(c) Kuinka paljon Omppukoneen kannattaa maksaa oraakkelille, joka ilmoittaa jokaisesta sarjasta, onko se "hyvä sarja" vai "huono sarja"?

Huomautus: Tehtävässä oletetaan, että Omppukone voi testata *täsmälleen* yhden piirin sarjasta. Tilanne muuttuu paljon mielenkiintoisemmaksi (eli vaikeammaksi), jos Omppukone voi testata niin monta piiriä kuin haluaa, joko samanaikaisesti tai peräkkäin. Innokas opiskelija voi halutessaan ratkaista Omppukoneen ongelman tässä laajennetussa muodossa.

3.7 Harjoitustehtävä. (a) Kuinka paljon esimerkissä 3.6.2 öljy-yhtiön kannattaisi korkeintaan maksaa geologille?

(b) Kuinka paljon esimerkissä 3.6.2 öljy-yhtiön kannattaisi korkeintaan maksaa oraakkelille, joka ilmoittaa onko paikassa P öljyä vai ei?

3.8 Harjoitustehtävä. (a) Päätöspuu on *vaihteleva*, jos mitään sattumasolmua ei välittömästi seuraa sattumasolmu eikä mitään päätössolmua välittömästi seuraa päätössolmu. Osoita, että jokainen päätöspuu on ekvivalentti jonkin vaihtelevan päätöspuun kanssa.

(b) Päätöspuu on *binäärinen*, jos jokainen solmu (päätös tai sattuma), haarautuu täsmälleen kahdeksi alipuuksi. Osoita, että jokainen päätöspuu on ekvivalentti binäärisen päätöspuun kanssa.

(c) Osoita, että jokainen päätöspuu on ekvivalentti jonkin vaihtelevan binäärisen puun kanssa.

3.9 Harjoitustehtävä. (a) Esitä esimerkkien 3.2.1 ja 3.4.1 päätöspuut vaihtelevina päätöspuina.

(b) Esitä esimerkkien 3.2.1 ja 3.4.1 päätöspuut vaihtelevina binäärisinä päätöspuina.

3.10 Harjoitustehtävä. Todista lause 3.7.1.

3.11 Harjoitustehtävä. Miksi päätöspuiden kohdalla käytetään ainoastaan odotusarvosääntöa ja hyötysääntöä? Miksi esimerkiksi optimistin sääntöä tai katumuksen kaihtamista ei käytetä?

Luku 4

Todennäköisyyksien estimointi

In God we trust; all others must bring data. – W. Edwards Deming

Torture numbers, and they'll confess to anything. – Gregg Easterbrook

It is the mark of a truly intelligent person to be moved by statistics.
– George Bernard Shaw

Tähän asti päätöksenteon pohjana olevat todennäköisyydet on "annattu taivaasta". Nyt esitämme menetelmiä todennäköisyyksien estimointiin datasta. Kaikki esitettävät menetelmät olettavat, että estimointiin käytetty data on (riittävän) *riippumatonta* ja *samoin jakautunutta*. Harjoitustehtävissä 4.3(b), 4.6 ja 4.7 näistä oletuksista luovutaan hieman.

4.1 Suhteellisten frekvenssien menetelmä

Varmaankin yksinkertaisin ja luontevin tapa estimoida todennäköisyyksiä on *suhteellisten frekvenssien menetelmä* eli käyttää empiiristä todennäköisyysjakaumaa. Tällöin satunnaiskokeeseen X liittyvän tapahtuman $\{X = x\}$ estimoitu todennäköisyys on sen esiintymiskertojen suhteellinen lukumäärä, eli frekvenssi, pitkässä toistojen sarjassa. Menetelmässä siis oletetaan, että toistetut satunnaiskokeet $X_1, X_2, \ldots$ ovat (riittävän) *riippumattomia* ja *samankaltaisia*. Menetelmä perustuu todennäköisyyden frekvenssitulkintaan ja sitä kautta suurten lukujen lakiin 1.3.19.

4.1.1 Määritelmä (Suhteellisten frekvenssien menetelmä)**.** Jos on saatu havainnot $x_1, \ldots, x_n$, niin *suhteellisten frekvenssien menetelmän* antama estimaatti $\hat{\mathrm{P}}(X = x)$ todennäköisyydelle $\mathrm{P}(X = x)$ on

$$\hat{\mathrm{P}}(X = x) \quad = \quad \hat{\mathrm{P}}(X = x | x_1, \ldots x_n) \quad = \quad \frac{n(x|x_1, \ldots, x_n)}{n},$$

missä $n(x|x_1, \ldots, x_n)$ on arvon x esiintymiskertojen lukumäärä havainnoissa $x_1, \ldots, x_n$.

4.1.2 Esimerkki. Jos havainnot ovat $x_1 = 1$, $x_2 = 0$, $x_3 = 0$, $x_4 = 1$, $x_5 = 0$, $x_6 = 42$, niin estimoitu todennäköisyysjakauma on

$$\begin{aligned} p_0 &= \mathrm{P}(X=0) = 3/6 = 0{,}5000, \\ p_1 &= \mathrm{P}(X=1) = 2/6 = 0{,}3333, \\ p_{42} &= \mathrm{P}(X=42) = 1/6 = 0{,}1667. \end{aligned}$$

Jos sitten saadaan uusi havainto $x_7 = 0$, niin uusi estimoitu todennäköisyysjakauma on

$$\begin{aligned} p_0 &= \mathrm{P}(X=0) = 4/7 = 0{,}5714, \\ p_1 &= \mathrm{P}(X=1) = 2/7 = 0{,}2857, \\ p_{42} &= \mathrm{P}(X=42) = 1/7 = 0{,}1429. \end{aligned}$$

Seuraava uusi havainto $x_8 = -1$, aiheuttaa uuden estimoitun todennäköisyysjakauman

$$\begin{aligned} p_{-1} &= \mathrm{P}(X=-1) = 1/8 = 0{,}1250, \\ p_0 &= \mathrm{P}(X=0) = 4/8 = 0{,}5000, \\ p_1 &= \mathrm{P}(X=1) = 2/8 = 0{,}2500, \\ p_{42} &= \mathrm{P}(X=42) = 1/8 = 0{,}1250. \end{aligned}$$

Jos vielä saadaan uusi havainto $x_9 = \pi$, niin uusi estimoitu todennäköisyysjakauma on

$$\begin{aligned} p_{-1} &= \mathrm{P}(X=-1) = 1/9 = 0{,}1111, \\ p_0 &= \mathrm{P}(X=0) = 4/9 = 0{,}4444, \\ p_1 &= \mathrm{P}(X=1) = 2/9 = 0{,}2222, \\ p_\pi &= \mathrm{P}(X=\pi) = 1/9 = 0{,}1111, \\ p_{42} &= \mathrm{P}(X=42) = 1/9 = 0{,}1111. \end{aligned}$$

Suhteellisten frekvenssien menetelmässä siis uusi data voi sekä muuttaa vanhoja todennäköisyyksiä että tuoda mukaan uusia tulosmahdollisuuksia. Ilmeistä on myös, että menetelmä suppenee hitaasti. Toisin sanoen, jos dataa on vähän, niin estimaatit voivat muuttua paljonkin tulevaisuudessa.

4.1.3 Esimerkki (Leipuri Pulla estimoi suhteellisilla frekvensseillä)**.** Leipuri Pulla myy pullia Kumputien Leipomossa. Hän paistaa pullat aamulla ja myy ne lounastauolla viereisen Ministeriön Erikoisosaston virkamiehille. Eilisiä pullia ei voi tänään enää myydä. Leipuri Pullan leipomo on auki maanantaista perjantaihin, muttei pyhinä tai aattoina.

Pullan paistaminen maksaa leipuri Pullalle 0,20 € pullalta, ja hän myy niitä 1,00 € kappalehintaan. Siten, jos leipuri Pulla paistaa aamulla i pullaa ja häneltä kysytäään lounastauolla j pullaa, niin hänen voittonsa on

$$r_{ij} \quad = \quad 1{,}00\,€ \cdot \min(i,j) - 0{,}20\,€ \cdot i.$$

Leipuri Pullan tulee päättää, kuinka monta pullaa pitää paistaa aamulla lounastaukoa varten, jotta hän saisi parhaan mahdollisen voiton.

Leipuri pulla on kerännyt seuraavan datan myymistään pullista:

	Lokakuu 2009				
Viikko	Ma	Ti	Ke	To	Pe
40				2	0
41	3	7	10	4	0
42	7	5	10	6	4
43	4	5	6	10	2
44	8	6	5	6	2

	Marraskuu 2009				
Viikko	Ma	Ti	Ke	To	Pe
45	8	4	3	3	3
46	4	4	6	8	2
47	8	5	8	2	0
48	3	10	9	5	2
49	7				

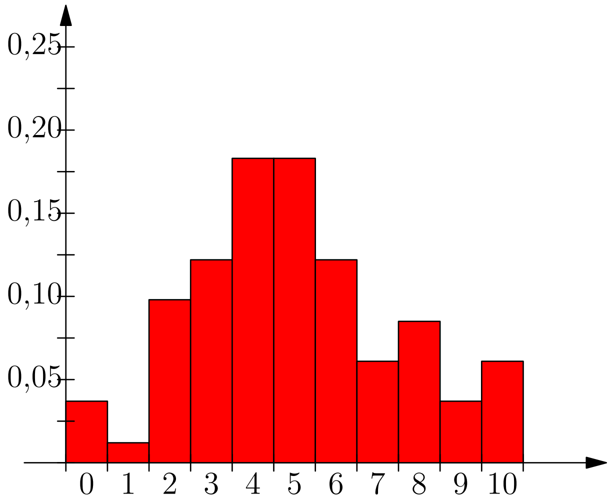
	Joulukuu 2009				
Viikko	Ma	Ti	Ke	To	Pe
49		4	4	6	2
50	7	5	5	10	3
51	4	4	5	6	1
52	5	4	4	–	–
53	7	5	9	–	

	Tammikuu 2010				
Viikko	Ma	Ti	Ke	To	Pe
53					–
1	3	4	–	4	2
2	9	8	5	6	4
3	5	6	5	6	3
4	3	8	5	5	3

Datapisteiden lukumäärä on 82 ja saatujen arvojen $0,\ldots,10$ suhteelliset frekvenssit $p_k = \mathrm{P}(X=k)$ $(k=0,\ldots,10)$ ovat

$$\begin{aligned} p_0 &= 3/82 = 0{,}037, & p_6 &= 10/82 = 0{,}122, \\ p_1 &= 1/82 = 0{,}012, & p_7 &= 5/82 = 0{,}061, \\ p_2 &= 8/82 = 0{,}098, & p_8 &= 7/82 = 0{,}085, \\ p_3 &= 10/82 = 0{,}122, & p_9 &= 3/82 = 0{,}037, \\ p_4 &= 15/82 = 0{,}183, & p_{10} &= 5/82 = 0{,}061. \\ p_5 &= 15/82 = 0{,}183, \end{aligned}$$

Graafisesti saatu empiirinen todennäköisyysjakauma on siis:



Esimerkin 4.1.3 leipuri Pulla vaihtoehdot ovat

$$a_i \quad = \quad \text{“Paistetaan aamulla } i \text{ pullaa”}.$$

Hän arvottaa ne riskineutraalisti, eli odotusarvon mukaan:

$$\begin{aligned} V(a_i) &= \mathrm{E}(a_i) = \sum_{j \in J} r_{ij}\, p_j \\ &= \sum_{j=0}^{10} \Big(\min(i,j) - 0{,}2i \Big)\, p_j. \end{aligned}$$

Jos leipuri Pulla ja uskoo suhteellisten frekvenssien menetelmään, hänen estimaattinsa todennäköisyyksielle on siis annettu edellä ja saamme numeeriset arvot

$$\begin{aligned} V(a_0) \quad = \quad & 0{,}00 \cdot 0{,}037 + 0{,}00 \cdot 0{,}012 + 0{,}00 \cdot 0{,}098 + 0{,}00 \cdot 0{,}122 + \\ & 0{,}00 \cdot 0{,}183 + 0{,}00 \cdot 0{,}183 + 0{,}00 \cdot 0{,}122 + 0{,}00 \cdot 0{,}061 + \\ & 0{,}00 \cdot 0{,}085 + 0{,}00 \cdot 0{,}037 + 0{,}00 \cdot 0{,}061 \quad = \quad 0{,}000, \end{aligned}$$

$$\begin{aligned} V(a_1) \quad = \quad & -0{,}20 \cdot 0{,}037 + 0{,}80 \cdot 0{,}012 + 0{,}80 \cdot 0{,}098 + 0{,}80 \cdot 0{,}122 + \\ & 0{,}80 \cdot 0{,}183 + 0{,}80 \cdot 0{,}183 + 0{,}80 \cdot 0{,}122 + 0{,}80 \cdot 0{,}061 + \\ & 0{,}80 \cdot 0{,}085 + 0{,}80 \cdot 0{,}037 + 0{,}80 \cdot 0{,}061 \quad = \quad 0{,}764, \end{aligned}$$

$$\begin{aligned} V(a_2) \quad = \quad & -0{,}40 \cdot 0{,}037 + 0{,}60 \cdot 0{,}012 + 1{,}60 \cdot 0{,}098 + 1{,}60 \cdot 0{,}122 + \\ & 1{,}60 \cdot 0{,}183 + 1{,}60 \cdot 0{,}183 + 1{,}60 \cdot 0{,}122 + 1{,}60 \cdot 0{,}061 + \\ & 1{,}60 \cdot 0{,}085 + 1{,}60 \cdot 0{,}037 + 1{,}60 \cdot 0{,}061 \quad = \quad 1{,}516, \end{aligned}$$

$$
\begin{aligned}
V(a_3) &= -0{,}60 \cdot 0{,}037 + 0{,}40 \cdot 0{,}012 + 1{,}40 \cdot 0{,}098 + 2{,}40 \cdot 0{,}122 + \\
&\quad 2{,}40 \cdot 0{,}183 + 2{,}40 \cdot 0{,}183 + 2{,}40 \cdot 0{,}122 + 2{,}40 \cdot 0{,}061 + \\
&\quad 2{,}40 \cdot 0{,}085 + 2{,}40 \cdot 0{,}037 + 2{,}40 \cdot 0{,}061 \quad = \quad 2{,}169,
\end{aligned}
$$

$$
\begin{aligned}
V(a_4) &= -0{,}80 \cdot 0{,}037 + 0{,}20 \cdot 0{,}012 + 1{,}20 \cdot 0{,}098 + 2{,}20 \cdot 0{,}122 + \\
&\quad 3{,}20 \cdot 0{,}183 + 3{,}20 \cdot 0{,}183 + 3{,}20 \cdot 0{,}122 + 3{,}20 \cdot 0{,}061 + \\
&\quad 3{,}20 \cdot 0{,}085 + 3{,}20 \cdot 0{,}037 + 3{,}20 \cdot 0{,}061 \quad = \quad 2{,}701,
\end{aligned}
$$

$$
\begin{aligned}
V(a_5) &= -1{,}00 \cdot 0{,}037 + 0{,}00 \cdot 0{,}012 + 1{,}00 \cdot 0{,}098 + 2{,}00 \cdot 0{,}122 + \\
&\quad 3{,}00 \cdot 0{,}183 + 4{,}00 \cdot 0{,}183 + 4{,}00 \cdot 0{,}122 + 4{,}00 \cdot 0{,}061 + \\
&\quad 4{,}00 \cdot 0{,}085 + 4{,}00 \cdot 0{,}037 + 4{,}00 \cdot 0{,}061 \quad = \quad 3{,}050,
\end{aligned}
$$

$$
\begin{aligned}
V(a_6) &= -1{,}20 \cdot 0{,}037 - 0{,}20 \cdot 0{,}012 + 0{,}80 \cdot 0{,}098 + 1{,}80 \cdot 0{,}122 + \\
&\quad 2{,}80 \cdot 0{,}183 + 3{,}80 \cdot 0{,}183 + 4{,}80 \cdot 0{,}122 + 4{,}80 \cdot 0{,}061 + \\
&\quad 4{,}80 \cdot 0{,}085 + 4{,}80 \cdot 0{,}037 + 4{,}80 \cdot 0{,}061 \quad = \quad 3{,}216,
\end{aligned}
$$

$$
\begin{aligned}
V(a_7) &= -1{,}40 \cdot 0{,}037 - 0{,}40 \cdot 0{,}012 + 0{,}60 \cdot 0{,}098 + 1{,}60 \cdot 0{,}122 + \\
&\quad 2{,}60 \cdot 0{,}183 + 3{,}60 \cdot 0{,}183 + 4{,}60 \cdot 0{,}122 + 5{,}60 \cdot 0{,}061 + \\
&\quad 5{,}60 \cdot 0{,}085 + 5{,}60 \cdot 0{,}037 + 5{,}60 \cdot 0{,}061 \quad = \quad 3{,}260,
\end{aligned}
$$

$$
\begin{aligned}
V(a_8) &= -1{,}60 \cdot 0{,}037 - 0{,}60 \cdot 0{,}012 + 0{,}40 \cdot 0{,}098 + 1{,}40 \cdot 0{,}122 + \\
&\quad 2{,}40 \cdot 0{,}183 + 3{,}40 \cdot 0{,}183 + 4{,}40 \cdot 0{,}122 + 5{,}40 \cdot 0{,}061 + \\
&\quad 6{,}40 \cdot 0{,}085 + 6{,}40 \cdot 0{,}037 + 6{,}40 \cdot 0{,}061 \quad = \quad 3{,}242,
\end{aligned}
$$

$$
\begin{aligned}
V(a_9) &= -1{,}80 \cdot 0{,}037 - 0{,}80 \cdot 0{,}012 + 0{,}20 \cdot 0{,}098 + 1{,}20 \cdot 0{,}122 + \\
&\quad 2{,}20 \cdot 0{,}183 + 3{,}20 \cdot 0{,}183 + 4{,}20 \cdot 0{,}122 + 5{,}20 \cdot 0{,}061 + \\
&\quad 6{,}20 \cdot 0{,}085 + 7{,}20 \cdot 0{,}037 + 7{,}20 \cdot 0{,}061 \quad = \quad 3{,}140,
\end{aligned}
$$

$$
\begin{aligned}
V(a_{10}) &= -2{,}00 \cdot 0{,}037 - 1{,}00 \cdot 0{,}012 + 0{,}00 \cdot 0{,}098 + 1{,}00 \cdot 0{,}122 + \\
&\quad 2{,}00 \cdot 0{,}183 + 3{,}00 \cdot 0{,}183 + 4{,}00 \cdot 0{,}122 + 5{,}00 \cdot 0{,}061 + \\
&\quad 6{,}00 \cdot 0{,}085 + 7{,}00 \cdot 0{,}037 + 8{,}00 \cdot 0{,}061 \quad = \quad 3{,}001.
\end{aligned}
$$

Riskeneutraali leipuri Pulla paistaa siis 7 pullaa aamulla.

4.1.4 Huomautus (Suhteellisten frekvenssien menetelmän hyviä ja huonoja puolia)**.** Hyvää suhteellisten frekvenssien menetelmässä on:

$\oplus$ se on helppo ymmärtää ja toteuttaa,

$\oplus$ se ei perustu mihinkään teoreettiseen malliin, joten mallivirheen riskiä ei ole.

Huonoa suhteellisten frekvenssien menetelmässä on:

$\ominus$ se ei mallita tilanteeseen mahdollisesti liittyvää erikoisluonnetta,

$\ominus$ se vaatii paljon dataa, jotta todennäköisyydet olisivat luotettavasti estimoidut.

4.2 Teoreettisen mallin sovittaminen

Tässä osiossa tarkastelemme ns. *parametristä tilastollista päättelyä.*

Joskus voidaan satunnaiskokeesta X olettaa jotain. Tyypillisesti oletetaan, että X on jakautunut jonkin tunnetun jakauman — mutta tuntemattoman parametrin — mukaan. Esimerkiksi kolikonheitto on $\mathrm{Bin}(1,\theta)$-jakautunut, missä θ — klaavan todennäköisyys — on tuntematon parametri. Samoin, jos heitetään kolikkoa 3 kertaa peräkkäin ja lasketaan klaavat, niin tulos on $\mathrm{Bin}(3,\theta)$-jakautunut. Satunnaiskokeen jakauman tyypi siis tiedetään, mutta täsmentävä parametritieto puuttuu. Tuntematon parametri pitää siis estimoida datasta. Menetelmässä oletetaan, että satunnaiskokeet $X_1, X_2, \ldots$ ovat (riittävän) *riippumattomia* ja *samankaltaisia*, ja että satunnaiskokeen X jakauma tunnetaan "parametria vaille". Menetelmä perustuu suurten lukujen lakiin.

Tuntemattoman parametrin estimointiin on olemassa ainakin kaksi kilpailevaa menetelmää: bayesläinen tilastollinen päättely ja suurimman uskottavuuden periaate. Bayesläistä tilastollista päättelyä emme tällä kurssilla käsittele.

4.2.1 Määritelmä (Suurimman uskottavuuden periaate)**.** *Suurimman uskottavuuden periaate* on, kuten nimestäkin käy ilmi, se, että uskomme satunnaisten tapahtumien tapahtuvan todennäköisimmällä mahdollisella tavalla. Toisin sanoin: mitä uskottavampi tapahtuma on, sitä enemmän siihen uskomme.

Tarkemmin suurimman uskottavuuden periaatteessa on kyse seuraavasta:

- on annettu data, eli havainnot, $x_1, \ldots, x_n$,
- on mallinnettu yksittäisen havainnon x todennäköisyys, tai todennäköisyystiheys[1], $p(x|\theta)$, missä parametri θ on tuntematon.

[1] *Todennäköisyystiheys* $p(x)$ tarkoittaa sitä, että todennäköisyys saada havainto väliltä $[a,b]$ on integraali

$$\int_a^b p(x)\,\mathrm{d}x,$$

eli todennäköisyys, että saatu havainto on Δ:n päässä pisteestä x on noin $p(x)\Delta$, kun Δ on pieni.

Datan $x_1, \ldots, x_n$ todennäköisyys, tai todennäköisyystiheys, jos θ tunnettaisiin olisi

$$p(x_1, \ldots, x_n|\theta) \quad = \quad p(x_1|\theta) \cdots p(x_n|\theta)$$

(tulomuoto johtuu riippumattomuusoletuksesta). Koska θ on tuntematon tulkitsemme todennäköisyysfunktion $p(x_1, \ldots, x_n|\theta)$ toisin päin datan *uskottavuutena* $L(\theta|x_1, \ldots, x_n)$.

4.2.2 Määritelmä (Suurimman uskottavuuden estimaattori)**.** Parametrin θ *suurimman uskottavuuden estimaattori* $\hat{\theta}$ on se tuntemattoman parametrin θ arvo, jolla uskottavuus $L(\theta|x_1, \ldots, x_n)$ maksimoituu:

$$\hat{\theta} \quad = \quad \underset{\theta}{\operatorname{argmax}}\, L(\theta|x_1, \ldots, x_n).$$

Käytännössä $\hat{\theta}$ lasketaan usein tyypillisellä “koulupojan optimointimenetelmällä”: derivoidaan uskottavuus $L(\theta|x_1, \ldots, x_n)$ parametrin θ suhteen, ja optimi $\hat{\theta}$ löytyy derivaatan nollakohdasta. Toisin sanoen optimi $\hat{\theta}$ on se piste, jossa

$$\frac{\mathrm{d}L}{\mathrm{d}\theta}(\hat{\theta}|x_1, \ldots, x_n) \quad = \quad 0.$$

4.2.3 Huomautus (Log-uskottavuus)**.** Koska tulojen derivointi on vaivalloista, tarkastellaan usein *log-uskottavuutta*

$$\ell(\theta|x_1, \ldots, x_n) \quad = \quad \ln L(\theta|x_1, \ldots, x_n).$$

Koska logaritmi on aidosti kasvava funktio on log-uskottavuuden maksimointi yhtäpitävää uskottavuuden maksimoinnin kanssa. Log-uskottavuutta on kuitenkin mukavampi derivoida, sillä se on summamuotoinen:

$$\begin{aligned} \ell(\theta|x_1, \ldots, x_n) &= \ln\Big(p(x_1|\theta) \cdots p(x_n|\theta)\Big) \\ &= \ln p(x_1|\theta) + \cdots + \ln p(x_n|\theta). \end{aligned}$$

Siten, käyttämällä derivointisääntöjä

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\theta}\Big(f(\theta) + g(\theta)\Big) &= \frac{\mathrm{d}}{\mathrm{d}\theta}f(\theta) + \frac{\mathrm{d}}{\mathrm{d}\theta}g(\theta), \\ \frac{\mathrm{d}}{\mathrm{d}\theta}\ln f(\theta) &= \frac{\frac{\mathrm{d}f}{\mathrm{d}\theta}(\theta)}{f(\theta)}, \end{aligned}$$

saamme

$$\frac{\mathrm{d}\ell}{\mathrm{d}\theta}(\theta|x_1, \ldots, x_n) \quad = \quad \frac{\frac{\mathrm{d}p}{\mathrm{d}\theta}(x_1|\theta)}{p(x_1|\theta)} + \cdots + \frac{\frac{\mathrm{d}p}{\mathrm{d}\theta}(x_n|\theta)}{p(x_n|\theta)}.$$

4.2.4 Huomautus (Teoreettisen mallin sovittamismenetelmän hyviä ja huonoja puolia)**.** Hyviä puolia teoreettisen mallin sovittamisessa on se, että se

- $\oplus$ tarvitsee vain vähän dataa,
- $\oplus$ mahdollistaa taustatietojen käytön mallinnuksessa.

Huonoja puolia taas ovat

- $\ominus$ mallivirheen riski,
- $\ominus$ joskus menetelä on vaikea ymmärtää tai toteuttaa.

Teoreettisia malleja

Esitämme kaksi mallia esimerkkeinä: binomimallin ja Poisson-mallin. Laskemme suurimman uskottavuuden estimaattorit näissä malleissa sekä vielä tarkastelemme miten nämän mallit toimivat leipuri Pullan esimerkin 2.4.8 tapauksessa.

4.2.5 Esimerkki (Binomimalli)**.** Tarkastelemme satunnaiskoetta, jossa on n riippumatonta samankaltaista toistoa ja jokaisella kerralla voidaan joko "onnistua" todennäköisyydellä θ tai "epäonnistua" todennäköisyydellä $1–\theta$. Oletamme, että n on tunnettu ja θ on tuntematon parametri.

Binomimallissa parametrin θ suurimman uskottavuuden estimaattori on "onnistumisten" lukumäärä jaettuna toistojen lukumäärällä. Perustelemme tämän. Olkoon data $x_1, \ldots, x_n$, missä $x_k = 1$, jos toistolla k onnistuttiin ja $x_k = 0$ muulloin. Tällöin "onnistumisten" lukumäärä jaettuna toistojen lukumäärällä on keskiarvo

$$\bar{x} \quad = \quad \frac{x_1 + \cdots + x_n}{n}.$$

Nyt $n\bar{x}$ on "onnistumisten" lukumäärä, ja siten[2]

$$p(x_1, \ldots, x_n|\theta) \quad = \quad \theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}}.$$

Saamme log-uskottavuuden

$$\begin{aligned}
\ell(\theta|x_1, \ldots, x_n) &= \ln\left(\theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}}\right) \\
&= \ln\left(\theta^{n\bar{x}}\right) + \ln\left((1-\theta)^{n-n\bar{x}}\right) \\
&= n\bar{x}\ln\theta + (n - n\bar{x})\ln(1-\theta).
\end{aligned}$$

[2]Oikeastaan binomimallissa yleensä oletetaan, että "onnistumisten" paikkoja $x_1, \ldots, x_n$ ei havaita, vaan ainoastaan niiden lukumäärä $n\bar{x}$. Tällöin oikea todennäköisyysfunktio on

$$p(n\bar{x}|\theta) \quad = \quad \binom{n}{n\bar{x}}\theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}}.$$

Koska binomikerroin $\binom{n}{n\bar{x}}$ on vakio parametrin θ suhteen ei tällä kuitenkaan ole mitään vaikutusta suurimman uskottavuuden analyysiin.

Siten

$$\begin{aligned}
\frac{\mathrm{d}\ell}{\mathrm{d}\theta}(\theta|x_1,\ldots,x_n) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\Big(n\bar{x}\ln\theta + (n-n\bar{x})\ln(1-\theta)\Big) \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\Big(n\bar{x}\ln\theta\Big) + \frac{\mathrm{d}}{\mathrm{d}\theta}\Big((n-n\bar{x})\ln(1-\theta)\Big) \\
&= n\bar{x}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\theta + (n-n\bar{x})\frac{\mathrm{d}}{\mathrm{d}\theta}\ln(1-\theta) \\
&= n\bar{x}\cdot\frac{1}{\theta} + (n-n\bar{x})\cdot(-1)\cdot\frac{1}{1-\theta} \\
&= \frac{n\bar{x}}{\theta} - \frac{n-n\bar{x}}{1-\theta} \\
&= \frac{(1-\theta)n\bar{x} - \theta(n-n\bar{x})}{\theta(1-\theta)} \\
&= \frac{n\bar{x} - \theta n\bar{x} - \theta n + \theta n\bar{x}}{\theta(1-\theta)} \\
&= \frac{n(\bar{x}-\theta)}{\theta(1-\theta)}.
\end{aligned}$$

Nyt

$$\frac{n(\bar{x}-\theta)}{\theta(1-\theta)} = 0$$

jos ja vain jos yläkerta $n(\bar{x}-\theta)=0$, eli $\theta=\bar{x}$.

Suurimman uskottavuuden estimaattori binomimallissa on siis otoskeskiarvo:

$$\hat{\theta} = \bar{x}.$$

4.2.6 Esimerkki (Binomimalli leipuri Pullalle)**.** Leipuri Pullan esimerkissä 4.1.3 voimme ajatella, että päivittäiset pullanmyynnit ovat binomijakautuneet parametrein 10 (tunnettu)[3] ja θ (tuntematon). Tällöin 82 päivän pullanmyynnit on binomijakautunut parametrein 820 ja θ. Suurimman uskottavuuden estimaattori parametrille θ on tällöin pullien kokonaismyynti jaettuna 820:lla:

$$\hat{\theta} = \frac{410}{820} = 0{,}5.$$

Erityisesti siis arvio todennäköisyydelle, että jonakin päivän myydään x pullaa ($x=0,1,\ldots,10$) on

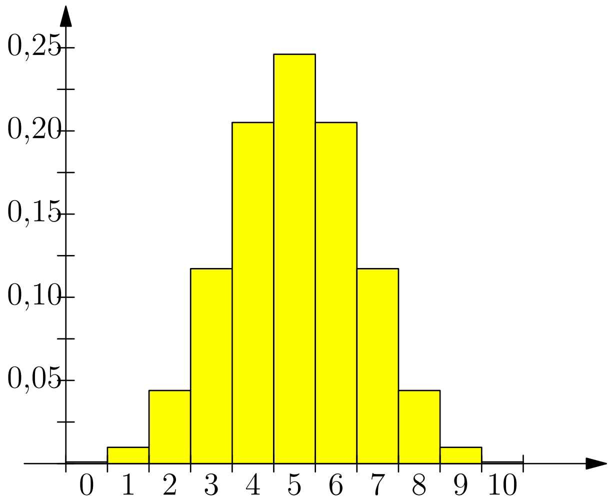
$$\binom{10}{x}0{,}5^{x}\cdot 0{,}5^{10-x} = \binom{10}{x}0{,}5^{10}.$$

Sijoittamalla luvut $x=0,\ldots,10$ yllä olevaan kaavaan saamme numeeriset arvot

[3] Parametrin n arvoksi on valittu rohkeasti suurimman havainnon arvo. Jos joskus tulevaisuudessa havaitaan suurempi arvo, esimerkiksi 42, aiheuttaa tämä ongelmia.

$$
\begin{aligned}
p_0 &= 0{,}0010, & p_4 &= 0{,}2051, & p_8 &= 0{,}0439,\\
p_1 &= 0{,}0098, & p_5 &= 0{,}2461, & p_9 &= 0{,}0098,\\
p_2 &= 0{,}0439, & p_6 &= 0{,}2051, & p_{10} &= 0{,}0010.\\
p_3 &= 0{,}1172, & p_7 &= 0{,}1172, & &
\end{aligned}
$$

Kuvallinen esitys tästä jakaumasta on (kuvan skaala on sama kuin suhteellisten frekvessien tapauksessa):



Esimerkin 4.1.3 leipuri Pullan vaihtoehdot ovat

$$a_i = \text{“Paistetaan aamulla } i \text{ pullaa”}.$$

Hän arvottaa ne riskineutraalisti, eli odotusarvon mukaan:

$$
\begin{aligned}
V(a_i) = \mathrm{E}(a_i) &= \sum_{j\in J} r_{ij}\, p_j \\
&= \sum_{j=0}^{10} \Big(\min(i,j) - 0{,}2i\Big) \cdot \binom{10}{j} 0{,}5^{10}
\end{aligned}
$$

($i = 0, \ldots, 10$), missä viimeinen yhtälö tulee siitä, että leipuri Pulla uskoo binomimalliin parametrein 10 ja 0,5.

Jätämme harjoitustehtäväksi 4.1(a) laskea leipuri Pullan arvotukset tässä tilanteessa ja määrätä hänen optimaalinen valintansa.

4.2.7 Esimerkki (Poisson-malli)**.** Binomimalli olettaa, että satunnaiskokeen tuloksen mahdolliset arvot $0, 1, \ldots, n$ on rajoitettu (ja tunnettu) joukko. Näin kuitenkin on varsin harvoin, erityisestikään ylärajan n kohdalla. Usein hyvä malli lukumäärämuuttujalle on Poisson-jakauma. Poisson-jakaumaan päädytään esimerkiksi $\mathrm{Bin}(n, \theta)$-jakaumien raja-arvona, kun annetaan $n \to \infty$ ja $\theta \to 0$ siten, että $n\theta \to \lambda$. [4] *Poisson-jakauma* parametrilla λ on

$$p(x|\lambda) \quad = \quad e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Poisson-parametri λ on tyypillisesti tuntematon. Johdamme nyt sen suurimman uskottavuuden estimaattorin.

Olkoot (riippumattomat) havainnot $x_1, \ldots, x_n$ Poisson(λ)-jakaumasta. Tällöin log-uskottavuus on

$$\begin{aligned}
&\ell(\lambda|x_1, \ldots, x_n) \\
&\quad = \ln p(x_1|\lambda) + \cdots + \ln p(x_n|\lambda) \\
&\quad = \ln\left(e^{-\lambda}\frac{\lambda^{x_1}}{x_1!}\right) + \cdots + \ln\left(e^{-\lambda}\frac{\lambda^{x_n}}{x_n!}\right) \\
&\quad = \big(\ln e^{-\lambda} + \ln\lambda^{x_1} - \ln x_1!\big) + \cdots + \big(\ln e^{-\lambda} + \ln\lambda^{x_n} - \ln x_n!\big) \\
&\quad = -n\lambda + (x_1 + \cdots + x_n)\ln\lambda - \big(\ln x_1! + \cdots + \ln x_n!\big).
\end{aligned}$$

Siten

$$\frac{\mathrm{d}\ell}{\mathrm{d}\lambda}(\lambda|x_1, \ldots, x_n) \quad = \quad -n + \frac{x_1 + \cdots + x_n}{\lambda} \quad = \quad 0,$$

jos ja vain jos

$$\lambda \quad = \quad \frac{x_1 + \cdots + x_n}{n} \quad = \quad \bar{x}.$$

Siten suurimman uskottavuuden estimaattori $\hat{\lambda}$ on Poisson-mallissakin otoskeskiarvo

$$\hat{\theta} \quad = \quad \bar{x}.$$

4.2.8 Esimerkki (Poisson-malli leipuri Pullalle)**.** Leipuri Pullan esimerkissä 4.1.3 $\hat{\lambda} = 410/82 = 5$, mikä on olennaisesti sama arvo kuin edellä binomimallin tapauksessa. Erityisesti siis arvio myydä x pullaa $(x = 0, 1, 2, \ldots)$ jonakin päivänä on

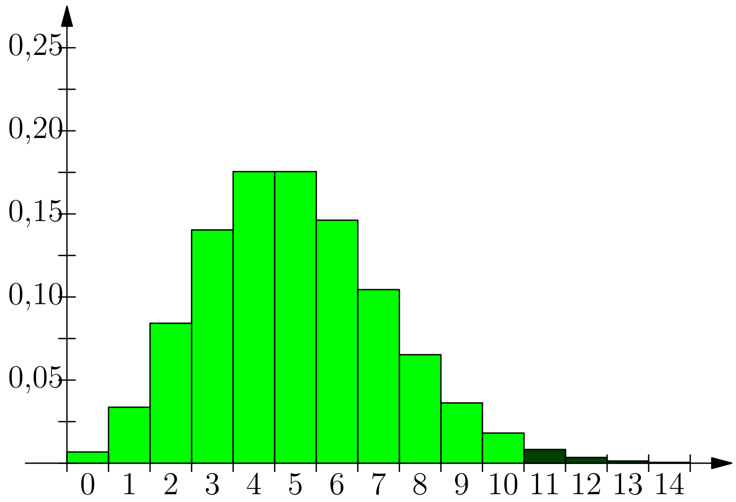
$$e^{-5}\frac{5^x}{x!}.$$

Tulosmahdollisuuksia on nyt ääretön määrä. Sijoittamalla yllä olevaan kaavaan x:n paikalle alkupäätä $x = 0, 1, 2, \ldots$ saamme numeeriset todennäköisyydet

[4] Tästä syystä Poisson-jakaumaa kutsutaan joskus leikkisästi "pienten lukujen laiksi". Vertaus liittyy tietysti "suurten lukujen lakiin" ja siihen liittyvään normaalijakaumaan.

$$
\begin{aligned}
p_0 &= 0{,}0067, & p_6 &= 0{,}1462, & p_{12} &= 0{,}0034,\\
p_1 &= 0{,}0337, & p_7 &= 0{,}1044, & p_{13} &= 0{,}0013,\\
p_2 &= 0{,}0842, & p_8 &= 0{,}0653, & p_{14} &= 0{,}0005,\\
p_3 &= 0{,}1403, & p_9 &= 0{,}0363, & p_{15} &= 0{,}0002,\\
p_4 &= 0{,}1755, & p_{10} &= 0{,}0181, & p_{16} &= 0{,}0000.\\
p_5 &= 0{,}1755, & p_{11} &= 0{,}0082, & &
\end{aligned}
$$

Kuvallinen esitys tämän jakauman *alusta* on (Kuvan skaala on sama kuin suhteellisten frekvessien ja binomimallin tapauksissa. Tummanvihreällä on merkitty muutamia tulosmahdollisuuksia, joita ei aikaisemmin pidetty mahdollisina):



Esimerkin 4.1.3 leipuri Pulla vaihtoehdot ovat

$$a_i \quad = \quad \text{“Paistetaan aamulla } i \text{ pullaa”}.$$

Hän arvottaa ne riskineutraalisti, eli odotusarvon mukaan:

$$
\begin{aligned}
V(a_i) \quad = \quad \mathrm{E}(a_i) \quad &= \quad \sum_{j\in J} r_{ij}\, p_j \\
&= \quad \sum_{j=0}^{\infty} \Big(\min(i,j) - 0{,}2i\Big)\, e^{-5}\frac{5^j}{j!}
\end{aligned}
$$

($i = 0,\ldots,10$), missä viimeinen yhtälö tulee siitä, että leipuri Pulla uskoo Poisson-malliin parametrilla 5. Jotta voisimme laskea edellisen kaavan äärettömän summan, harrastamme hieman algebraa:

$$
\begin{aligned}
V(a_i) &= \sum_{j=0}^{\infty} \Big(1{,}00 \cdot \min(i,j) - 0{,}20 \cdot i\Big) \cdot e^{-5} \frac{5^j}{j!} \\
&= e^{-5} \sum_{j=0}^{\infty} \min(i,j) \frac{5^j}{j!} - 0{,}2ie^{-5} \sum_{j=0}^{\infty} \frac{5^j}{j!} \\
&= e^{-5} \sum_{j=0}^{\infty} \min(i,j) \frac{5^j}{j!} - 0{,}2i \\
&= e^{-5} \sum_{j=0}^{i} j \frac{5^j}{j!} + ie^{-5} \sum_{j=i+1}^{\infty} \frac{5^j}{j!} - 0{,}2i \\
&= e^{-5} \sum_{j=0}^{i} j \frac{5^j}{j!} + i \left(1 - e^{-5} \sum_{j=0}^{i} \frac{5^j}{j!}\right) - 0{,}2i.
\end{aligned}
$$

Tämä lauseke voidaan laskea kynällä ja paperilla (tai tietokoneella, jos et ole Työn Sankari™).

Jätämme harjoitustehtäväksi 4.1(b) laskea leipuri Pullan arvotukset tässä tilanteessa ja määrätä hänen optimaalinen valintansa.

4.3 Pearson–Tukey-menetelmä

Joskus data on turhan hienojakoista tarpeisiimme nähden. Esimerkiksi päätöspuissa turhan hienojakoinen data aiheuttaa sattumasolmuissa paljon haarautumisia. Hienojakoisesta datasta pääsee eroon *luokittelemalla*, tavalla tai toisella. Yksi tapa luokitella dataa on Pearson–Tukey-menetelmä. Oletamme edelleen, että data on (riittävän) *riippumatonta* ja *samoin jakautunutta*.

4.3.1 Määritelmä (Pearson–Tukey-menetelmä)**.** *Pearson–Tukey-menetelmä* tiivistää todennäköisyysjakauman, joko empiirisen tai teoreettisen, kolmen pisteen jakaumaksi siten, että jakauma 0,05-fraktiili ja 0,95-faktiili saavat molemmat todennäköisyyden 0,185 ja 0,5-fraktiili, eli mediaani, saa todennäköisyyden 0,630.

4.3.2 Huomautus. Pearson–Tukey-menetelmä on *ad hoc*[5], mutta sen on havaittu toimivan kohtalaisen hyvin, jos tuntematon jakauma on yksihuippuinen ja symmetrinen.

4.3.3 Huomautus. Huomattavaa Pearson–Tukey-menetelmässä on, että sen todennäköisyydet — 0,185, 0,630 ja 0,185 — ovat aina samat. Muuttuvat suureet

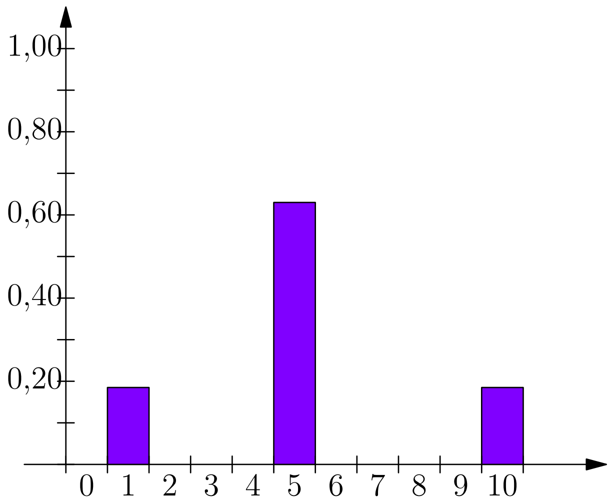
[5] Ad hoc tarkoittaa “tätä varten”. Se on hieno tapa sanoa, että meillä ei ole mitään hyvinperusteltua syytä toimia niin kuin toimimme.

ovat fraktiilien arvot. Toisin sanoen jakauman muoto on aina samankaltainen, mutta paikka vaihtelee.

4.3.4 Esimerkki (Pearson–Tukey-menetelmä leipuri Pullalle)**.** Leipuri Pullan esimerkissä 4.1.3 datapisteitä oli 82. Siten 0,05-fraktiili on suurin neljästä pienimmästä havainnosta. Tämä on 1, sillä pienimmät havainnot olivat $0, 0, 0, 1$. Vastaavasti neljä suurinta havaintoa olivat $10, 10, 10, 10$. Siten 0,95-fraktiili on 10. Mediaani taas on järjestyksessä 41. havainto. Tämä on 5. Siten Pearson–Tukey-aproksimaatio pullanmyyntitodennäköisyyksille on

$$\begin{aligned} p_1 &= 0{,}185, \\ p_5 &= 0{,}630, \\ p_{10} &= 0{,}185. \end{aligned}$$

Kuvallisesti Pearson–Tukey-todennäköisyysjakauma on siis (kuva on eri skaalalla kuin kolme edellistä kuvaa):



Jos esimerkin 4.1.3 leipuri Pulla arvottaa vaihtoehdot

$$a_i \quad = \quad \text{“Paistetaan aamulla } i \text{ pullaa”}$$

riskineutraalisti, eli odotusarvon mukaan, ja hän uskoo Pearson–Tukey-menetelmään, hänen arvotuksensa saadaan kaavasta

$$\begin{aligned} V(a_i) &= \Big(\min(i,1) - 0{,}2i\Big) \cdot 0{,}185 \\ &\quad + \Big(\min(i,5) - 0{,}2i\Big) \cdot 0{,}630 \\ &\quad + \Big(\min(i,10) - 0{,}2i\Big) \cdot 0{,}185. \end{aligned}$$

Jätämme harjoitustehtäväksi 4.1(c) laskea leipuri Pullan arvotukset tässä tilanteessa ja määrätä hänen optimaalinen valintansa. Huomautamme, että aluksi saattaa tuntua luontevalta valita päätösjoukoksi a_1, a_5, a_{10}. Tämä ei kuitenkaan ole välttämättä oikea tapa lähestyä ongelmaa. Joskus nimittäin on hyvä hyväksyä aina hieman tappiota, jotta saataisiin *keskimääräisesti* paras tulos.

4.3.5 Huomautus (Pearson–Tukey-menetelmän hyviä ja huonoja puolia)**.** Pearson–Tukey menetelmä on hyvä, sillä se

⊕ on helppo toteuttaa,

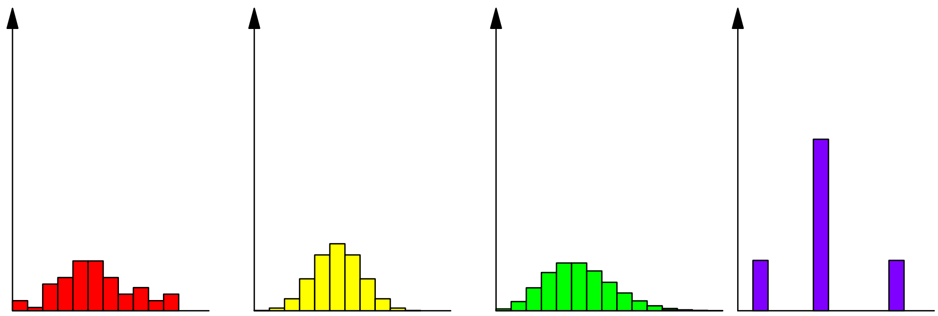
⊕ tuottaa pienen otosavaruuden ja soveltuu siten hyvin esimerkiksi päätöspuihin.

Pearson–Tukey-menetelmä on huono, sillä se

⊖ on erittäin karkea,

⊖ toimii huonosti epäsymmetrisille tai monihuippuisille jakaumille.

4.3.6 Esimerkki (Pullan eri estimaatit)**.** Esitämme vielä, vertailun vuoksi, saadut neljä jakauma-arviota rinnakkain samalla skaalalla:



4.4 Harjoitustehtäviä lukuun 4

4.1 Harjoitustehtävä. Tarkastelemme esimerkin 4.1.3 leipuri Pullaa. Kuinka monta pullaa tulee leipuri Pullan paistaa aamuisin, kun hän on riskineutraali, ja uskoo että (potentiaalisesti) myytyjen pullien lukumäärä lounastauolla on

(a) binomijakautunut parametrein $n = 10$ ja $p = 0{,}5$,

(b) Poisson-jakautunut parametrilla 5,

(c) Pearson–Tukey-jakautunut fraktiiliparametrein $1, 5, 10$?

4.2 Harjoitustehtävä. Tarkastelemme esimerkin 4.1.3 leipuri Pullaa. Oletamme, että leipuri Pullalla on vain lokakuun 2009 data käytettävissään. Oletamme lisäksi, että leipuri Pulla on riskineutraali päätöksenteossaan. Kuinka monta pullaa leipuri Pullan tulee paistaa aamuisin, kun hän estimoi pullanmyyntitodennäköisyydet

(a) suhteellisten frekvenssien menetelmällä,
(b) binomimallilla käyttäen suurimman uskottavuuden periaatetta,
(c) Poisson-mallilla käyttäen suurimman uskottavuuden periaatetta,
(d) Pearson–Tukey-menetelmällä?

4.3 Harjoitustehtävä. Satunnaismuuttuja X on *geometrisesti jakautunut* parametrilla θ, jos

$$\mathrm{P}(X=n) \quad = \quad (1-\theta)^n\theta,$$

$n = 0, 1, 2, \ldots$.

(a) Millaista tilannetta geometrisesti jakautunut satunnaismuuttuja kuvaa?
(b) Johda parametrin θ suurimman uskottavuuden estimaattori.

4.4 Harjoitustehtävä. Sinikka Sateen tulee päättää ottaako aamulla sateenvarjo mukaan vai ei. Hän on riskineutraali harmituksen suhteen. Jos hän ottaa sateenvarjon, eikä sada, häntä harmittaa 1 verran. Jos hän ei ota sateenvarjoa, ja sataa, häntä harmittaa 2 verran. Muissa tapauksissa häntä ei harmita ollenkaan.

Sinikka Sade on kerännyt dataa sadepäivistä (1 tarkoittaa “sataa”, 0 tarkoittaa “ei sada”):

Lokakuu 2009							
Vk	Ma	Ti	Ke	To	Pe	La	Su
40				1	0	0	0
41	0	0	1	1	1	1	1
42	1	0	0	0	1	1	1
43	1	0	0	0	0	0	0
44	0	0	0	0	0	1	1

Marraskuu 2009							
Vk	Ma	Ti	Ke	To	Pe	La	Su
45	0	0	1	0	0	0	1
46	1	1	0	0	0	0	0
47	1	1	1	0	0	1	1
48	0	0	0	0	0	0	0
49	0						

(a) Tulisiko Sinikka Sateen ottaa sateenvarjo aamuisin mukaan?
(b) Voisiko Sinikka Sade päättää paremmin ottaako sateenvarjo huomenna mukaan vai ei käyttäen tietoa tämän päivän säästä? Jos tänään sataa, tulisiko Sinikka Sateen ottaa huomenna sateenvarjo mukaan? Entä jos tänään ei sada?
Vihje: Kannattaa estimoida ehdolliset todennäköisyydet huomenna sataa/ei sada ehdolla tänään sataa/ei sada.

4.5 Harjoitustehtävä. Sijoittaja S.:llä on 10.000€ sijoitettavaa. Hän ei halua hajauttaa sijoitustaan, vaan haluaa sijoittaa joko Riskiin, Varmaan, tai Säästöön. Säästö antaa varmasti 1% tuoton. Hillosanomista sijoittaja S. on lukenut, että Varman ja Riskin odotetut tuotot (μ) ja tuottojen “volatiliteetit” (σ) ovat

$$\begin{aligned}\mu_{\text{Varma}} &= 3\%,\\ \sigma_{\text{Varma}} &= 10\%,\end{aligned}$$

$$
\begin{aligned}
\mu_{\text{Riski}} &= 5\%, \\
\sigma_{\text{Riski}} &= 20\%.
\end{aligned}
$$

Sijoittaja S. on antanut itselleen kertoa, että tämä tarkoittaa sitä, että tuotto on normaalisti jakautunut satunnaismuuttuja odotusarvolla μ ja varianssilla σ^2.

Mihin kohteeseen tulisi sijoittaja S.:n sijoittaa rahansa, kun hän on

(a) riskineutraali,
(b) hyödyn maksimoija hyötyfunktiolla $u(r) = \ln r$.

Käytä analyysissäsi laajennettua Pearson–Tukey-menetelmää, ja vertaa sitä "todelliseen" normaalijakaumaan.

4.6 Harjoitustehtävä. Esimerkin 4.1.3 data on generoitu seuraavasti: ensiksi pyhäpäivät, viikonloput ja aatot on poistettu. Sitten perjantaisin myydyt pullat ovat riippumattomia Poisson(2)-jakautuneita satunnaismuuttujia ja maanantaista torstaihin myydyt pullat ovat riippumattomia Poisson(6)-jakautuneita satunnaismuuttujia. (Aikariippuvutta ei siis ollut, mutta epähomogeenisuutta oli. Toisin sanoen havainnot olivat riippumattomia, mutteivät samankaltaisia.)

(a) Piirrä myytyjen pullien todennäköisyysjakaumat (ma–to ja pe).
(b) Piirrä "umpimähkään" valittuna päivänä myytyjen pullien todennäköisyysjakauma.

4.7 Harjoitustehtävä. Nyt, kun tiedät harjoitustehtävässä 4.6 kerrotun totuuden, niin kuinka monta pullaan tulisi esimerkin 4.1.3 leipuri Pullan paistaa amulla, kun hän on

(a) optimisti,
(b) katumuksen kaihtaja,
(c) riskineutraali,
(d) riskinkaihtaja hyötyfunktiolla $u(r) = \ln(1 + r)$?

Luku 5

Hyötyteoriaa

> Michael: “I don't know anyone who could get through the day without two or three juicy rationalizations. They're more important than sex.”
> Sam Weber: “Ah, come on. Nothing's more important than sex.”
> Michael: “Oh yeah? Ever gone a week without a rationalization?”
>
> – The Big Chill

> The correctness of a decision can't be judged from the outcome. Nevertheless, that's how people assess it. A good decision is one that's optimal at the time it's made, when the future is by definition unknown. Thus, correct decisions are often unsuccessful, and vice versa. – Howard Marks

> Almost anything can be attacked as a failure, but almost anything can be defended as not a significant failure. Politicians do not appreciate the significance of 'significant'. – Sir Humphrey Appleby

5.1 Suhtautuminen riskiin

5.1.1 Huomautus (Riski on epävarmuutta)**.** Puhekielessä riskillä tarkoitetaan usein riskiä epäonnistua. Teoreetikoille riski tarkoittaa ilkikurisesti vain satunnaisuutta eli epävarmuutta. Siten on myös olemassa riski onnistua yli odotusten, tai vaikkapa riski voittaa lotossa.

Riskiprofiili ja ankkuripiste

Seuraavalla esimerkillä yritämme valaista henkilön suhtaumista riskiin suhteessa henkilön omaan lähtötilanteeseen eli ankkuripisteeseen.

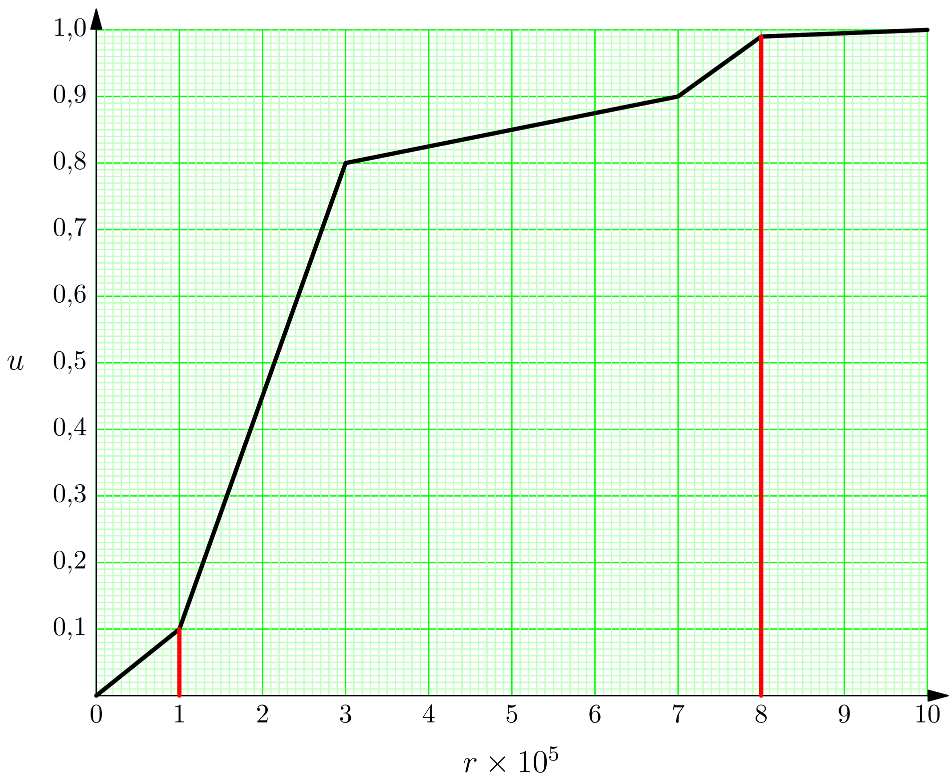
5.1.2 Esimerkki (Pikkurillin arvo)**.** Sadistinen miljonääri tarjoaa Oiva Opiskelijalle ja Tarmo Toimitusjohtajalle satunnaiskoetta, jossa he

- menettävät todennäköisyydellä 0,5 vasemman käden pikkurillinsä,
- saavat todennäköisyydellä 0,5 200.000€.

Oiva ottaa tarjouksen vastaan ja Tarmo ei ota. Onko tarinan opetus, että Tarmo arvostaa pikkurilliään enemmän kuin Oiva?

Tarmo Toimitusjohtaja voi toki arvostaa pikkurilliään enemmän kuin Oiva Opiskelija. Toisaalta on selvää, että sekä Oiva Opiskelija että Tarmo Toimitusjohtaja eivät voi olla riskineutraaleja ja arvostaa pikkurillejään yhtä paljon. On kuitenkin mahdollista, että he arvostavat yhtä paljon pikkurillejään ja että heillä on sama *riskiprofiili* eli he ovat molemmat hyödyn maksimoijia samalla hyötyfunktiolla u. Ero päätöksentekoon tulee siitä, että heillä on eri lähtövarallisuus eli ankkurointipiste r. Toisin sanoen he katsovat samaa hyötyfunktiota u eri kohdista r.

Olkoon molempien hyötyfunktio u seuraavan kuvan mukainen:



Oletamme, että molemmat arvostavat pikkurillejään 100.000€ verran. Tämä tarkoittaa sitä, että jos he olisivat riskineutraaleja, niin he molemmat suostuisivat innolla sadistisen miljonäärin tarjoukseen. Olkoon sitten Oiva Opiskelijan varallisuus 100.000€ ja Tarmo Toimitusjohtajan varallisuus 800.000€. Nämä varallisuudet on merkitty hyötyfunktion kuvaajaan punaisilla pisteviivoilla.

Nyt Oiva Opiskelijan vaihtoehdot ja niitä vastaavat hyödyt ovat:

1. Ei oteta tarjousta vastaan. Tällöin hyöty

$$u(100.000) \;=\; 0{,}1.$$

2. Otetaan tarjous vastaan. Tällöin keskimääräinen hyöty on

$$\begin{aligned}
&u(300.000)\cdot 0{,}5 + u(0)\cdot 0{,}5\\
&= 0{,}8\cdot 0{,}5 + 0\cdot 0{,}5\\
&= 0{,}4.
\end{aligned}$$

Oiva Opiskelija siis hyväksyy tarjouksen. Tarmo Toimitusjohtajan vaihtoehdot ja niitä vastaavat hyödyt ovat:

1. Ei oteta tarjousta vastaan. Tällöin hyöty

$$u(800.000) \;=\; 0{,}99.$$

2. Otetaan tarjous vastaan. Tällöin keskimääräinen hyöty on

$$\begin{aligned}
&u(1.000.000)\cdot 0{,}5 + u(700.000)\cdot 0{,}5\\
&= 1\cdot 0{,}5 + 0{,}9\cdot 0{,}5\\
&= 0{,}95.
\end{aligned}$$

Tarmo Toimitusjohtaja siis hylkää tarjouksen.

5.1.3 Esimerkki (Agitointia ja propagandaa)**.** Olettakaamme, että kansalaiset ovat riskiä kaihtavia samalla hyöytyfunktiolla u. Yhteisesti ollaan päätetty, että kansallisvarallisuus pitää jakaa niin, että kokonaishyöty

$$U \;=\; \sum_{j\in J} u(r_j)$$

maksimoituu. Yllä J on kansalaisten joukko, ja r_j on kansalaiselle j jaettu osuus kokonaisvarallisuudesta. Kysymys on siis, miten jakaa kokonaisvarallisuus

$$\sum_{j\in J} r_j = 100\%$$

osiin r_j, $j\in J$, kansalaisten kesken siten, että kokonaishyöty U maksimoituu.

Tarkastelemme ongelmaa tilanteessa, jossa kansakunta koostuu vain kahdesta kansalaisesta. Yleisen tilanteen tarkastelun jätämme harjoitustehtäväksi 5.2.

Teemme tavallisen normalisoinnin: $u(0) = 0$ ja $u(1) = 1$. Tällöin tilanne, jossa ensimmäinen kansalainen saa osuuden p kokonaisvarallisuudesta, johtaa kokonaisyhyötyyn

$$U(p) = u(p) + u(1-p).$$

Funktion maksimi löytyy tunnetusti joko välin päätepisteistä, tai sitten derivaatan nollakohdista. Välin päätepisteissä saamme

$$\begin{aligned} U(0) &= u(0)+u(1) &= 0+1 &= 1, \\ U(1) &= u(1)+u(0) &= 1+0 &= 1. \end{aligned}$$

Mitä derivaatan nollakohtiin tulee, niin

$$\frac{\mathrm{d}U}{\mathrm{d}p}(p) = \frac{\mathrm{d}u}{\mathrm{d}p}(p) - \frac{\mathrm{d}u}{\mathrm{d}p}(1-p)$$

on nolla, jos

$$\frac{\mathrm{d}u}{\mathrm{d}p}(p) = \frac{\mathrm{d}u}{\mathrm{d}p}(1-p).$$

Tämä taas tapahtuu, jos $p = 1/2$. Tämä tarkoittaa varallisuuden tasajakoa!

Jos siis hyväksymme kokonaishyödyn ja yleisen riskin välttämisen moraaliseksi ohjenuoraksi, niin varallisuuden tasajako on moraalisesti oikea valinta. Toki tämä esimerkki tarkasteli vain staattista tilannetta: dynaamisessa tilanteessa asia ei ole aivan näin yksinkertainen ☺.

5.2 Arpajaiset, preferenssit ja hyödyt

Hyötyteoriaa ja rationaalista päätöksentekoa esitetään usein niin sanottujen arpajaisten avulla. Itse asiassa arpajaiset eivät ole mitään sen kummallisempaa kuin satunnaismuuttujia. Siten tämä uusi käsite on periaatteessa tarpeeton, mutta otamme sen kuitenkin käyttöön “perinteen velvoittamina”.

Yksinkertaiset Arpajaiset

5.2.1 Määritelmä (Arpajaiset, yksinkertaiset arpajaiset)**.** *Arpajaisilla*, tai *yksinkertaisilla arpajaisilla*, tarkoitetaan diskreettiä satunnaismuuttujaa L, joka saa arvot $r_1, \ldots, r_n$ todennäköisyyksin $p_1, \ldots, p_n$. Arpajaisille L käytetään myös merkintää

$$L = L(p_1, r_1; p_2, r_2; \ldots; p_n, r_n),$$

missä todennäköisyydet p_j ja niitä vastaavat palkkiot r_j on kirjoitettu eksplisiittisesti näkyviin.

5.2.2 Huomautus. Arpajaiset

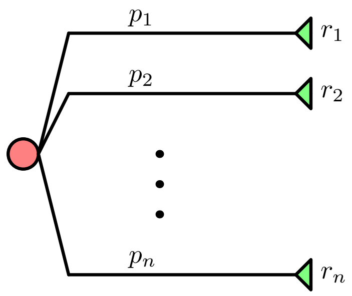
$$L = L(p_1, r_1; p_2, r_2; \ldots; p_n, r_n).$$

voidaan tulkita tilanteeksi, jossa tarjotaan

- palkkiota r_1 todennäköisyydellä p_1,
- palkkiota r_2 todennäköisyydellä p_2,

  $\vdots$
- palkkiota r_n todennäköisyydellä p_n.

Kuvallisesti arpajaiset $L = L(p_1, r_1; p_2, r_2; \ldots; p_n, r_n)$ tarkoittaa siis lehtiin päättyvää sattumasolmua



5.2.3 Huomautus (Arpajaiset ja päätösmatriisit)**.** Luvun 2 kielellä arpajaiset L_i voisi tarkoittaa valinnan a_i seurausta. Tällöin

$$L_i \quad = \quad L(p_1, r_{i1}; \ldots; p_n, r_{im}),$$

jos $J = \{1, \ldots, m\}$.

5.2.4 Esimerkki. Pekka ja Liisa heittävät reilua kolikkoa. Jos tulee klaava, niin Pekka saa Liisalta euron, ja jos tulee kruuna, niin Liisa saa Pekalta euron. Liisan kannalta kyseessä on arpajaiset $L(0{,}5, 1; 0{,}5, -1)$. Pekan kannalta tilanne on sama.

Yhdistetyt arpajaiset

Seuraava märitelmä on luonteeltaan *rekursiivinen*[1]:

5.2.5 Määritelmä (Yhdistetyt arpajaiset)**.** *Yhdistetyt arpajaiset* ovat satunnaismuuttujia tai -kokeita, jossa palkkiot voivat olla arpajaisia.

5.2.6 Huomautus. Esimerkiksi merkintä

$$L = L(p_1,\ L'(p_1', r_1'; p_2', r_2'; \ldots; p_{n'}', r_{n'}');\ p_2, r_2; \ldots; p_n, r_n)$$

tarkoittaa yhdistettyjä arpajaisia, jossa tarjotaan

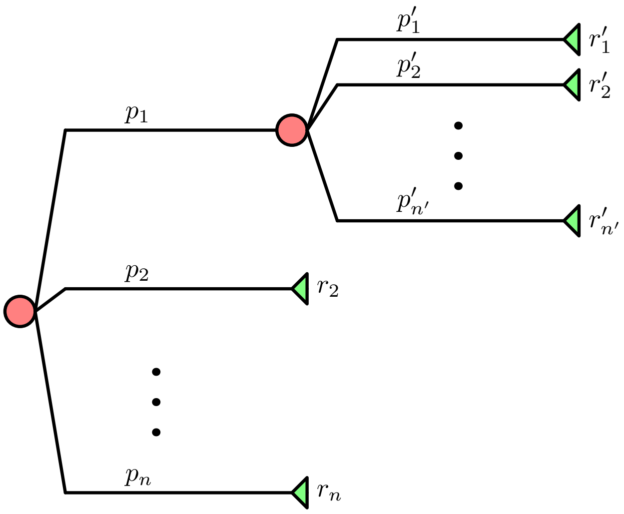
- uusia arpajaisia $L' = L'(p_1', r_1'; p_2', r_2'; \ldots; p_{n'}', r_{n'}')$ todennäköisyydellä p_1,
- palkkiota r_2 todennäköisyydellä p_2,

[1] Sanakirjamääitelmä rekursiiviselle on: “rekursiivinen: ks. rekursiivinen” ☺.

$\vdots$

- palkkiota r_n todennäköisyydellä p_n.

Kuvallisesti edellä määritellyt yhditetyt arpajaiset L tarkoittaa siis sattumapuuta



5.2.7 Esimerkki. Pekka ja Liisa heittävät reilua kolikkoa. Jos tulee kruuna, niin Liisa saa Pekalta euron. Jos taas tulee klaava, niin Pekka herrasmiehenä antaa Liisalle uuden mahdollisuuden. Jos taas tulee klaava, niin Pekka saa Liisalta euron. Liisan kannalta kyseessä on yhdistetyt arpajaiset $L(0{,}5, 1; 0{,}5, L(0{,}5, 1; 0{,}5-1))$. Pekan kannalta kyseessä on yhdistetyt arpajaistet $L(0{,}5, L(0{,}5, 1; 0{,}5, -1); 0{,}5, -1)$.

5.2.8 Määritelmä (Lyhennysmerkintä arpajaisille)**.** Jos L_1 ja L_2 ovat arpajaisia ja $q \in (0, 1)$, niin merkintä

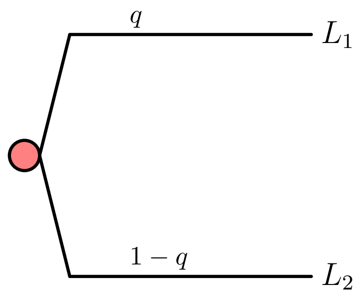
$$L \quad = \quad qL_1 + (1-q)L_2$$

tarkoittaa yhdistettyjä arpajaisia L, joissa ensin valitaan joko arpajaiset L_1 todennäköisyydellä q tai arpajaiset L_2 todennäköisyydellä $1-q$, ja sitten suoritetaan valitut arpajaiset: joko L_1 jai L_2.

5.2.9 Esimerkki. Käyttämällä lyhennysmerkintää 5.2.8 yhdistetyille arpajaisille huomaamme, että

$$qL_1 + (1-q)L_2 \quad = \quad L(q, L_1; 1-q, L_2).$$

Kuvallisesti $L = qL_1 + (1-q)L_2$ tarkoittaa siis seuraavaa sattumapuun haaraa:



5.2.10 Esimerkki. Esimerkin 5.2.7 Liisan tilanne

$$L(0{,}5, 1; 0{,}5, L(0{,}5, 1; 0{,}5 - 1))$$

voidaan siis Määritelmän 5.2.8 avulla kirjoittaa myös muodossa

$$0{,}5L(1, 1) + 0{,}5L(0{,}5, 1; 0{,}5, -1),$$

tai jopa muodossa

$$0{,}5L(1, 1) + 0{,}5\Big(0{,}5L(1, 1) + 0{,}5L(1, -1)\Big).$$

Jos vielä otamme käyttöön lyhennysmerkinnän $L(r) = L(1, r)$ “varmoille” arpajaisille, niin päädymme muotoon

$$0{,}5L(1) + 0{,}5\Big(0{,}5L(1) + 0{,}5L(1) + 0{,}5L(-1)\Big).$$

Lopuksi, jos hyväksymme kaavat[2]

$$\begin{aligned} p(q_1L_1 + q_2L_2) &= pq_1L_1 + pq_2L_2, \\ p_1L + p_2L &= (p_1 + p_2)L, \end{aligned}$$

niin Esimerkin 5.2.7 Liisan tilanne voidaan kuvata kaavalla

$$0{,}75L(1) + 0{,}25L(-1).$$

Tämä kaava sanoo siis yksinkertaisesti, että Liisa saa yhden euron (Pekalta) todennäköisyydellä 0,75 ja menettää yhden euron (Pekalle) todennäköisyydellä 0,25.

[2]Rationaalinen päätöksentekijä hyväksynee nämä kaavat. Tästä puhumme lisää myöhemmin. Sen sijaan kaavaa $pL(q) = pqL(1)$ ei tule hyväksyä.

Preferenssit ja indifferenssit

5.2.11 Määritelmä (Preferenssi ja indifferenssi)**.** Perferenssi- ja indifferenssimerkinnät $\prec$, $\preceq$, $\sim$,$\succ$ ja $\succeq$ tarkoittavat:

(i) Mikäli päätöksentekijä valitsee mieluummin arpajaiset L_1 kuin arpajaiset L_2, niin sanomme että päätöksentekijä *preferoi* arpajaisia L_1 yli arpajaisten L_2. Tällöin merkitsemme $L \prec L'$.

(ii) Mikäli päätöksentekijä on valitsee yhtä mielellään arpajaiset L_1 tai L_2, niin sanomme että päätöksentekijä on *indifferentti* arpajaisten L_1 ja L_2 välillä. Tällöin merkitsemme $L_1 \sim L_2$.

(iii) Merkintä $L_1 \preceq L_2$ tarkoittaa $(L_1 \prec L_2) \vee (L_1 \sim L_2)$.

(iv) Merkintä $L_1 \succ L_2$ tarkoittaa $\neg(L_1 \preceq L_2)$.

(v) Merkintä $L_1 \succeq L_2$ tarkoittaa $\neg(L_1 \prec L_2)$.

5.2.12 Huomautus (Arpajaisperefenssit, palkkiomatriisit ja arvofunktiot)**.** Luvun 2 kielellä arpajaiset L_i vastaavat päätöksen a_i mahdollisia seurauksia $\{r_{ij}; j \in J\}$. Siten päätössääntö $V : A \to \bar{\mathbb{R}}$ vastaa preferenssejä arpajaisille $\{L_i; i \in I\} = \{a_i; i \in I\}$. Preferenssit tulevat ekvivalenssista

$$L_{i_1} \prec L_{i_2} \quad \text{jos ja vain jos} \quad V(a_{i_2}) < V(a_{i_2}).$$

Preferenssit hyödyistä

Jokaista hyötyfunktiota $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ voidaan panna vastaamaan arpajaispreferenssit ja - indifferenssit seuraavan määritelmän antamalla tavalla:

5.2.13 Määritelmä (Preferenssit hyötyfunktiosta)**.** Olkoon $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ hyötyfunktio. Päätöksentekijä preferoi arpajaisia L_2 yli arpajaisten L_1, eli $L_1 \prec L_2$, jos

$$\mathrm{E}\,(u(L_1)) \quad < \quad \mathrm{E}\,(u(L_2))$$

Samoin päätöksentekijä on indifferentti arpajaisten L_1 ja L_2 välillä, eli $L_1 \sim L_2$, jos

$$\mathrm{E}\,(u(L_1)) \quad = \quad \mathrm{E}\,(u(L_2))$$

5.2.14 Esimerkki. Pekka Päätöksentekijällä on logaritminen hyötyfunktio $u(r) = \ln(1+r)$. Hänelle tarjotaan arpajaisia

1. $L_1 = L(1, 1.000€)$, eli varmaa palkkiota 1.000€, tai
2. $L_2 = L(0{,}5,\ 2.000€\ ;\ 0{,}5,\ 0€)$, eli 50%:n mahdollisuutta saada palkkio 2.00€ ja 50:n prosentin mahdollisuutta saada palkkio 0€ .

Kummat arpajaiset Pekka Päätöksentekijä valitsee?

Pekka Päätöksentekijän hyödyt arpajaisilla L_1 ja L_2 ovat

$$\begin{aligned}
\mathrm{E}(u(L_1)) &= u(1.000) \\
&= \ln 1.001 \\
&= 6{,}9088, \\
\mathrm{E}(u(L_2)) &= u(2.000)\cdot 0{,}5 + u(0)\cdot 0{,}5 \\
&= \ln 2.001\cdot 0{,}5 + \ln 1\cdot 0{,}5 \\
&= 7{,}6014\cdot 0{,}5 + 0\cdot 0{,}5 \\
&= 3{,}8007.
\end{aligned}$$

Siis $\mathrm{E}(u(L_2)) < \mathrm{E}(u(L_1))$, eli $L_2 \prec L_1$, eli Pekka Päätöksentekijä valitsee "varmat" arpajaiset L_1.

Hyödyt ja preferenssit kyselyistä

Esitämme nyt pitkähkön esimerkin kautta miten arpajaispreferenssit voidaan rakentaa kyselyiden avulla ja miten niistä rakentuu hyötyfunktio u, tai ainakin osa siitä.

5.2.15 Esimerkki (Preferenssit kyselemällä)**.** Päätöksentekijä halua laittaa seuraavat arpajaiset järjestykseen:

$$\begin{aligned}
L_1 &= L(1, 10.000), \\
L_2 &= L(0{,}50, 30.000; 0{,}50, 0), \\
L_3 &= L(1, 0), \\
L_4 &= L(0{,}02, -10.000, 0{,}98, 500).
\end{aligned}$$

Aloitamme havaitsemalla suurimman mahdollisen palkkion $r_+ = 30.000$ ja pienimmän mahdollisen palkkion $r_- = -10.000$. Muut mahdolliset palkkiot ovat (laskevassa) suuruusjärjestyksessä $r_1 = 10.000$, $r_2 = 500$ ja $r_3 = 0$.

Päätöksentekijältä kysytään "subjektiivista" todennäköisyyttä q_j, jolla vallitsisi indifferenssi

$$L(1, r_j) \quad\sim\quad q_j L(1, 30.000) + (1-q_j) L(1, -10.000).$$

Oletamme, että palkkiolle $r_1 = 10.000$ päätöksentekijämme valitsee kysyttäessä $q_1 = 0{,}90$. Hän on siis valinnut indifferenssin

$$L(1, 10.000) \quad\sim\quad 0{,}90 L(1, 30.000) + 0{,}10 L(1, -10.000). \tag{5.2.16}$$

Palkiolle $r_2 = 500$ hän on kysyttäessä valinnut indifferenssin

$$(5.2.17) \qquad L(1, 500) \;\sim\; 0{,}62L(1, 30.000) + 0{,}38L(1, -10.000).$$

Siis $q_2 = 0{,}62$. Kysyttäessä palkkiosta $r_3 = 0$ hän valitsee todennäköisyyden $q_3 = 0{,}60$ eli indifferenssin

$$(5.2.18) \qquad L(1, 0) \;\sim\; 0{,}60L(1, 30.000) + 0{,}40L(1, -10.000).$$

Käyttämällä kyselyistä saatuja indifferenssejä päätöksentekijä voi rakentaa arpajaiset L_1', L_2', L_3' ja L_4' siten, että $L_1 \sim L_1'$, $L_2 \sim L_2'$, $L_3 \sim L_3'$ ja $L_4 \sim L_4'$, ja lisäksi arpajaisissa L_1', L_2', L_3 ja L_4' ainoat palkkiomahdollisuudet ovat $r_+ = 30.000$ ja $r_- = -10.000$. Nämä arpajaiset taas on helppo laittaa järjestykseen kyselyihin vastattujen "subjektiivisten" todennäköisyyksien q_1, q_2 ja q_3 avulla seuraavalla tavalla.

Indifferenssisä (5.2.16) näemme välittömästi, että $L_1 \sim L_1'$, missä

$$L_1' \;=\; 0{,}90L(1, 30.000) + 0{,}10L(1, -10.000).$$

Indifferenssistä (5.2.18) näemme, että $L_2 \sim L_2''$, missä

$$L_2'' \;=\; 0{,}5L(1, 30.000) + 0{,}5L(0{,}60, 30.000; 0{,}40, -10.000).$$

L_2'' ovat siis yhdistetyt arpajaiset, missä

- $r_+ = 30.000$ saadaan todennäköisyydellä $0{,}50 + 0{,}50 \cdot 0{,}60 = 0{,}80$,
- $r_- = -10.000$ saadaan todennäköisyydellä $0{,}50 \cdot 0{,}40 = 0{,}20$.

Siten $L_2 \sim L_2'' \sim L_2'$, missä

$$L_2' \;=\; 0{,}80L(1, 30.000) + 0{,}20L(1, -10.000).$$

Indifferenssistä (5.2.18) näemme, että $L_3 \sim L_3'$, missä

$$L_3' \;=\; 0{,}60L(1, 30.000) + 0{,}40L(1, -10.000).$$

Indifferenssistä (5.2.17) seuraa, että $L_4 \sim L_4''$, missä

$$L_4'' \;=\; 0{,}02L(1, -10.000) + 0{,}98L(0{,}62, 30.000; 0{,}38, -10.000).$$

L_4'' ovat siis yhdistetyt arpajaiset, missä

- $r_+ = 30.000$ saadaan todennäköisyydellä $0{,}98 \cdot 0{,}62 = 0{,}6076$,
- $r_- = -10.000$ saadaan todennäköisyydellä $0{,}02 + 0{,}98 \cdot 0{,}38 = 0{,}3924$.

Siten $L_4 \sim L_4'' \sim L_4'$, missä

$$L_4' \quad = \quad 0{,}6076L(1, 30.000) + 0{,}3924L(1, -10.000).$$

On siis saatu arpajaiset

$$\begin{aligned}
L_1' &= 0{,}90L(1, 30.000) + 0{,}10L(1, -10.000), \\
L_2' &= 0{,}80L(1, 30.000) + 0{,}20L(1, -10.000), \\
L_3' &= 0{,}60L(1, 30.000) + 0{,}40L(1, -10.000), \\
L_4' &= 0{,}6076L(1, 30.000) + 0{,}3924L(1, -10.000).
\end{aligned}$$

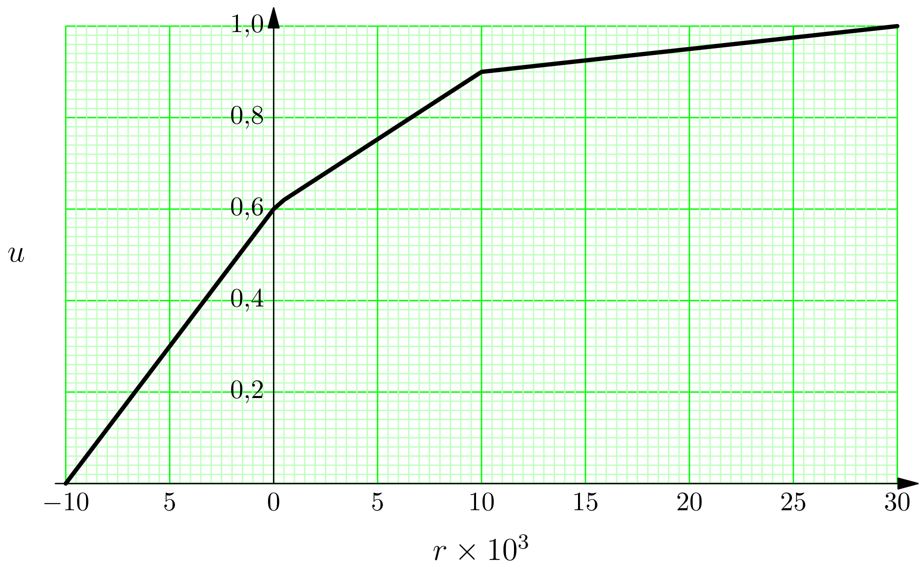
Nämä arpajaiset on ilmeisen helppo järjestää paremmuusjärjestykseen: $L_3' \prec L_4' \prec L_2' \prec L_1'$. Koska pätee indifferenssit $L_i' \sim L_i$, $i = 1, 2, 3, 4$, niin sama paremmuusjärjestys pätee myös alkuperäisille arpajaisille: $L_3 \prec L_4 \prec L_2 \prec L_1$. Ongelma on ratkaistu!

5.2.19 Esimerkki (Hyödyt kyselemällä)**.** Esimerkissä 5.2.15 preferenssit saatiin tietyllä kyselymenetelmällä. Tämä kyselymenetelmä paljastaa myös päätöksentekijän hyötyfunktion, tai ainakin osan siitä. Nimittäin voimme tulkita "subjektiiviset" todennäköisyydet q_j palkkoita r_j vastaaviksi hyödyiksi ja rakentaa päätöksentekijän hyötyfunktion asettamalla

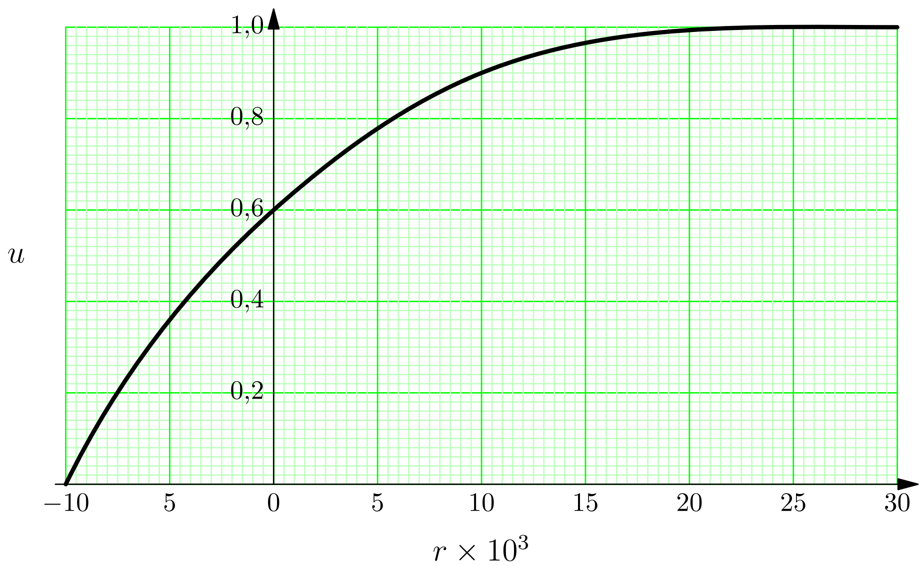
$$\begin{array}{rclclcl}
u(30.000) &=& u(r_+) &=& 1, && \\
u(10.000) &=& u(r_1) &=& q_1 &=& 0{,}90, \\
u(500) &=& u(r_2) &=& q_2 &=& 0{,}62, \\
u(0) &=& u(r_3) &=& q_3 &=& 0{,}60, \\
u(-10.000) &=& u(r_-) &=& 0. &&
\end{array}$$

Tämä tietysti antaa edellä esitettyä "puutapaa" helpomman tavan asettaa arpajaiset L_1, L_2, L_3 ja L_4 preferenssijärjestykseen: ensin määrätään hyötyfunktion $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ arvo $u(r)$ kaikissa tarvittavissa palkkiopisteissä r, ja sitten asetetaan arpajaiset L_1, L_2, L_3 ja L_4 preferenssijärjestykseen odotetun hyödyn mukaan.

Seuraavassa hahmotelma päätöksentekijän hyötyfunktiosta:



Edellä hahmotellussa kuvassa hyötyfunktio u on laskettu pisteiden $r_- = -10.000$, $r_3 = 0$, $r_2 = 500$, $r_1 = 10.000$ ja $r_+ = 30.000$ ulkopuolella lineaarsesti interpoloiden. Tämä tekee hyötyfunktioon teräviä kulmia, mikä saattaa olla käytännön tulkinnan kannalta kyseenalaista. Sileämpi sovitus annettuihin pisteisiin $r_- = -10.000$, $r_3 = 0$, $r_2 = 500$, $r_1 = 10.000$ ja $r_+ = 30.000$ saadaan aikaan esimerkiksi käyttämällä piirrosohjelmissa suosittuja (kasvaviksi ehdollistettuja) 3. asteen Bézierin splinejä (muitakin menetelmiä toki on):



Esimerkkien 5.2.15 ja 5.2.19 ongelma ja sen ratkaisualgoritmi voidaan kuvata yleisesti seuraavasti:

5.2.20 Määritelmä (Kyselyalgoritmi preferensseille ja hyödyille).
Ongelma: On annettu arpajaiset $L_1, L_2, \ldots, L_m$, jotka pitää laittaa preferenssijärjestykseen.
Ratkaisu:

(i) Järjestä kaikki mahdolliset palkkiot laskevaan suuruusjärjestykseen $r_+ > r_1 > \cdots > r_n > r_-$ (mahdollisia palkkioita on siis $n+2$ kappaletta).
(ii) Jokaiselle palkkiolle rji, $j = +, 1, \ldots, n, -$, *kysy päätöksentekijältä* "subjektiivista" todennäköisyyttä q_j, jolla hänelle pätee indifferenssi

$$L(1, r_j) \quad \sim \quad q_j L(1, r_+) + (1 - q_j) L(1, r_-).$$

(Väkisinkin muuten palkkiota r_+ vastaa $q_+ = 1$ ja palkkiota r_- vastaa $q_- = 0$).
(iii) Määritä päätöksentekijän hyötyfunktio $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ pisteissä r_j kaavalla $u(r_j) = q_j$. (Voit laajentaa hyötyfunktion $u(r)$ muihin pisteisiin lineaarisesella interpoloinnilla tai jollakin muulla menetelmällä, jos haluat. Tämä ei ole kuitenkaan ongelman kannalta tarpeellista.)
(iv) Määrää arpajaisten $L_1, L_2, \ldots, L_m$ järjestys hyötyfunktion $u(r)$ avulla asettamalla kaikille $k_1, k_2 = 1, \ldots, m$

$$L_{k_1} \prec L_{k_2} \quad \text{jos ja vain jos} \quad \mathrm{E}\left(u(L_{k_1})\right) < \mathrm{E}\left(u(L_{k_2})\right).$$

ja

$$L_{k_1} \sim L_{k_2} \quad \text{jos ja vain jos} \quad \mathrm{E}\left(u(L_{k_1})\right) = \mathrm{E}\left(u(L_{k_2})\right).$$

5.2.21 Huomautus (Hyötyfunktion arvojoukko). Algoritmi 5.2.20 rakentaa hyötyfunktion, jonka arvot ovat välillä $[0, 1]$: huonoin mahdollinen palkkio r_- saa hyödyn $u(r_-) = 0$ ja paras mahdollinen palkkio r_+ saa hyödyn $u(r_+) = 1$. Itse asiassa, koska hyöytfunktio on aina kasvava, edellinen algoritmi tuotaa aina hyötyfunktion, jonka voi tulkita todennäköisyysjakauman kertymäfunktioksi.

Aikaisemmin olemme esittäneet hyötyfunktioita, joissa palkkiot eivät ole rajoittuneet välille $[0, 1]$. Tyypillisin esimerkki on ollut logaritminen hyötyfunktio $u(r) = \ln(1 + r)$, joka saa siis arvoja väliltä $[0, \infty)$, jos palkkiot r ovat positiivisia. Tässä ei ole kuitenkaan mitään ristiriitaa. Nimittäin hyötyfunktioiden arvot sinänsä eivät ole mielenkiintoisia, vaan ainoastaan niiden arvojen järjestykset. Esimerkiksi positiivisesti *affiinimuunnettu*[3] hyötyfunktio

$$u(r) \quad = \quad \alpha u_0(r) + \beta,$$

[3] Affiinimuunnos funktiosta $f(x)$ on $\alpha f(x) + \beta$, missä $\alpha, \beta \in \mathbb{R}$. Affiinimuunnoksessa siis muunnetaan mittakaavaa, eli skaalaa, (α) ja suuntaa (α positiivinen tai negatiivinen), sekä paikkaa, eli lokaatiota, (β).

$\alpha > 0$, $\beta \in \mathbb{R}$, saa aikaan täsmälleen samat preferenssit kuin alkuperäinen hyötyfunktio $u_0(r)$. Toisin sanoen:

$$\mathrm{E}\big(u_0(L_1)\big) \quad\leq\quad \mathrm{E}\big(u_0(L_2)\big)$$

jos ja vain jos

$$\mathrm{E}\big(u(L_1)\big) \quad\leq\quad \mathrm{E}\big(u(L_2)\big).$$

Erityisesti, jos suurin mahdollinen palkkio on r_+ ja pienin mahdolinen palkkio on r_-, niin mikä tahansa hyötyfunktio $u_0(r)$ voidaan normeerata arvovälille $[0,1]$ muunnoksella

$$u(r) \quad=\quad \frac{u_0(r)-u_0(r_-)}{u_0(r_+)-u_0(r_-)}.$$

5.3 Von Neumann–Morgerstern -aksioomat

Vuonna 1944 Von Neumann ja Morgenstern muotoilivat *rationaalisen päätöksenteon* aksioomat.

Lähtökohtana on tietysti, että kaikki mahdolliset palkkiot voidaan esittää reaalilukuina, jolloin ne toteuttavat kaikki reaalilukujen aksioomat, kuten vaikkapa sen, että jokaisesta reaalilukuparista r_1, r_2 voidaan sanoa päteekö $r_1 < r_2$, $r_2 < r_1$ vai $r_1 = r_2$.

5.3.1 Määritelmä (Von Neumann–Morgenstern -aksioomat)**.** Päätösentekijän preferenssit ovat *rationaalisia*, jos ne toteuttavat seuraavat *Von Neumann–Morgenstern -aksioomat*:

(i) Kaikille arpajaisille L_1 ja L_2 pätee täsmälleen yksi relaatioista

  (a) $L_1 \prec L_2$,
  (b) $L_2 \prec L_1$,
  (c) $L_1 \sim L_2$.

  Lisäksi kaikille arpajaisille L_1, L_2 ja L_3 pätee:

$$\text{jos} \quad L_1 \prec L_2 \text{ ja } L_2 \prec L_3, \quad \text{niin} \quad L_1 \prec L_3.$$

  Vielä lisäksi kaikille arpajaisille L_1, L_2 ja L_3 pätee:

$$\text{jos} \quad L_1 \sim L_2 \text{ ja } L_2 \sim L_3, \quad \text{niin} \quad L_1 \sim L_3.$$

(ii) Jos $L_1 \sim L_2$, niin kaikille todennäköisyyksille $0 \leq q \leq 1$ ja kaikille arpajaisille L pätee

$$qL_1 + (1-q)L \quad\sim\quad qL_2 + (1-q)L.$$

(iii) Kaikille arpajaisille L_1, L_2 ja L_3, joille $L_1 \prec L_2$ ja $L_2 \prec L_3$ on olemassa jokin todennäköisyys $0 \leq q \leq 1$ siten, että

$$L_2 \quad \sim \quad qL_1 + (1-q)L_3.$$

5.3.2 Huomautus (Von Neumann–Morgenstern -aksioomista)**.** (i) *Järjestysaksiooma* 5.3.1(i) sanoo, että kaikkia arpajaisia voidaan verrata toisiinsa ja että $\prec$ on *aito järjestysrelaatio* kuten esimerkiksi relaatio $<$, ja että $\sim$ on *ekvivalenssirelaatio* kuten esimerkiksi relaatio $=$. Erityisesti järjestysaksiooma tekee merkinnöistä $L_1 \prec L_2 \prec L_3$ ja $L_1 \sim L_2 \sim L_3$ mielekkäitä.

(ii) *Riippumattomuusaksioomaa* 5.3.1(ii) tarvitaan, jos halutaan samaistaa arpajaiset ja yhdistetyt arpajaiset.

(iii) *Jatkuvuusaksioomaa* 5.3.1(iii) tarvitaan, jotta jokaista epävarmuutta sisältävää arpajaista saataisiin vastaamaan jokin varma palkkio.

Von Neumann–Morgenstern -aksioomat ovat yhtäpitäviä hyötyyn perustuvan päätössäännön kanssa. Itse asiassa suuri tulos on:

5.3.3 Lause (Von Neumann–Morgenstern -hyötylause)**.** (i) *Jos päätöksentekijä noudattaa Von Neumann–Morgenstern -aksioomia 5.3.1, on olemassa jokin sellainen hyötyfunktio $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$, että päätöksentekijä preferoi arpajaisia L_2 yli arpajaisten L_1 jos ja vain jos* $\mathrm{E}(u(L_1)) < \mathrm{E}(u(L_2))$.

(ii) *Kääntäen, jos päätöksentekijä arvottaa arpajaiset hyötysäännöllä jonkin hyötyfunktion $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ mukaan, noudattaa päätöksentekijä automaattisesti Von Neumann–Morgenstern -aksioomia 5.3.1.*

Todistus. Harjoitustehtävä 5.8. □

5.4 Hyötyfunktion estimointi

Indifferenssikyselyt

Esimerkki 5.2.19 ja sitä seuraava algoritmi antaa jo tavan hyöytyfunktion estimointiin kyselyjen avulla. Esitämme nyt samankaltaisen kyselymetodin.

5.4.1 Määritelmä (Kyselyalgoritmi hyödyille puolituksilla)**.** Haluamme estimoida päätöksentekijän hyöytyfunktion u välillä $[r_-, r_+]$ eli huonoimmasta mahdollisesta palkkiosta parhaaseen mahdolliseen.

(i) Koska voimme aina normeerata hyötyfunktion mille tahansa arvovälille haluamme, valitsemme $u(r_-) = 0$ ja $u(r_+) = 1$.

(ii) Sitten kysymme päätöksentekijältä "subjektiivista" todennäköisyyttä $x_{1/2}$, jolla hänelle pätee indifferenssi

$$L(1, x_{1/2}) \quad \sim \quad \frac{1}{2}L(1, r_+) + \frac{1}{2}L(1, r_-).$$

Tällöin $u(x_{1/2}) = 1/2$. Nimittäin edellisen indifferenssin nojalla

$$\begin{aligned} u(x_{1/2}) &= \mathrm{E}\left(u\left(L(1, x_{1/2})\right)\right) \\ &= \mathrm{E}\left(u\left(\frac{1}{2}L(1, r_+) + \frac{1}{2}L(1, r_-)\right)\right) \\ &= u(r_+) \cdot 0{,}5 + u(r_-) \cdot 0{,}5 \\ &= 0{,}5. \end{aligned}$$

Olemme siis löytäneet ensimmäisen puolittajapisteen $x_{1/2}$.

(iii) Seuraavaksi puolitamme "alavälin" $[0, 1/2]$, eli kysymme puolittajapistettä $x_{1/4}$, jolle pätee indifferenssi

$$L(1, x_{1/4}) \quad \sim \quad \frac{1}{2}L(1, x_{1/2}) + \frac{1}{2}L(1, r_-).$$

Tällöin $u(x_{1/4}) = 1/4$. Nimittäin edellisen indifferenssin nojalla

$$\begin{aligned} u(x_{1/4}) &= \mathrm{E}\left(u\left(L(1, x_{1/4})\right)\right) \\ &= \mathrm{E}\left(u\left(\frac{1}{2}L(1, x_{1/2}) + \frac{1}{2}L(1, r_-)\right)\right) \\ &= u(x_{1/2}) \cdot 0{,}5 + u(r_-) \cdot 0{,}5 \\ &= 0{,}5 \cdot 0{,}5 \\ &= 0{,}25. \end{aligned}$$

(iv) Seuraavaksi puolitamme "ylävälin" $[1/2, 1]$, eli kysymme puolittajapistettä $x_{3/4}$, jolle pätee indifferenssi

$$L(1, x_{3/4}) \quad \sim \quad \frac{1}{2}L(1, r_+) + \frac{1}{2}L(1, x_{1/2}).$$

Tällöin $u(x_{3/4}) = 3/4$. (Perustelu on samakanltainen kuin edellä.)

(v) Jatkamalla puolituskyselyjä riittävän pitkälle saamme hahmoteltua päätöksentekijän hyötyfunktion u arvot $u(r)$ mielivaltaisen tarkasti, eli arvot $u(x_{1/2}) = 1/2$, $u(x_{1/4}) = 1/4$, $u(x_{3/4}) = 3/4$, $u(x_{1/8}) = 1/8$, $u(x_{3/8}) = 3/8$, $u(x_{5/8}) = 5/8$, $u(x_{7/8}) = 7/8$, $u(x_{1/16}) = 1/16$, jne.

5.4.2 Esimerkki. Päivi Päättäjälle tarjotaan palkkiota väliltä $[-10,30]$. Esitä kyselyt, joiden avulla Päivin hyötyfunktio voidaa hahmotella.

Normeeraamalla voimme asettaa Päivin hyötyfunktiossa $u(-10) = 0$ ja $u(30) = 1$. Sitten kysymme indifferenssipistettä $x_{1/2}$, jolle pätee

$$L(1,x_{1/2}) \quad \sim \quad \frac{1}{2}L(1,30)+\frac{1}{2}L(1,-10).$$

Päivi vastaa $x_{1/2} = -3{,}4$. Tiedämme nyt siis, että $u(-3{,}4) = 1/2$. Seuraavaksi kysymme indifferenssipistettä $x_{1/4}$, jolle pätee

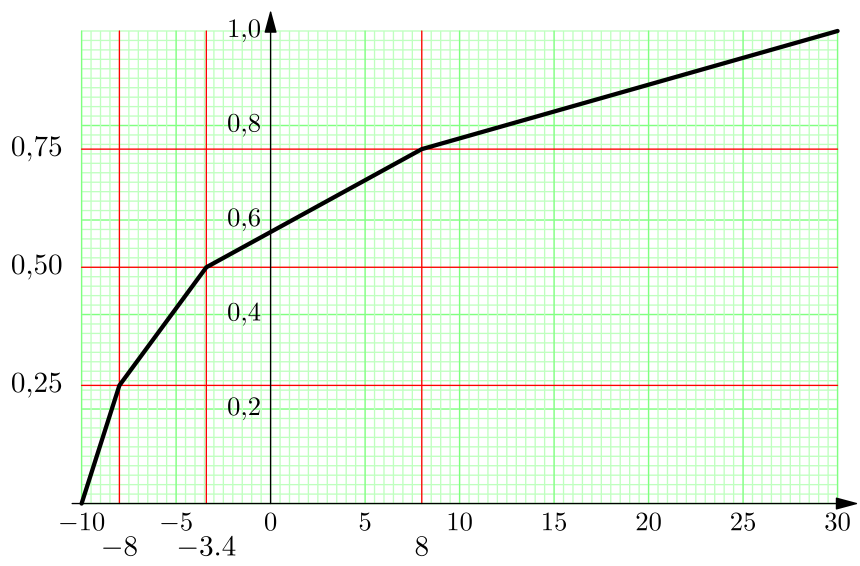
$$L(1,x_{1/4}) \quad \sim \quad \frac{1}{2}L(1,-3{,}4)+\frac{1}{2}L(1,-10).$$

Päivi vastaa $x_{1/4} = -8$. Tiedämme nyt siis, että $u(-8) = 1/4$. Seuraavaksi kysymme indifferenssipistettä $x_{3/4}$, jolle pätee

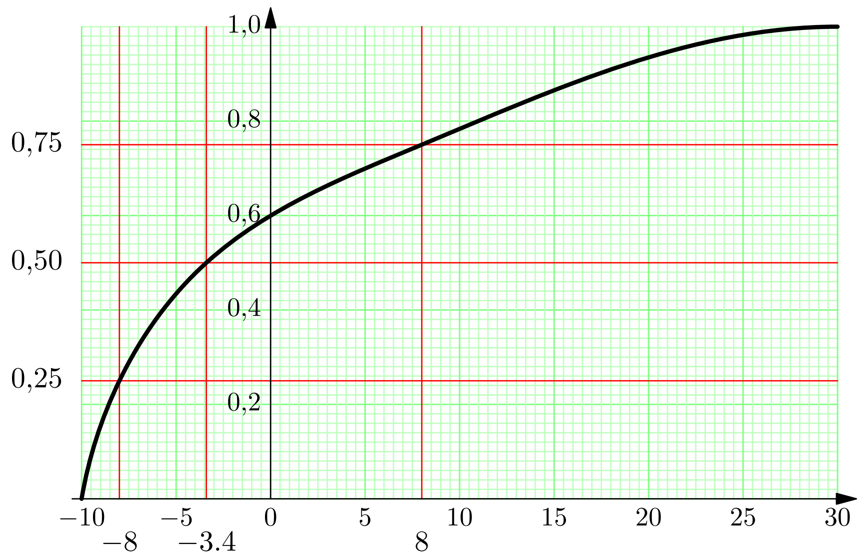
$$L(1,x_{3/4}) \quad \sim \quad \frac{1}{2}L(1,30)+\frac{1}{2}L(1,-3{,}4).$$

Päivi vastaa $x_{3/4} = 8$. Tiedämme nyt siis, että $u(8) = 3/4$.

Lineaarisesti interpoloiden hahmotelmamme Päivin hyötyfunktiosta on:



Sama hahmotelma, mutta kasvavilla Bézierin splineillä lineaarisen interpoloinnin sijaan:



Teoreettisen mallin sovittaminen

Tässä osiossa oletamme, että palkkiot ovat positiivisia ja mahdollisesti ylhäältä rajoittamattomia. Samoin rakentamamme hyötyfunktiot ovat positiivisia ja mahdollisesti ylhäältä rajoittamattomia.

Joskus voidaan postuloida hyötyfunktiolle jokin tietty muoto. Voidaan esimerkiksi olettaa "vakio absoluuttinen riskiaversio" (CARA) tai "vakio suhteellinen riskiaversio" (CRRA).

5.4.3 Määritelmä (Riskiaversio)**.** Olkoon $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ hyötyfunktio. Siihen liittyvä *absoluuttinen riskiaversio* (ARA) on

$$\mathrm{ARA}_u(r) \quad = \quad -\frac{\mathrm{d}^2 u}{\mathrm{d}r^2}(r) \Big/ \frac{\mathrm{d}u}{\mathrm{d}r}(r)$$

ja siihen liittyvä *suhteellinen riskiaversio* (RRA) on

$$\mathrm{RRA}_u(r) \quad = \quad r\mathrm{ARA}_u(r).$$

5.4.4 Huomautus (Riskiaversio, konkaavisuus ja konveksisuus)**.** Jos absoluuttinen riskiaversio ARA_u on positiivinen funktio, on alkuperäinen hyötyfunktio u *konkaavi*, ja siten päätöksentekijä on *riskinkaihtaja*, eli "riskiaversi". Negatiivinen absoluuttinen riskiaversio vastaa *konveksia* hyötyfunktiota ja *riskinrakastamista*.

Seuraavaksi osoitamme, miten oletus "vakiosta absoluuttisesta riskiaversiosta" johtaa parametriseen malliin nimeltä *eksponenttinen hyöty*:

5.4.5 Lause (CARA on exponenttinen hyöty)**.** *Olkoon päätöksentekijän absoluuttinen riskiaversio* ARA_u *vakio. Tällöin hänen hyötyfunktionsa on muotoa*

$$u(r) \quad = \quad 1 - e^{-r/R}$$

jollakin positiivisella parametrilla R.

Todistus. Koska

$$\mathrm{ARA}_u(r) \quad = \quad -\frac{\mathrm{d}^2 u}{\mathrm{d}r^2}(r) \bigg/ \frac{\mathrm{d}u}{\mathrm{d}r}(r)$$

on vakio, niin sen integraali on affiini:

$$\int_0^r \frac{\frac{\mathrm{d}^2 u}{\mathrm{d}x^2}(x)}{\frac{\mathrm{d}u}{\mathrm{d}x}(x)}\,\mathrm{d}x \quad = \quad ar + b.$$

Mutta merkitsemällä $v(x) = \frac{\mathrm{d}u}{\mathrm{d}x}(x)$ voimme ratkaista yllä olevan integraalin:

$$\int_0^r \frac{\frac{\mathrm{d}v}{\mathrm{d}x}(x)}{v(x)}\,\mathrm{d}x \quad = \quad \int_0^r \frac{1}{v(x)}\mathrm{d}v(x) \quad = \quad \ln v(r) - \ln v(0).$$

Siten, vaihtamalla vakioiden nimiä näemme, että

$$\ln v(r) \quad = \quad ar + b.$$

Tämä siis tarkoittaa sitä, että

$$v(r) \quad = \quad e^{ar+b},$$

eli, kun valitsemme $u(0) = 0$, niin

$$\begin{aligned} u(r) &= \int_0^r v(x)\,\mathrm{d}x \\ &= \int_0^r e^{ax+b}\,\mathrm{d}x \\ &= \frac{e^b}{a}\int_0^r e^{ax}\,\mathrm{d}(ax) \\ &= \frac{e^b}{a}\left(e^{ar} - 1\right). \end{aligned}$$

Väite seuraa tästä valitsemalla $a = -e^b$ ja merkitsemällä $R = -1/a$. Valinta $a = -e^b$ voidaan tehdä, sillä muilla valinnoilla saatu hyötyfunktio olisi vain affinimuunnos tästä valinnasta. □

5.4.6 Huomautus (Riskitoleranssi ja sen estimointi)**.** Lauseen 5.4.5 parametria R kutsutaan *riskitoleranssiksi* Se voidaan estimoida (approksimatiivisesti) kyselyllä

1. et saa mitään,
2. saat R€ todennäköisyydellä $1/2$ ja $-R/2$€ todennäköisyydellä $1/2$.

Se R, jolla olet indifferentti näiden vaihtoehtojen suhteen on riskitoleranssisi ja määrää siten ekponenttisen hyötyfunktiosi.

5.4.7 Esimerkki (Vakio suhteellinen riskiaversio CRRA)**.** Lause 5.4.5 kertoo kaikki höytyfunktiot, joilla on vakio absoluuttinen riskiarversio (CARA). Vakio suhteellinen riskiaversio (CRRA) on kolmella luokalla hyötyfunktioita. Yksi luokka on

$$u(r) \quad = \quad r^{\alpha}, \quad 0 < \alpha < 1.$$

Näille hyötyfunktioille $\mathrm{RRA}_u(r) = 1 - \alpha$.

5.5 Harjoitustehtäviä lukuun 5

5.1 Harjoitustehtävä. (a) Kuinka halvalla esimerkin 5.1.2 Oiva Opiskelija olisi valmis riskeeraamaan pikkurillinsä?

(b) Kuinka paljon esimerkin 5.1.2 Tarmo Toimitusjohtajalle pitäisi pikkurillistä maksaa, jotta hän olisi valmis riskeeraamaan sen?

5.2 Harjoitustehtävä. Tarkastelemme esimerkkiä 5.1.3

(a) Etsi kokonaishyödyn maksimoiva jakosuhde r_j, $j \in J$, kokonaisvarallisuudelle, kun kansalaisilla on sama konkaavi hyötyfunktio u.

(b) Etsi kokonaishyödyn maksimoiva jakosuhde r_j, $j \in J$, kokonaisvarallisuudelle, kun kansalaisilla on sama konveksi hyöytyfunktio u.

5.3 Harjoitustehtävä. Olkoon Leena Lokin hyöytyfunktio $u(r) = \ln r$. Leena Lokilla on 20.000€. Hänelle tarjotaan arpajaisia

$$\begin{aligned} L_1 &= L(1, -1.000€), \\ L_2 &= 0{,}9L(1, 0€) + 0{,}1L(1, -10.000), \end{aligned}$$

missä palkkiot ovat voittoja/tappioita.

Kumman arpajaisista Leena Lokki valitsee?

5.4 Harjoitustehtävä. Pekka Päättäjä on päättänyt itselleen seuraavat indifferenssit:

$$\begin{aligned} L(1, 10) &\sim 0{,}5L(1, 0) + 0{,}5L(1, 20), \\ L(1, 12) &\sim 0{,}4L(1, 10) + 0{,}6L(1, 20), \\ L(1, 15) &\sim 0{,}2L(1, 10) + 0{,}8L(1, 20). \end{aligned}$$

(a) Mitä pystyt sanomaan Pekka Päättäjän hyötyfunktiosta?
(b) Onko Pekka Päättäjä riskinrakastaja vai riskinkaihtaja?

5.5 Harjoitustehtävä. Oiva Opiskelijan on päätettävä osallistuako kurssille "Operaatioanalyysi" vai "Johdatus tilastotieteeseen". Oiva arvioi, että jos hän osallistuu kurssille "Operaatioanalyysi", niin todennäköisyydellä 10% hän saa arvosanan *5*, todennäköisyydellä 40% arvosanan *4* ja todennäköisyydellä 50% arvosanan *3*. Samoin Oiva arvelee, että jos hän osallistuu kurssille "Johdatus tilastotieteeseen", niin todennäköisyydet arvosanoille ovat: 70% arvosanalle *4*, 25% arvosanalle *3* ja 5% arvosanalle *2*.

Oiva on indifferentti arpajaisten $L(1,\ \mathit{3})$ ja $L(0{,}25\,,\ \mathit{5}\ ;\ 0{,}75\,,\ \mathit{2})$ suhteen, sekä arpajaisten $L(1,\ \mathit{4})$ ja $L(0{,}70\,,\ \mathit{5}\ ;\ 0{,}30\,,\ \mathit{2})$ suhteen.

Jos Oiva Opiskelija haluaa optimoida arvosanasta saadun keskimääräisen (eli odotetun) hyödyn, niin kumman kursseista "Operaatioanalyysi" vai "Tilastotieteen johdantokurssi" hän valitsee?

5.6 Harjoitustehtävä. Esimerkkien 5.2.15 ja 5.2.19 ratkaisussa ja sitä seuraavassa yleisessä algoritmissa käytettiin kaikkia Von Neumannin ja Morgensternin aksioomia.

(a) Etsi jokaiselle Von Neumann–Morgenstern -aksioomalle kohta, missä sitä käytettiin.
(b) Mikä ratkaisualgoritmissa menee pieleen, jos jokin Von Neumannin ja Morgensternin aksioomista ei pitäisikään paikkaansa?
(c) Onko jokin Von Neumann–Morgenstern -aksioomista mielestäsi erityisen kyseenalainen? Jos mielestäsi useakin niistä on kyseenalainen, niin mikä on mielestäsi kaikkein kyseenalaisin?

5.7 Harjoitustehtävä. Lasse Lottoaja lottoaa yhden rivin joka viikko. Hänellä on myös 200.000 € kotivakuutus, josta hän maksaa 100 € vuosittaista vakuutusmaksua. Kotivakuutuksessa on 500 € omavastuu.

(a) Onko Lasse Lottoaja riskinkaihtaja?
(b) Onko Lasse Lottoaja riskinrakastaja?
(c) Onko Lasse Lottoaja epärationaalinen?

5.8 Harjoitustehtävä. Todista Von Neumann–Morgenstern -hyötylause 5.3.3.

5.9 Harjoitustehtävä. (a) Von Neumann–Morgenstern -aksioomista 5.3.1 ja -hyötylauseesta 5.3.3 *ei seuraa*, että hyötyfunktio u on aidosti kasvava tai edes jatkuva. Perustele tämä.
(b) Perustele, että Von Neumann–Morgenstern -rationaalinen päätöksentekijä voi olla pöhkö.

(c) Seuraako Von Neumann–Morgenstern -aksioomista 5.3.1 ja hyötylauseesta 5.3.3, että hyötyfunktio u on kasvava?

5.10 Harjoitustehtävä. Miten Von Neumann–Morgenstern -aksioomia 5.3.1 tulee muuttaa, jotta lauseen 5.3.3 hyötyfunktio olisi jatkuva ja aidosti kasvava?

5.11 Harjoitustehtävä. Perustele, miksi huomautuksen 5.4.6 kysely estimoi riskitoleranssin R approximatiivisesti.

5.12 Harjoitustehtävä. Etsi kaikki höytyfunktiot u, joilla on vakio suhteellinen riskiaversio (CRRA), eli

$$\mathrm{RRA}_u(r) \quad = \quad \text{vakio.}$$

Luku 6

Hyötyteorian kritiikki

> Your moral feelings are attached to frames, to descriptions of reality rather than to reality itself. – Daniel Kahneman

> It is important that students bring a certain ragamuffin, barefoot irreverence to their studies; they are not here to worship what is known, but to question it. – Jacob Bronowski

> Just because you and I are arguing, doesn't mean one of us is right.
> – Professor S.

Von Neumann–Morgensternläisessä hyötyteoriassa on sellainen pikkuisen ikävä puoli, että käytännössä ihmiset eivät näytä noudattavan sen aksioomia. Yksi keskeinen kritiikki heidän hyötyteoriaa vastaan on Tverskyn ja Kahnemanin vuonna 1981 esittämät *prospektiteoria* ja *ankkurointiefekti*.

6.1 Prospektiteoria

6.1.1 Esimerkki (Allais'n paradoksi)**.** Tarkastelemme paradoksia, jonka Maurice Allais esitti vuonna 1953. Paradoksin tarkoitus oli lähinnä kritisoida Von Neumannin ja Morgensternin riippumattomuusaksioomaa.

Herra A.:n on valittava seuraavien arpajaisten väliltä (palkkiot ovat euroja):

$$\begin{aligned} L_1 &= L(1, 1.000.000), \\ L_2 &= L(0{,}10\,,\ 5.000.000\ ;\ 0{,}89,\ 1.000.000\ ;\ 0{,}01\,,\ 0). \end{aligned}$$

Herra A., kuten valtaosa muistakin ihmisistä, valitsee arpajaiset L_1. Seuraavaksi herra A.:n on valittava seuraavien arpajaisten väliltä:

$$\begin{aligned} L_3 &= L(0{,}11\,,\ 1.000.000\ ;\ 0{,}89\,,\ 0), \\ L_4 &= L(0{,}10\,,\ 5.000.000\ ;\ 0{,}90\,,\ 0). \end{aligned}$$

Herra A., kuten valtaosa muistakin ihmisistä, valitsee arpajaiset L_4.

Tarkastelemme, millainen herra A.:n hyötyfunktion tulisi olla. Voimme normeerata $u(5.000.000) = 1$ ja $u(0) = 0$. Merkitsemme $q = u(1.000.000)$. Herra A. valitsee arpajaiset L_1 mieluummin kuin arpajaiset L_2. Tämä tarkoittaa Von Neumann–Morgensterniläisittäin sitä, että

$$q \quad > \quad 0{,}10 \cdot 1 \ + \ 0{,}89 \cdot q \ + \ 0{,}01 \cdot 0$$

eli $q > 0{,}90909$. Toisaalta herra A. valitsee arpajaiset L_4 mieluummin kuin arpajaiset L_3. Tämä tarkoittaa Von Neumann-Morgensterniläisittäin sitä, että

$$0{,}10 \cdot 1 \ + \ 0{,}90 \cdot 0 \quad > \quad 0{,}11 \cdot q \ + \ 0{,}89 \cdot 0$$

eli $q < 0{,}90909$.

Herra A.:lla ei siis voi olla hyötyfunktiota, sillä jos hänellä sellainen olisi, niin olisi myös olemassa luku q, joka on *sekä* suurempi kuin 0,90909 *että* pienempi kuin 0,90909. (Tämä argumentti on *reductio ad absurdum* eli perustelu ristiriidan kautta.) Koska hyötyfunktion olemassaolo on yhtäpitävää Von Neumann–Morgenstern -rationaalisuuden kanssa, johtopäätös on, että herra A., kuten valtaosa muistakin ihmisistä, on epärationaalinen.

Herra A.:n epärationaalisuus voidaan selittää pois esimerkiksi Tverskyn ja Kahnemanin esittämän *prospektiteorian* avulla. Prospektiteoriassa ideana on korvata todennäköisyydet p *prospekteilla* $\Pi(p)$, missä prospektifunktio $p \mapsto \Pi(p)$ muuttuu nopeammin, kun p on lähellä ykköstä tai nollaa. Tällä on tarkoitus mallintaa matemaattisesti se havaittu psykologinen tosiseikka, että ihmisillä on tapana antaa suurille ja pienille todennäköisyyksille liian suuri paino. Ihmiset esimerkiksi tyypillisesti pitävät muutosta todennäköisyydestä 0,01 todennäköisyyteen 0,02 huomattavasti merkittävämpänä kuin muutosta todenäköisyydestä 0,41 todennäköisyyteen 0,42, vaikka molemmissa tapauksissa todennäköisyyksien muutos on sama 1 %-yksikkö.

Prospektiteorian mukainen päätössääntö melkein kuten hyötyteorian mukainen päätössääntö. Ainoa ero on, että summassa todennäköisyydet p on korvattu prospekteilla $\Pi(p)$. Toisin sanoen jokaista hyötyfunktiota u ja prospektifunktiota Π vastaa arpajaispreferenssit seuraavan määritelmän 6.1.2 kuvaamalla tavalla.

6.1.2 Määritelmä (Prospektipäätöksenteko)**.** Päätöksentekijä preferoi arpajaisia L' yli arpajaisten L, eli $L \prec L'$, jos ja vain jos

$$\begin{aligned}
\mathrm{E}_\Pi\Big(u(L)\Big) &= \sum_{j \in J} u(r_j)\,\Pi(p_j) \\
&< \sum_{j \in J} u(r'_j)\,\Pi(p'_j) \\
&= \mathrm{E}_\Pi\Big(u(L')\Big).
\end{aligned}$$

Samoin päätöksentekijä on indifferentti arpajaisten L ja L' välillä, eli $L \sim L'$, jos ja vain jos

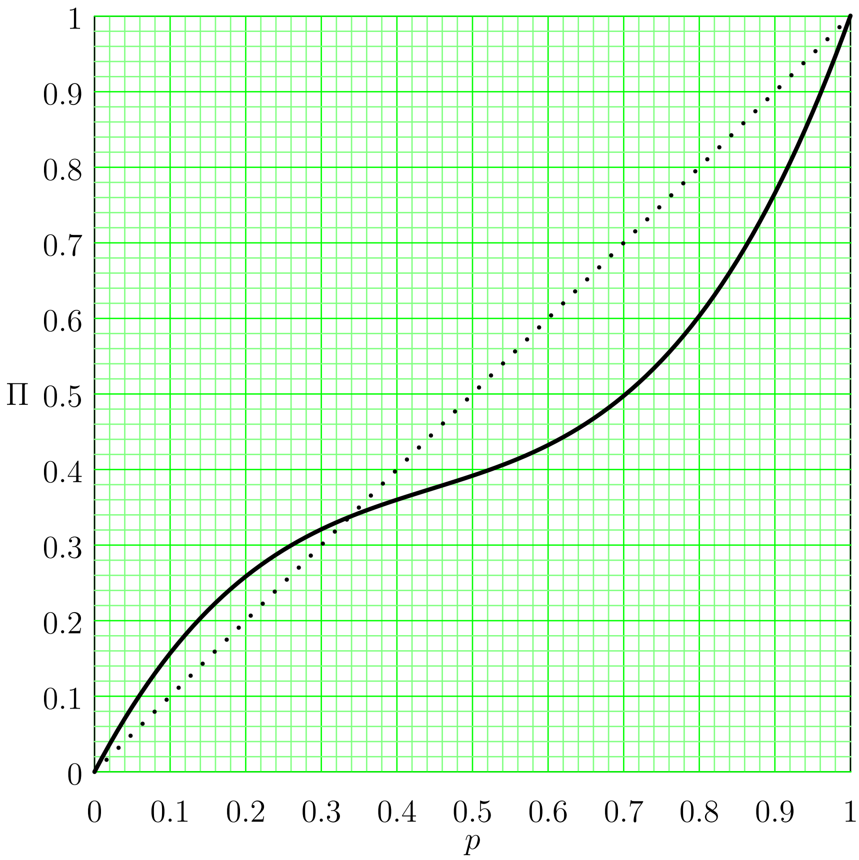
$$\begin{aligned} \mathrm{E}_\Pi\Big(u(L)\Big) &= \sum_{j\in J} u(r_j)\,\Pi(p_j) \\ &= \sum_{j\in J} u(r'_j)\,\Pi(p'_j) \\ &= \mathrm{E}_\Pi\Big(u(L')\Big). \end{aligned}$$

6.1.3 Esimerkki. Tarkastelemme esimerkin vuoksi prospektifunktiota

$$\Pi(p) \quad = \quad 1{,}89799p - 3{,}55995p^2 + 2{,}662549p^3.$$

Tällä prospektifunktiolla on se toivottu ominaisuus, että se muuttuu nopeammin nollan ja ykkösen lähellä ja hitaammin "keskitodennäköisyyksillä". Muuten valittu prospektifunktio on täysin ad hoc, eli sille ei ole mitään teoreettisia perusteluja. Emme käsittele tällä kurssilla prospektifunktion empiiristä estimointia.

Kuvallisesti siis tarkastelemme prospektifunktiota



Nyt (hyötyinä mitatut) *varmuusvastineet*[1] ovat

$$\begin{aligned}
\mathrm{E}_\Pi\left(u(L_1)\right) &= q\cdot\Pi(1)\\
&= q,\\
\mathrm{E}_\Pi\left(u(L_2)\right) &= 1\cdot\Pi(0{,}10) \;+\; q\cdot\Pi(0{,}89) \;+\; 0\cdot\Pi(0{,}01)\\
&= \Pi(0{,}10) \;+\; q\cdot\Pi(0{,}89)\\
&= 0{,}15686 \;+\; q\cdot 0{,}74639,\\
\mathrm{E}_\Pi\left(u(L_3)\right) &= q\cdot\Pi(0{,}11) \;+\; 0\cdot\Pi(0{,}89)\\
&= q\cdot\Pi(0{,}11)\\
&= q\cdot 0{,}16925,\\
\mathrm{E}_\Pi\left(u(L_4)\right) &= 1\cdot\Pi(0{,}10) \;+\; 0\cdot\Pi(0{,}90)\\
&= \Pi(0{,}10)\\
&= 0{,}15686.
\end{aligned}$$

Herra A.:n arpajaispreferenssit $L_1 \succ L_2$ ja $L_4 \succ L_3$ asettavat ehdot $\mathrm{E}_\Pi(u(L_1)) > \mathrm{E}_\Pi(u(L_2))$ ja $\mathrm{E}_\Pi(u(L_4)) > \mathrm{E}_\Pi(u(L_3))$. Toisin sanoen

$$\begin{aligned}
q &> 0{,}15686 + q\cdot 0{,}74639,\\
0{,}15686 &> q\cdot 0{,}16925.
\end{aligned}$$

Näistä saamme q:lle ehdot

$$0{,}61851 \;<\; q \;<\; 0{,}92679.$$

Johtopäätös on, että prospektiteoria selittää paradoksin.

6.2 Ankkurointiefekti

6.2.1 Määritelmä (Ankkurointiefekti)**.** Ankkuroinnilla (engl. framing) tarkoitetaan sitä, että päätöksentekijä riskipreferenssit kääntyvät ympäri, kun voitosta siirrytään tappioihin. Päätöksentekijä siis ankkuroituu nykytilanteeseen ja on tyypillisesti riskinrakastaja tappioille (vaikkei sitä usein huomaakaan) ja riskinkaihtaja voitoille.

6.2.2 Huomautus. Huomattavaa on, että ankkuroinnissa on kyse *kognitiivisesta harhasta*, missä täsmälleen samaan tilanteeseen suhtaudutaan eri tavalla, jos tilanne *muotoillaan* eri tavalla. Pahaiten tämä käy ilmi harjoitustehtävässä 6.1.

[1] Arpajaisten *varmuusvastine* on se luku — hyöty tai palkkio — jolla päätöksentekijä on valmis luopumaan oikeudestaan tai velvollisuudestaan osallistua arpajaisiin.

6.2.3 Esimerkki (Ankkurointiparadoksi)**.** Olkoot arpajaiset *voittoineen tai tappioineen*

$$\begin{aligned} L_1 &= L(1\,,\,250\,€), \\ L_2 &= L(0{,}25\,,\,1.000\,€\,\,;\,\,0{,}75\,,\,0\,€), \\ L_3 &= L(1\,,\,-750\,€), \\ L_4 &= L(0{,}75\,,\,-1.000\,€\,\,;\,\,0{,}25\,,\,0\,€). \end{aligned}$$

84% kaikista ihmisistä valitsee L_1:n mieluummin kuin L_2:n ja 87% ihmisistä valitsee L_4:n mieluummin kuin L_3:n.

Huomaamme aluksi, vertailun vuoksi, että riskineutraalin päätöksentekijän preferenssit olisivat $L_4 \sim L_3 \prec L_2 \sim L_1$. Koska kyseessä on tappiot/voitot, niin preferenssi $L_2 \prec L_1$ tarkoittaa, että päätöksentekijä on riskinkaihtaja voitoille ja samoin preferenssi $L_3 \prec L_4$ tarkoittaa, että päätöksentekijä on riskinrakastaja tappioille. Koska tämän on pädettävä kaikille päätöksentekijän mahdollisille varallisuuksille, hyötyfunktion on oltava affiini. Mutta hyötyfunktio ei voi olla affiini, sillä $L_3 \not\sim L_4$.

Ankkurointiefekti selittää tämän paradoksin: voittojen suhteen ollaan riskinkaihtajia ja tappioiden suhteen ollaan riskinrakastajia.

6.3 Suhtautuminen kritiikkiin

Tverskyn ja Kahnemanin Von Neumann–Morgenstern -hyötyteorian kritiikkiin voi suhtautua monellakin tavalla. Kaksi toisistaan täysin poikkeavaa tapaa on:

I Tverskyn ja Kahnemanin havainnot osoittavat, että ihmiset eivät toimi Von Neumann–Morgenstern -hyötyteorian mukaan. Siten hyötyteoria ei vastaa todellisuutta, ja sen antama hyöty (☺) on vähintäänkin kyseenalainen!

II Tversky, Kahneman ja kumppanit eivät ole havainnoillaan osoittaneet mitään muuta kuin sen, että ihmiset toimivat epärationaalisesti. Tämän ei pitäisi olla mikään järisyttävä uutinen kenellekään!

6.4 Harjoitustehtäviä lukuun 6

6.1 Harjoitustehtävä. Sinun on valmistauduttava possuköhän etenemiseen Suomessa. Oletetavaa on, että possuköhä tappaa 600 ihmistä. Voit valita kahdesta rokotusohjelmasta:

Ohjelma I 200 ihmistä säästyy.
Ohjelma II Todennäköisyydellä 1/3 600 ihmistä säästyy.

Kumman rokotusohjelman valitset?

Entä kumman rokotusohjelman valitset, jos vaihtoehdot ovat

Ohjelma I’ 400 ihmistä kuolee.
Ohjelma II’ Todennäköisyydellä 2/3 600 ihmistä kuolee.

6.2 Harjoitustehtävä. Tarkastelemme *Ellsbergin paradoksia.* Laatikossa on 90 palloa. Palloista 30 on punaisia, ja loput palloista ovat joko keltaisia tai mustia. Laatikosta nostetaan yksi pallo umpimähkään. Seuraavat palkkiovaihtoehdot ovat nyt tarjolla:

Vaihtoehto 1 Saat 1.000 €, jos tulee punainen pallo.
Vaihtoehto 2 Saat 1.000 €, jos tulee keltainen pallo.
Vaihtoehto 3 Saat 1.000 €, jos tulee keltainen tai musta pallo.
Vaihtoehto 4 Saat 1.000 €, jos tulee punainen tai musta pallo.

Useimmat ihmiset valitsevat vaihtoehdon 1 mieluummin kuin vaihtoehdon 2 ja vaihtoehdon 3 mieluummin kuin vaihtoehdon 4. Miksi tämä on epärationaalista?

6.3 Harjoitustehtävä. Tarkastelemme *Tverskyn ja Kahnemanin paradoksia.* Olkoon arpajaiset *tappioinen tai voittoinen*

$$
\begin{aligned}
L_1 &= L(0{,}001\,,\ 5.000\$\ ;\ 0{,}999\,,\ 0\$), \\
L_2 &= L(1\,,\ 5\$), \\
L_3 &= L(0{,}001\,,\ -5.000\$;\ 0{,}999\,,\ 0\$), \\
L_4 &= L(1\,,\ -5\$).
\end{aligned}
$$

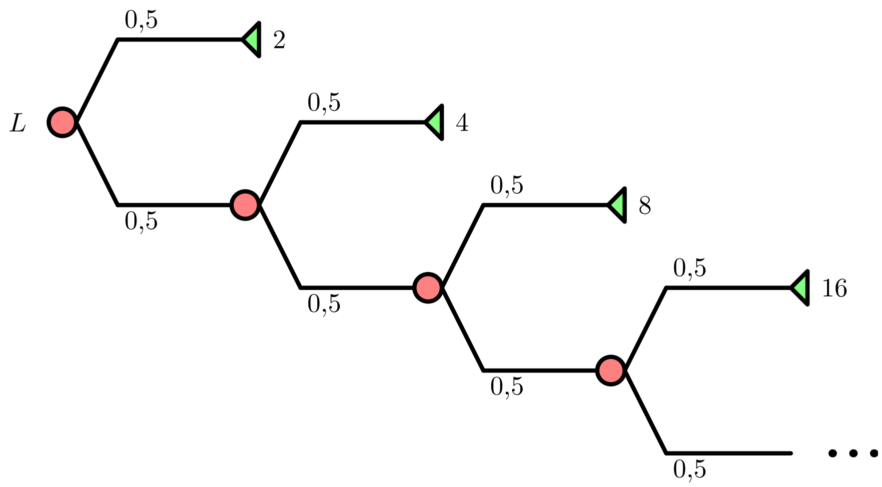
Tversky ja Kahneman pyysivät 72 henkilöä valitsemaan arpajaisten L_1 ja L_2 sekä L_3 ja L_4 väliltä. Yli 75% valitsi arpajaiset L_1 mieluummin kuin arpajaiset L_2 ja arpajaiset L_4 mieluummin kuin arpajaiset L_3.

(a) Miten riskineutraali päätöksentekijä arvottaisi arpajaiset?
(b) Miten riskiä rakastava päätöksentekijä arvottaisi arpajaiset?
(c) Miten riskiä kaihtava päätöksentekijä arvottaisi arpajaiset?
(d) Miten sinä järjestäisit arpajaiset?
(e) Miksi Tverskyn ja Kahnemanin koe on ristiriidassa Von Neumann–Morgenstern -hyötyteorian kanssa?
(f) Miten prospektiteoria tai ankkurointiefekti voi selittää Tverskyn ja Kahnemanin kokeen?

6.4 Harjoitustehtävä. Tarkastelemme *Pietarin paradoksia.* Olkoon L seuraavat (mahdollisesti äärettömät) arpajaiset: Reilua kolikkoa heitetään kunnes tulee ensimmäinen klaava. Jos klaava tulee n:nnellä heitolla, pelaaja saa 2^n euroa. Toisin sanoen L on äärettömät arpajaiset

$$\begin{aligned} L &= L\left(2^{-1}, 2^1 \ ; \ 2^{-2}, 2^2 \ ; \ 2^{-3}, 2^3 \ ; \ 2^{-4}, 2^4 \ ; \ \ldots\right) \\ &= L\left(0{,}5, 2 \ ; \ 0{,}25, 4 \ ; \ 0{,}125, 8 \ ; \ 0{,}0625, 16 \ ; \ \ldots\right). \end{aligned}$$

Sama kuvallisesti:



(a) Mikä on suurin hinta, jonka riskineutraali päätöksentekijä suostuu maksamaan osallistumismaksuna arpajaisista L?
(b) Mikä on suurin hinta, jonka hyötyfunktion $u(r) = \log_2 r$ omaava päätöksentekijä on valmis maksamaan arpajaisista L?
(c) Mikä on suurin hinta, jonka hyötyfunktion $u(r) = \ln r$ omaava päätöksentekijä on valmis maksamaan arpajaisista L?

6.5 Harjoitustehtävä. Kokeessa sinulle tarjotaan kaksi mahdollisuutta:

Vaihtoehto 1 saat 5 €,
Vaihtoehto 2 saat 1.000 €.

Voit valita kumman tahansa, mutta jos valitset epärationaalisesti, niin kokeen järjestäjä antaa sinulle 1.000.000 €. Miten valitset?

6.6 Harjoitustehtävä. (a) Etsi tästä kirjasta vähintään yksi painovirhe ja vähintään yksi “aito” virhe ja lähetä vastauksesi luennoijalle osoitteeseen `tommi.sottinen@uwasa.fi` otsikolla `ORMS2020-virheita`.
(b) Anna palautetta (kehuja, haukkuja, parannusehdotuksia) tästä kurssista ja/tai näistä luentomuistiinpanoista luennoijalle osoitteeseen `tommi.sottinen@uwasa.fi` otsikolla `ORMS2020-palaute`.

Kirjallisuutta

[1] Robert Clemen (1996) *Making Hard Decisions: An Introduction to Decision Analysis*, Second edition, Duxbury Press.

[2] Franz Kafka (1915) *Muodonmuutos.*

[3] Mika Koskenoja (2002) *Sattuman matematiikkaa I – klassinen todennäköisyys*, Solmu 2002:2.

[4] Howard Raiffa (1968) *Decision Analysis: Introductory Lectures on Choices under Uncertainty*, Random House.

[5] Tommi Sottinen (2008) *Todennäköisyysteoria: Teoria mitasta, mitallisuudesta, mitattomuudesta ja riippumattomuudesta*, Luentomuistiinpanot.

[6] Nassim Taleb (2004) *Satunnaisuuden hämäämä*, Terra Cognita.

[7] Nassim Taleb (2007) *Musta joutsen: Erittäin epätodennäköisen vaikutus*, Terra Cognita.

[8] Pekka Tuominen (1993) *Todennäköisyyslaskenta I*, 2. tarkistettu painos, Limes.

[9] Wayne Winston (2004) *Operations Research: Applications and Algorithms*, Fourth edition, Brooks/Cole.

Hakemisto