



OPPIMISMALLIT KNIGHTILAISEN EPÄVARMUUDEN VALLITESSA

Simón Albert Álvarez

Pro gradu -tutkielma
2026

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Tarkastajat:

Prof. Jukka Lempa

FT Harto Saarinen

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä

TURUN YLIOPISTO, Matematiikan ja tilastotieteen laitos

Pro gradu -tutkielma

Pääaine: Matematiikka

Tekijä: Simón Albert Álvarez

Otsikko: Oppimismallit Knightilaisen epävarmuuden vallitessa

Ohjaaja: Prof. Jukka Lempa

Sivumäärä: 52 sivua + liitteet 6 sivua

Aika: 2026

Tutkielmassa tarkastellaan oppimista Knightilaisen epävarmuuden vallitessa jatkuvan ajan stokastisessa mallissa. Lähtökohtana on Chenin ja Epsteinin robusti ja aikakonsistentti kehikko, jota laajennetaan oppimiseen tuntemattomasta parametrista Brownin liikkeeseen perustuvan prosessin avulla. Epävarmuus mallinnetaan priorijoukkoina, ja päätöksentekijän preferenssit määritellään maxmin-periaatteen mukaisesti.

Keskeinen osa tutkielmaa on mallin rakenteen systemaattinen esitys sekä useiden kirjallisuudessa vain luonnosteltujen tulosten täsmällinen todistaminen. Oppimisiongelma muotoillaan optimaalisen pysäytyksen ongelmana, jossa päätöksentekijä voi lykätä päätöksen valintaa havainnoidakseen lisäinformaatiota kustannusta vastaan. Tarkastelua sovelletaan kahteen erityistapaukseen: Ellsbergin paradoksin dynaamiseen versioon sekä robustiin kahden hypoteesin peräkkäiseen testaukseen Wienerprosessin tuntemattomasta ajautumasta. Molemmissa tapauksissa johdetaan eksplisiittiset ratkaisut ja analysoidaan, miten epävarmuuden aste vaikuttaa oppimisen kannattavuuteen ja optimaaliseen pysäytysstrategiaan.

Asiasanat: Ambiguiteetti (Knightilainen epävarmuus), oppiminen epävarmuuden vallitessa, robusti päätöksenteko, optimaalinen pysäytys, aikakonsistenssi, Hamilton-Jacobi-Bellman yhtälö (HJB-yhtälö).

Sisällys

1	Johdanto	1
2	Rekursiivinen hyötymalli oppimisen kanssa	3
2.1	Suorakulmaisuus ja sen perustelu	3
2.2	Mallin rakenne	6
3	Optimaalinen oppiminen	12
3.1	Mallin rakenne ja yleinen ongelmanmuotoilu	12
3.2	Oppiminen ja Ellsbergin uurnat	15
3.3	Robusti peräkkäinen hypoteesin testaus	22
4	Yleisempi tulos	25
5	Todistukset ja muut tulokset	30
5.1	Lauseen 5 todistus	30
5.2	Lemma 2:n todistus	42
5.3	Sovelluksien todistukset	48
5.4	Lauseen 4 todistus	49
5.5	Nash-tasapainotulkinta Ellsbergin mallissa	50

1 Johdanto

Tutkielmassa tarkastellaan oppimista Knightilaisen epävarmuuden alaisuudessa jatkuvan ajan stokastisessa ympäristössä. Knightilainen epävarmuus viittaa tilanteeseen, jossa epävarmuutta ei voida kuvata yksittäisellä todennäköisyysjakaumalla. Frank Knight[16] teki erottelun mitattavan riskin ja sellaisen epävarmuuden välillä, jossa todennäköisyyksiä ei tunneta. Jälkimmäistä kutsutaan nykyisin usein ambiguuteetiksi.

Perinteisissä todennäköisyysmalleissa päätöksentekijä arvioi vaihtoehtoja yhden priorijakauman avulla. Knightilaisen epävarmuuden tapauksessa tämä ei ole mahdollista: päätöksentekijä ei tiedä, mikä todennäköisyysmalli on oikea. Tällöin uskomuksia kuvataan joukkona mahdollisia priorijakaumia. Koska uskomukset parametrin arvosta muodostavat joukon eivätkä yksittäistä jakaumaa, päätöksentekijä arvioi vaihtoehtoja robustisti. Hän tarkastelee kunkin toimintavaihtoehdon odotettua hyötyä kaikkien sallittujen mallien suhteen ja perustaa päätöksensä pahimman tapauksen arvioon. Toisin sanoen valinta perustuu robustiin päätäntään, jossa mahdollisia priorivalintoja tarkastellaan pahimman tapauksen näkökulmasta. Näin myös oppimispäätös (eli kuinka kauan signaalia seurataan) määräytyy robustin optimointiperiaatteen mukaisesti.

Tämän tutkielman matemaattinen kehikko perustuu pääosin Epstein ja Jin[12] esittämään malliin "Optimal Learning under Robustness and Time-Consistency", joka puolestaan rakentuu Chenin ja Epsteinin[7] jatkuvan ajan robustille ja aikakonsistentille hyötymallille. Tutkielmassa esitetään kyseinen rakenne systemaattisesti, täsmennetään useita teknisiä yksityiskohtia sekä johdetaan eksplisiittisesti useita tuloksia, jotka kirjallisuudessa esitetään tiiviimmässä muodossa. Analyysi pohjautuu pääosin jatkuvan ajan suodatusteoriaan sekä Girsanovin muunnokseen.

Tarkastellaan päätöksentekijää, joka tekee valinnan tilanteessa, jossa tuotto riippuu sekä satunnaisista ulkoisista tekijöistä että tuntemattomasta parametrilla θ . Parametri voi saada kaksi mahdollista arvoa, mutta päätöksentekijä ei tiedä, kumpi niistä tulee toteutumaan. Lisäksi hänellä ei ole yksikäsitteistä todennäköisyysarviota parametrin arvosta, vaan pelkästään hänen alkuperäisiä uskomuksiaan kuvaava priorijoukko.

Ennen lopullista päätöstä päätöksentekijä voi hankkia lisäinformaatiota havainnoimalla signaaliprosessia, jonka kehitys riippuu tuntemattoman parametrin todellisesta arvosta. Havainnoinnista aiheutuu kustannus aikayksikköä kohden, minkä vuoksi oppiminen ei ole ilmaista. Päätöksentekijän ongelmaksi muodostuu tällöin tasapainottaa informaation arvo ja sen hankkimisen kustannus.

Dynaamisessa ympäristössä robusti arviointi ei kuitenkaan automaattisesti takaa aikakonsistenttisuutta. Ilman lisärakennetta voi syntyä tilanne, jossa alkuhetkellä optimaalinen suunnitelma ei säily optimaalisena myöhäisemmässä ajankohdassa parametria koskevan informaation kertyessä ajassa. Tässä tutkielmassa hyödynnetään Chenin ja Epsteinin kehikkoa, jossa priorijoukolle asetettu rakenne-ehto varmistaa, että päätöksentekijän arviointiperiaate säilyy johdonmukaisena ajan kuluessa. Tämä mahdollistaa oppimisen ja robustin päätöksenteon yhdistämisen yhtenäiseksi dynaamiseksi malliksi.

Mallia sovelletaan kahteen teorianmuodostuksen kannalta keskeiseen erityista-

paukseen. Ensimmäinen on Ellsbergin paradoksin dynaaminen versio. Klassinen Ellsbergin[5] koe havainnollistaa, että päätöksentekijät suosivat usein riskillistä vaihtoehtoa ambiguuteetin sijaan. Tässä tutkielmassa tarkastellaan, miten tämä käyttäytyminen muuttuu, kun päätöksentekijällä on mahdollisuus oppia epäselvän vaihtoehdon ominaisuuksista. Kuten myöhemmin tutkielmassa toteutettu analyysi osoittaa, riittävän suuren alkuperäisen epävarmuuden tapauksessa oppiminen ei ole optimaalista lainkaan, kun taas pienemmän epävarmuuden tilanteessa oppimista jatketaan, kunnes signaali antaa riittävän vahvaa näyttöä parametrin arvosta.

Toinen sovellus koskee kahden hypoteesin peräkkäistä testausta tilanteessa, jossa ajautumaparametri on tuntematon ja siihen liittyy ambiguuteettia. Myös tässä tapauksessa saadaan eksplisiittinen rajarakenne optimaalisen pysäytyksen karakterisointiin. Robusti ratkaisu poikkeaa oleellisesti klassisesta Bayesilaisesta testauksesta, sillä päätös perustuu pahimman tapauksen posterioriin eikä yksittäiseen prioriin.

Tutkielman painopiste ei rajoitu vain mallin esittelyyn, vaan erityisesti sen rakenteen täsmälliseen perusteluun ja täydentämiseen. Priorijoukon suorakulmaisuuuden rekursiivinen konstruktio johdetaan eksplisiittisesti, ja dynaamisen min-iteroinnin rakenne osoitetaan yksityiskohtaisesti. Lisäksi johdetaan ja perustellaan useita tuloksia, jotka alkuperäisessä artikkelissa esitetään ilman täydellisiä todistuksia. Kapaleessa 5 useita Epstein ja Jin todistuksia täydennetään ja tuodaan esille tuloksia, joihin kyseisten tutkimusten tekstissä vain viitataan. Näin tutkielmassa pyritään matemaattisesti loogiseen ja johdonmukaiseen kehikkoon, joka selventää robustin oppimismallin rakennetta ja sen dynaamisia ominaisuuksia.

Tämän tutkielman kirjoituksessa on hyödynnetty ChatGPT-tekoälytyökalua (versio 5.2) kielenhuoltoon ja LaTeX:in kirjoittamiseen.

2 Rekursiivinen hyötymalli oppimisen kanssa

Päätöksentekijä (jatkossa PT) toimii ympäristössä, jossa tuottoihin liittyy epävarmuutta sekä ulkoisista satunnaistekijöistä että tuntemattomasta parametrivakiosta $\theta \in \Theta$. Edellinen vastaa klassista stokastista riskiä, kun taas jälkimmäinen kuvaa Knightilaista epävarmuutta, jota mallinnetaan priorijoukkona. Jokaista parametrin θ arvoa vastaa oma todennäköisyysjakaumansa tapausavaruudessa Ω , mutta päätöksentekijä ei tiedä, mikä näistä jakaumista kuvaa todellisuutta. Tämän vuoksi hänen alkuperäisiä uskomuksiaan kuvaa priorijoukko \mathcal{M}_0 , joka ei koostu yksittäisestä jakaumasta vaan joukosta mahdollisia priorijakaumia. Tämä heijastaa lähtökohtaista malliepävarmuutta.

Ajan myötä päätöksentekijä saa informaatiota havainnoimalla signaalia, jonka dynamiikka riippuu parametrin θ todellisesta arvosta. Havaintojen perusteella uskomuksia päivitetään Bayesin säännön mukaisesti, jolloin muodostuu posteriorijoukko \mathcal{M}_t . Vastaavasti voidaan määritellä predikttiivisten priorien joukko \mathcal{P}_t , joka koostuu niistä tilajoukon Ω todennäköisyysjakaumista, jotka saadaan yhdistämällä parametrijakaumat \mathcal{M}_t ja signaalin ehdollinen jakauma.

Päätöksentekijä arvioi vaihtoehtojaan maxmin-periaatteen mukaisesti: kunkin vaihtoehdon arvo määräytyy sen odotushyödyn infimumina uskomusjoukon \mathcal{P}_t yli. Toisin sanoen päättäjät valitsee vaihtoehdon, joka maksimoi pahimman tapauksen odotushyödyn niiden todennäköisyysjakaumien suhteen, jotka kuuluvat hänen uskomusjoukkoonsa.

Tällainen päätöksentekijä ei kuitenkaan yleisesti ole aikakonsistenttia. Alkuhetkellä optimaalinen päätös ei välttämättä pysy optimaalisena myöhemmin, kun informaatiota on kertynyt ja posteriorijoukko on päivittynyt. Aikakonsistenssi voi rikkoutua, koska eri ajanhetkillä tarkasteltavat pahimman tapauksen jakaumat eivät välttämättä ole keskenään yhteensopivia.

Aikakonsistenssin varmistamiseksi priorijoukkoa voidaan laajentaa siten, että se on suljettu sellaisten yhdistelmien suhteen, joissa eri ajanhetkien ja eri mahdollisten tilojen ehdollisia todennäköisyysjakaumia voidaan liittää yhteen yhdeksi kokonaiseksi todennäköisyysjakaumaksi. Tällöin muodostuva priorijoukko on suorakulmainen. Suorakulmaisuus merkitsee, että päätöksentekijän uskomusrakenne on dynaamisesti johdonmukainen: posterioripäivitykset eri ajanhetkillä ovat keskenään yhteensopivia siten, että optimaaliset suunnitelmat säilyttävät voimassaolonsa ajan kuluessa. Tämä ominaisuus on keskeinen rekursiivisesti määriteltujen preferenssien ja robustin oppimisen analyysissa.

2.1 Suorakulmaisuus ja sen perustelu.

Olkoon (Ω, \mathcal{F}_T) mitallinen avaruus, missä Ω on tapausavaruus ja \mathcal{F}_T sitä vastaava σ -algebra. Olkoon $\mathcal{P} \subset \Delta(\Omega, \mathcal{F}_T)$ priorijoukko, missä

$$\Delta(\Omega, \mathcal{F}_T) = \{ P : P \text{ on todennäköisyysmitta avaruudessa } (\Omega, \mathcal{F}_T) \}.$$

Olkoon $p \in \mathcal{P}$. Olkoon p_t^ω mitan p säännöllinen ehdollinen todennäköisyys ajanhetkellä t , eli kuvaus

$$p_t^\omega(A) = p(A | \mathcal{F}_t)(\omega), \quad \omega \in \Omega, A \in \mathcal{F}_T.$$

Tällöin määritellään

$$\mathcal{P}_t(\omega) = \{p_t^\omega : p \in \mathcal{P}\}.$$

Määritelmä 1. Suorakulmainen priorijoukko. Joukko \mathcal{P} on suorakulmainen, jos kaikilla t ja $\omega \in \Omega$ pätee

$$\mathcal{P}_t(\omega) = \left\{ Q \in \Delta(\Omega, \mathcal{F}_T) \mid Q(A) = \int_{\Omega} Q_{t+1}^{\omega'}(A) m(d\omega') \text{ kaikilla } A \in \mathcal{F}_T \right\}, \quad (1)$$

missä

- m on todennäköisyysmitta siten, että $m \in \mathcal{Q}_t(\omega) \subset \Delta(\Omega, \mathcal{F}_{t+1})$,
- jokaiselle $\omega' \in \Omega$ pätee $Q_{t+1}^{\omega'} \in \mathcal{P}_{t+1}(\omega')$.

Toisin sanoen jokainen ajan t posteriori saadaan valitsemalla mielivaltainen seuraavan askeleen marginaalijakauma m sekä jokaiselle $\omega' \in \Omega$ jokin sallittu posteriorijakauma $p_{t+1}(\omega')$, ja liittämällä nämä yhteen. Muita posteriorijakaumia ei ole.

Lause 1 (Rekursiivinen konstruktio ja yksikäsitteisyys). *Määritellään rekursiivisesti $t = T - 1, \dots, 0$ siten, että $\mathcal{P}_t(\omega)$ koostuu kaikista mitoista, jotka toteuttavat yhtälön (1), missä $m \in \mathcal{Q}_t(\omega)$ ja $\omega' \mapsto Q_{t+1}^{\omega'}$ on mitallinen valinta siten, että $Q_{t+1}^{\omega'} \in \mathcal{P}_{t+1}(\omega')$ kaikilla $\omega' \in \Omega$.*

Tällöin on olemassa yksikäsitteinen priorijoukko \mathcal{P} , jonka yhden askeleen ehdolliset mitat ovat täsmälleen $\{\mathcal{Q}_t\}$ ja joka toteuttaa ehdon (1). Lisäksi \mathcal{P} on konvekssi. Lisäksi \mathcal{P} on konvekssi ja kompakti heikossa topologiassa, eli topologiassa, jossa $Q_n \rightarrow Q$ jos ja vain jos*

$$\int f dQ_n \rightarrow \int f dQ$$

kaikilla rajoitetuilla mitallisilla funktioilla f .

Todistuksen ydin. (\subseteq) Olkoon $Q \in \mathcal{P}$. Tällöin mitan Q ehdolliset jakaumat tuottavat hajotelman

$$Q(A) = \int_{\Omega} Q_{t+1}^{\omega'}(A) m(d\omega'), \quad A \in \mathcal{F}_T,$$

missä $m \in \mathcal{Q}_t(\omega)$ ja $\omega' \mapsto Q_{t+1}^{\omega'}$ on mitallinen valinta siten, että $Q_{t+1}^{\omega'} \in \mathcal{P}_{t+1}(\omega')$. Tästä seuraa, että $Q \in \mathcal{P}_t(\omega)$.

(\supseteq) Olkoon $m \in \mathcal{Q}_t(\omega)$ ja olkoon $\omega' \mapsto Q_{t+1}^{\omega'}$ mitallinen valinta siten, että $Q_{t+1}^{\omega'} \in \mathcal{P}_{t+1}(\omega')$ kaikilla $\omega' \in \Omega$. Tällöin peräkkäisten ehdollisten jakaumien perhe määrittää Ionescu-Tulcean laajennuslauseen nojalla yksikäsitteisen todennäköisyysmitan Q avaruudessa (Ω, \mathcal{F}_T) ; ks. [1, Proposition 2]. Tämä mitta toteuttaa esityksen

$$Q(A) = \int_{\Omega} Q_{t+1}^{\omega'}(A) m(d\omega'), \quad A \in \mathcal{F}_T,$$

joten $Q \in \mathcal{P}$.

Yksikäsitteisyys seuraa suoraan siitä, että annetut yhden askeleen ehdolliset jakaumat määräävät mitan yksikäsitteisesti Ionescu-Tulcean laajennuslauseen nojalla. Konveksisuus ja kompaktius seuraavat vastaavasti ehtojakaumien konveksisuudesta ja kompaktisuudesta sekä rakenteen sulkeutuvuudesta näiden yhdistelmien suhteen. \square

Lemma 1 (Dynaaminen min-iterointi). *Suorakulmaisuudesta seuraa, että jokaiselle rajoitetulle satunnaismuuttujalle X ja $t < T$ pätee*

$$\operatorname{ess\,inf}_{Q \in \mathcal{P}} E_Q[X \mid \mathcal{F}_t] = \operatorname{ess\,inf}_{m \in \mathcal{Q}_t(\omega)} E_m \left[\operatorname{ess\,inf}_{Q \in \mathcal{P}} E_Q[X \mid \mathcal{F}_{t+1}] \mid \mathcal{F}_t \right]$$

Lemmassa oleva yhtälö perustuu siihen, että suorakulmaisuuuden nojalla priorijoukko on suljettu yhdistelmien suhteen: jokainen $Q \in \mathcal{P}$ voidaan esittää muodossa

$$Q(A) = \int_{\Omega} Q_{t+1}^{\omega'}(A) m(d\omega'),$$

missä $m \in \mathcal{Q}_t(\omega)$ ja $Q_{t+1}^{\omega'} \in \mathcal{P}_{t+1}(\omega')$. Tämän esityksen ansiosta pahimman tapauksen odotusarvo voidaan minimoida kahdessa vaiheessa: ensin ajanhetkellä $t + 1$ ja sen jälkeen ajanhetkellä t .

Tämä johtaa aikakonsistenttiin preferenssirakenteeseen, koska päätöksentekijän valinta voidaan tehdä rekursiivisesti siten, että myöhemmät optimoinnit eivät muuta aiempia päätöksiä. Toisin sanoen optimointi ajanhetkellä t voidaan hajottaa peräkkäisiksi optimointiongelmiksi ajanhetkillä $t + 1, t + 2, \dots$, mikä on dynaamisen ohjelmoinnin peruseriaate.

Kolmion simpleksi -esimerkki. Kun $\mathcal{F}_1 = \sigma(\{G\}, \{R, B\})$ ja sallitaan $P(G) \in [g, \bar{g}]$ sekä $P(R \mid R \cup B) \in [\alpha, \beta]$, suorakulmaisuuus tuottaa kaikkien prosessin priorijakaumien parametriesityksen. Merkitään $m = P(G)$ ja $r = P(R \mid R \cup B)$. Tällöin

$$P(R \mid R \cup B) = \frac{P(R)}{P(R \cup B)} = \frac{P(R)}{1 - P(G)},$$

mistä seuraa

$$P(R) = (1 - P(G)) P(R \mid R \cup B) = (1 - m)r.$$

Vastaavasti

$$P(B) = (1 - m)(1 - r).$$

Siis

$$P(G) = m, \quad P(R) = (1 - m)r, \quad P(B) = (1 - m)(1 - r),$$

missä $(m, r) \in [g, \bar{g}] \times [\alpha, \beta]$.

Tässä simpleksi-esimerkissä tämä tarkoittaa, että vaikka alkuhetkellä pahin tapaus vastaisi tiettyä prioria P , myöhemmässä vaiheessa pahin tapaus voidaan valita erikseen ehdolla G ja ehdolla R , eikä sen tarvitse olla sama jakauma kuin alkuhetkellä. Tarkemmat tiedot suorakulmaisuudesta löytyvät lähteestä[8].

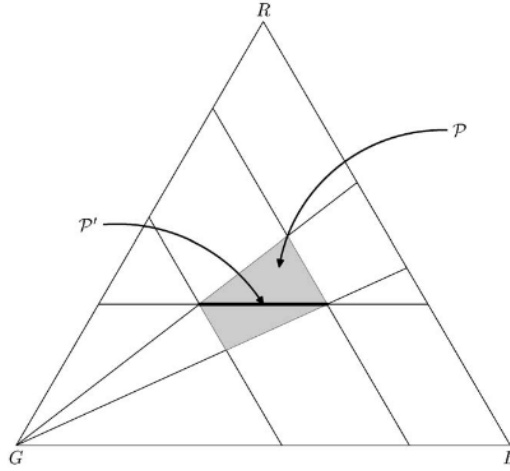


Fig. 1. Rectangular sets.

Kuva 1: Suorakulmaiset joukot. Kuvan mukainen kolmion simpleksi havainnollistaa priorijoukon rakenteen, jossa jokainen sallittu todennäköisyys voidaan muodostaa yhdistämällä marginaali- ja ehtojakaumia eri solmuissa. Kuva lähteestä[8]

2.2 Mallin rakenne

Olkoon $(\Omega, \mathcal{G}_\infty, P_0)$ todennäköisyysavaruus ja $W = (W_t)_{0 \leq t < \infty}$ yksiulotteinen Brownin liike, joka generoi filtraation $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$, missä (\mathcal{G}_t) on kasvava filtraatio ja $\mathcal{G}_\infty = \sigma(\bigcup_{t \geq 0} \mathcal{G}_t)$. Oletetaan, että kaikki todennäköisyysavarauudet ovat täydellisiä ja että kaikki filtraatiot on täydennetty tavanomaisella tavalla.¹ Mitan P_0 rooli on ainoastaan määrittellä ne tapahtumat joiden todennäköisyys on nolla.

Chen ja Epstein (jatkossa CE)[7] esittivät jatkuvan ajan mallin, jossa ambiguuteetti mallinnetaan joukolla todennäköisyysmittoja \mathcal{P}_0 , jotka ovat absoluuttisesti jatkuvia referenssimittaan P_0 nähden. Kullakin $P \in \mathcal{P}_0$ on siihen liittyvä tiheysprosessi η_t , joka määrittää mitan poikkeaman suhteessa P_0 :

$$\frac{dP^\eta}{dP_0} = \exp\left(-\int_0^T \eta_s dW_s - \frac{1}{2} \int_0^T |\eta_s|^2 ds\right).$$

Tämä muoto vastaa Girsanovin lauseessa esiintyvää eksponentiaalista tiheysprosessia, jonka avulla mitta P^η määritellään siten, että prosessi $W_t^\eta = W_t + \int_0^t \eta_s ds$ on Brownin liike mitan P^η alaisuudessa. Girsanovin lauseen mukaan, jos mitta P^η määritellään tiheysprosessin avulla muodossa

$$\frac{dP^\eta}{dP_0} = \exp\left(-\int_0^T \eta_s dW_s - \frac{1}{2} \int_0^T |\eta_s|^2 ds\right),$$

niin prosessi

$$W_t^\eta = W_t + \int_0^t \eta_s ds$$

¹Filtraation sanotaan olevan *täydennetty tavanomaisella tavalla*, jos se täyttää niin kutsutut *yleiset ehdot*: se on täydellinen (kaikki nollatapahtumat sisällytetään) ja oikealta jatkuva ($\mathcal{G}_t = \bigcap_{s > t} \mathcal{G}_s$ kaikilla $t \geq 0$). Näin varmistetaan, että Brownin liike ja stokastinen integraali ovat hyvin määriteltyjä.

on Brownin liike mitan P^η alaisuudessa.

Kullakin hetkellä t sallittujen ajautumaprosessin $\eta_t(\omega)$ arvojen joukkoa merkitään $\Xi_t(\omega) \subset \mathbb{R}^d$. Oletamme lisäksi, että joukot $\Xi_t(\omega)$ ovat konveksisia ja kompakteja. Nämä oletukset takaavat, että pahimman tapauksen ajautuma on hyvin määritelty ja että vastaavat optimointiongelmat saavuttavat ratkaisun. Tarkastelemme kuitenkin tässä tutkielmassa yksiulotteista tapausta $d = 1$, joka vastaa yhtä Brownin liikkeen komponenttia. Näin muodostuva joukko

$$\Xi = \{ \eta = (\eta_t) \mid \eta_t(\omega) \in \Xi_t(\omega) \, dt \otimes dP_0 \text{ m.v.} \}$$

määrittelee CE:n mukaisen todennäköisyysmittojen joukon

$$\mathcal{P}_0 = \{ P^\eta : \eta \in \Xi \}. \quad (2.2)$$

Tämä rakenne muodostaa CE-mallin perustan, jota Epstein ja Ji laajentavat sisällyttämällä siihen oppimisen tuntemattomasta parametrilla havaittavan signaalin avulla. Seuraavaksi Epstein ja Ji laajentavat Chenin ja Epsteinin mallia sisällyttämällä siihen oppimisdynamiikan stokastisen prosessin muodossa. Heidän kehikossaan PT tarkastelee yhtälössä (2.2) määriteltyä todennäköisyysmittojen joukkoa, joka kuvaa vaihtoehtoisia uskomuksia ajavan prosessin jakaumasta. Jokainen $\eta_t(\omega) \in \Xi_t(\omega)$ voidaan tulkita ehdollisena hypoteesina prosessin W ajautumasta hetkellä t . Girsanovin lauseen nojalla voidaan määritellä uusi Brownin liike mitan P^η alla:

$$dW_t^\eta = \eta_t \, dt + dW_t \quad (2.1)$$

missä W^η toimii ajavana satunnaisprosessina mitan P^η alla. Näin muodostuva kehikko mahdollistaa ajan tason mukaiset uskomukset ja PT:n preferenssit, jotka voidaan määritellä eksplisiittisesti yhtälöiden (2.3)-(2.4) avulla.

Yhtälöt (2.3) ja (2.4) määrittelevät päätöksentekijän preferenssit. Jotta nämä olisivat hyvin määriteltyjä, rajataan ensin satunnaismuuttujien luokka, jolla hyötyfunktio on hyvin määritelty. Epstein ja Ji seuraavat CE asetelmaa ja määrittelevät joukon $L(\tau)$ pysäytysajalle τ seuraavasti:

$$L(\tau) = \left\{ \xi \mid \xi \text{ on } \mathcal{G}_\tau\text{-mitallinen ja } \sup_{Q \in \mathcal{P}_0} \mathbb{E}_Q[|\xi|] < \infty \right\}.$$

Tämä ehto takaa, että satunnaismuuttuja ξ on integroituvissa kaikilla mahdollisilla todennäköisyysmitoilla $Q \in \mathcal{P}_0$, mikä puolestaan varmistaa, että odotusarvot ja infimumit seuraavissa yhtälöissä (2.3) ja (2.4) ovat hyvin määriteltyjä:

$$U_0(\xi) = \inf_{Q \in \mathcal{P}_0} \mathbb{E}_Q[\xi] = - \sup_{Q \in \mathcal{P}_0} \mathbb{E}_Q[-\xi] \quad (2.3)$$

$$U_t(\xi) = \text{ess inf}_{Q \in \mathcal{P}_0} \mathbb{E}_Q[\xi \mid \mathcal{G}_t] \quad (2.4)$$

Yhtälöt (2.3) ja (2.4) määrittelevät päätöksentekijän preferenssit siten, että kunkin vaihtoehdon arvo määräytyy sen odotushyödyn pienimpänä mahdollisena arvona kaikkien mittojen $Q \in \mathcal{P}_0$ yli. Ajan 0 hyöty $U_0(\xi)$, joka on esitetty yhtälössä (2.3), määrittää satunnaisten tuoton ξ arvon päätöksentekijälle, joka suhtautuu epävarmuuteen kielteisesti. Hän maksimoi hyödyn valitsemalla sen mitan $Q \in \mathcal{P}_0$, joka tuottaa alhaisimman odotetun arvon. Tämä muoto vastaa Gilboa–Schmeidlerin

(1989)[13] klassista maxmin-mallia, mutta jatkuvan ajan versiossa ja CE:n määrittelemällä priorijoukolla.

Yhtälö (2.4) puolestaan laajentaa tämän dynaamiseen asetelmaan. Siinä hyöty määritellään ehdollisesti kullakin ajanhetkellä t käyttäen oleellista infimumia, joka valitsee kaikkien mahdollisten mittojen joukosta sen, joka tuottaa pienimmän ehdollisen odotusarvon annetussa informaatioissa \mathcal{G}_t . Tämä varmistaa preferenssien aikakonsistenttiuden, koska Lemman 1 mukainen rekursiivinen esitys mahdollistaa optimoinnin hajottamisen peräkkäisiin vaiheisiin. Näin myöhemmät arviot eivät johda ristiriitaan aikaisempien valintojen kanssa.

Yhdessä yhtälöt (2.3) ja (2.4) siis kuvaavat robustia malliepävarmuuden sekä ajassa päivittyvän informaation huomioon ottavaa päätöksentekijää, jonka päätöksenteko on johdonmukaista.

Ehto

$$U_t(\xi) = \operatorname{ess\,inf}_{Q \in \mathcal{P}_0} \mathbb{E}_Q[\xi \mid \mathcal{G}_t]$$

voidaan tulkita robustina vastineena klassiselle iteroitujen odotusarvojen laille:

$$\mathbb{E}[\xi] = \mathbb{E}[\mathbb{E}[\xi \mid \mathcal{G}_t]].$$

Tässä asetelmassa odotusarvon roolin korvaa pessimistinen odotusarvo, joka määritellään infimumina kaikkien mahdollisten mittojen $Q \in \mathcal{P}_0$ yli. Näin saadaan aikakonsistentti hyötyrakenne:

$$U_t(\xi) = \operatorname{ess\,inf}_{Q \in \mathcal{P}_0} \mathbb{E}_Q[U_{t'}(\xi) \mid \mathcal{G}_t], \quad 0 \leq t < t'. \quad (2.5)$$

Rekursiivinen määritelmä (2.5) takaa aikakonsistenssin siten, että päätöksentekijän hyötyarviot eri ajanhetkillä ovat keskenään yhteensopivia. Nykyhetken arvio tulevasta hyödystä perustuu samaan mittojen joukkoon \mathcal{P}_0 ja samaan pessimistiseen odotusarvoperiaatteeseen kuin tulevaisuudessa tehty arvio. Näin ollen aiemmin tehty optimaalinen päätös säilyy optimaalisena myös myöhempinä ajanhetkinä.

Seuraavaksi määritellään mallioppimisen perusrakenne, jossa PT oppii ajan myötä tuntemattomasta parametrissa $\theta \in \Theta \subset \mathbb{R}$. Olkoon (Ω, \mathcal{F}) mitallinen avaruus, ja $\{\mathcal{F}_t\}_{t \geq 0}$ siihen liittyvä filtraatio. Oletetaan lisäksi joukko keskenään ekvivalentteja todennäköisyysmittoja $\{P^\mu : \mu \in \mathcal{M}_0\}$, joista jokainen vastaa mahdollista ennakkouskomusta parametrin θ jakaumasta. Mittajoukko \mathcal{M}_0 kuvastaa siten PT:n epävarmuutta parametrin arvoa kohtaan: jokainen $\mu \in \mathcal{M}_0$ määrittää todennäköisyysjakauman $\mu(A) = P^\mu(\{\theta \in A\})$ kaikille Borel-joukoille $A \subset \Theta$.

Oletetaan lisäksi, että mallissa on standardi Brownin liike $B = (B_t)_{t \geq 0}$, jonka generoima filtraatio on \mathcal{F}_t^B ja joka on riippumaton parametrin θ arvosta kunkin P^μ alla. Tämä Brownin liike toimii havaittavan signaaliprosessin $Z = (Z_t)_{t \geq 0}$ ajavana lähteenä. Signaali määritellään muodossa

$$Z_t = \int_0^t \theta ds + \int_0^t \sigma dB_s = \theta t + \sigma B_t, \quad (2.6)$$

missä $\sigma > 0$ on tunnettu vakio.

Koska päätöksentekijä havaitsee ainoastaan prosessin Z_t realisaatiot, määritellään filtraatio $\{\mathcal{G}_t\}$ signaalin Z generoimana. Olettamalla, että signaalin rakenne

tunnetaan, voidaan kullakin hetkellä t päivittää prioria $\mu \in \mathcal{M}_0$ Bayesin säännön avulla, jolloin saadaan posteriori μ_t . Tämä priorien päivitys tuottaa posteriorijoukkojen prosessin $(\mathcal{M}_t)_{t \geq 0}$, joka kuvaa PT:n ajan myötä tarkentuvaa, mutta edelleen epävarmaa uskomusrakennetta parametrin θ suhteen. Näin määritelty (Z_t) vastaa mallin viimeistä rakennuspalikkaa ja toimii oppimisen dynamiikan perustana.

Seuraavaksi [12] esittävät vaiheittain, miten heidän aiempaan dynaamisesti johdonmukaiseen epävarmuuskehikkoonsa (CE-malliin) tuodaan mukaan oppiminen stokastisen suodatusteorian avulla. Tämä edellyttää havaintoprosessin ja todennäköisyysmittojen "uudelleenrakentamista" siten, että päätöksentekijä voi päivittää uskomuksiaan tuntemattomasta parametrin θ reaaliajassa samalla kun säilytetään mallin robustius ja ajallinen johdonmukaisuus. Epstein ja Ji toteuttavat tämän askeleittain:

Askel 1: Innovaatioesityksen muodostaminen. Lähtökohtana on havaintoprosessi

$$dZ_t = \theta dt + \sigma dB_t,$$

jossa θ on tuntematon ajautuma. Standardin suodatusteorian (Liptser & Shiryaev)[18] mukaisesti tuntematon θ korvataan sen posteriori-estimaatilla $\hat{\theta}_t^\mu = \int \theta d\mu_t$, jolloin

$$dZ_t = \hat{\theta}_t^\mu(Z_t) dt + \sigma \left(dB_t + \frac{\theta - \hat{\theta}_t^\mu(Z_t)}{\sigma} dt \right). \quad (2.7)$$

Tämä voidaan kirjoittaa muodossa

$$dZ_t = \hat{\theta}_t^\mu(Z_t) dt + \sigma d\tilde{B}_t^\mu,$$

missä innovaatio-prosessi \tilde{B}_t^μ on määritelty

$$d\tilde{B}_t^\mu = dB_t + \frac{\theta - \hat{\theta}_t^\mu(Z_t)}{\sigma} dt.$$

Innovaatioprosessi \tilde{B}_t^μ on (\mathcal{G}_t) -adaptoitu, eli sen arvo ajanhetkellä t riippuu vain siihen mennessä kertyneestä informaatiosta (toisin sanoen \tilde{B}_t^μ on \mathcal{G}_t -mittallinen), ja se on standardi Brownin liike todennäköisyysavaruudessa $(\Omega, \mathcal{G}_\infty, P^\mu)$. Tämä vaihe erottaa havaintoprosessin deterministisen osan ja satunnaisen osan, jolloin oppiminen voidaan esittää stokastisena prosessina. Yhtälö voidaan myös kirjoittaa muodossa

$$d\tilde{B}_t^\mu = -\frac{1}{\sigma} \hat{\theta}_t^\mu(Z_t) dt + \frac{1}{\sigma} dZ_t,$$

mikä osoittaa, että (Z_t/σ) voidaan tulkita havaintoprosessina ja $(-\hat{\theta}_t^\mu(Z_t)/\sigma)$ vastaavana ajautumana CE-mallin yhtälössä (2.1).

Askel 2: Referenssimitan määrittäminen. Seuraavaksi Epstein ja Ji määrittelevät todennäköisyysmitan P_0 , jonka alla prosessi Z_t/σ on $\{\mathcal{G}_t\}$ -adaptoitu Brownin liike. Girsanovin lauseen avulla saadaan mittamuunnos

$$\frac{dP_0}{dP^\mu} \Big|_{\mathcal{G}_t} = \exp \left\{ -\frac{1}{2\sigma^2} \int_0^t (\hat{\theta}_s^\mu(Z_s))^2 ds - \frac{1}{\sigma} \int_0^t \hat{\theta}_s^\mu(Z_s) d\tilde{B}_s^\mu \right\},$$

joka voidaan kirjoittaa vaihtoehtoisesti myös muodossa:

$$\frac{dP_0}{dP^\mu} \Big|_{\mathcal{G}_t} = \exp \left\{ \frac{1}{2\sigma^2} \int_0^t (\hat{\theta}_s^\mu(Z_s))^2 ds - \frac{1}{\sigma^2} \int_0^t \hat{\theta}_s^\mu(Z_s) dZ_s \right\}.$$

Tämä muunnos standardoi kaikki mahdolliset priorimittaukset yhteiseen referenssimittaan P_0 nähden. Askel on analoginen CE-mallin perusmittaan, jossa W_t toimii yhteisenä Brownin liikkeenä.

Ensimmäinen esitys on Girsanovin lauseessa esiintyvä eksponentiaalinen martingaali. Toinen muoto saadaan korvaamalla differentiaali $d\tilde{B}_t^\mu = \frac{1}{\sigma} dZ_t - \frac{1}{\sigma} \hat{\theta}_t^\mu dt$, mutta tässä esityksessä integraali on kirjoitettu prosessin Z_t suhteen, eikä se enää muodosta eksponentiaalista martingaalia.

Askel 3: Askel 3: Referenssimittan muunnos ja ennustavien prioreiden muodostus.

Kun referenssimitta P_0 on määritelty sitä muunnetaan lievästi priorijoukon \mathcal{M}_0 avulla. Kullekin $\mu \in \mathcal{M}_0$ määritellään todennäköisyysmitta P_0^μ siten, että

$$\frac{dP_0^\mu}{dP_0} \Big|_{\mathcal{G}_t} = \exp \left\{ -\frac{1}{2\sigma^2} \int_0^t (\hat{\theta}_s^\mu(Z_s))^2 ds + \frac{1}{\sigma^2} \int_0^t \hat{\theta}_s^\mu(Z_s) dZ_s \right\}.$$

Tämä mittamuunnos säilyttää havaintoprosessin jakauman mutta vaihtaa sen ajautumatermiä. Girsanovin lauseen nojalla voidaan osoittaa, että prosessi

$$d\tilde{B}_t^\mu = -\frac{1}{\sigma} \hat{\theta}_t^\mu(Z_t) dt + \frac{1}{\sigma} dZ_t$$

on Brownin liike mitan P_0^μ alla. Vaikka $P^\mu \neq P_0^\mu$, ne indusoivat identtisen jakauman prosessille Z_t . Tämä johtuu siitä, että \tilde{B}_t^μ on $\{\mathcal{G}_t\}$ -adaptoitu Brownin liike sekä mittojen P^μ että P_0^μ suhteen, ja heikkojen ratkaisujen² yksikäsitteisyyden nojalla stokastisella differentiaaliyhtälöllä (Øksendal, Theorem 8.6.3 ja Lemma 5.3.1)[23] on tällöin sama jakauma kummankin mitan alla. Näin ollen mittamuunnos ei muuta havaintoprosessin jakaumaa, vaan ainoastaan sen ajautumadynamiikkaa.

Koska mallissa vain signaalin jakaumalla on merkitystä, mittamuutoksia P^μ ja P_0^μ ei tarvitse erottaa toisistaan. Tästä seuraa, että CE-kehikon elementit voidaan määrittää seuraavasti: referenssimitta P_0 ja Brownin liike W_t on määritelty askelissa 2 ja 1, ja lisäksi määritellään joukko

$$\Xi_t = \left\{ -\frac{\hat{\theta}_t^\mu}{\sigma} : \mu \in \mathcal{M}_0 \right\}, \quad \hat{\theta}_t^\mu = \int \theta d\mu_t, \quad (2.8)$$

²Heikolla ratkaisulla tarkoitetaan stokastisen differentiaaliyhtälön ratkaisua, jossa myös todennäköisyysavaruus, filtraatio ja Brownin liike kuuluvat osaksi ratkaisua. Tällöin yksikäsitteisyys tarkoittaa, että kaikkien ratkaisujen jakaumat ovat samat.

jossa integraali on tavallinen Lebesgue-integraali mitan μ_t suhteen (eli $\hat{\theta}_t^\mu = \mathbb{E}_{\mu_t}[\theta]$).

Tämän avulla muodostuu ennustavien prioreiden joukko $\mathcal{P}_0 = \{P_0^\mu : \mu \in \mathcal{M}_0\}$, jonka avulla päätöksentekijän hyöty voidaan kirjoittaa dynaamisesti johdonmukaisessa muodossa

$$U_t(\xi) = \operatorname{ess\,inf}_{P_0^\mu \in \mathcal{P}_0} \mathbb{E}_{P_0^\mu}[\xi \mid \mathcal{G}_t],$$

mikä kuvaa päätöksentekijän reaaliaikaista oppimista prioreihin liittyvän epävarmuuden puitteissa.

Seuraavaksi Epstein ja Ji korostavat, että heidän laajennuksensa CE-malliin perustuu keskeisesti niin sanotun heikon formuloinnin omaksumiseen vahvan formuloinnin sijasta. Tämä lähestymistapa ei sido analyysiä tiettyyn satunnaisprosessin polkuun, vaan määrittää prosessin jakauman mittateoreettisesti. Toisin sanoen heikko ratkaisu sallii sen, että Brownin liikettä ei tarkastella yksittäisenä shokkina, vaan keinona määrittellä signaalin $(Z_t - \theta t)/\sigma$ (askel 1) jakauma normaalina ehdollisesti parametrille θ .

Tämä muotoilu tekee mahdolliseksi mallintaa päätöksentekoa tilanteissa, joissa informaatio on epätäydellistä, kuten juuri tämän tutkielman käsittelemää oppimismallia. Øksendalin mukaan stokastisen differentiaaliyhtälön ratkaisut ovat heikosti yksikäsitteisiä, jolloin kaikki ratkaisut ovat identtisiä jakauman suhteen (aikaisemmin mainittu Lemma 5.3.1). Tätä periaatetta käytetään myös Girsanovin lauseen yhteydessä (Øksendal, Theorem 8.6.5), jossa mittamuunnoksen jälkeen saatu prosessi on Brownin liike uuden mitan alla ja siten sillä on sama jakauma kuin alkuperäisellä Brownin liikkeellä. Tähän tulokseen Epstein ja Ji tukeutuvat osoittaessaan, että heidän konstruktiollaan säilyy havaintoprosessin jakauma eri mittamuunnosten alla. Näin heikko formulointi antaa sekä matemaattisesti joustavan tavan yhdistää oppiminen ja robusti päätöksenteko heidän CE-mallissaan.

3 Optimaalinen oppiminen

3.1 Mallin rakenne ja yleinen ongelmanmuotoilu

PT valitsee yhden toimenpiteen joukosta $A = \{a_0, a_1, a_2\}$. Palkkiot ovat epävarmoja ja riippuvat tuntemattomasta parametrusta θ . Ennen valintaa PT voi hankkia lisätietoa θ :n arvosta seuraamalla signaaliprosessia Z , joka määräytyy yhtälön (2.6) mukaisesti, missä $\sigma > 0$ on annettu vakio. Oppiminen aiheuttaa kustannuksen $c > 0$ aikayksikköä kohti.

Olkoon $(\Omega, \mathcal{F}, (\mathcal{F}_t))$ taustalla oleva todennäköisyysavaruus, missä filtraatio (\mathcal{F}_t) kuvaa päätöksentekijän informaatiota ajanhetkeen t mennessä. Ellei toisin mainita, kaikki prosessit oletetaan (\mathcal{F}_t) -adaptoiduiksi.

Jos PT lopettaa oppimisen ajanhetkellä t , hänen ehdollinen odotettu hyötynsä on X_t . Ajatellaan, että X_t on se hyöty, jonka PT voi kyseisellä hetkellä saavuttaa valitsemalla optimitoimenpiteen joukosta A sen hetkisten uskomustensa perusteella.

PT on eteenpäin katsova, ja hänen ajanhetken 0 uskomuksensa signaalin tulevista jakaumista on esitetty todennäköisyysmittojen joukkona $\mathcal{P}_0 \subset \Delta(\Omega, \mathcal{G}_\infty)$, joka kuvaa epävarmuutta θ :n arvoa koskevien prioreiden välillä. PT tarkastelee pysäytysstrategioita, jotka esitetään satunnaisajankohtina τ , ja jotka rajoitetaan vaatimuksella $\sup_{Q \in \mathcal{P}_0} \mathbb{E}_Q \tau < \infty$. Pysäytysstrategioiden joukkoa merkitään merkinnällä Γ .

Koska PT on maxmin-päätöksentekijä, optimaalinen pysäytysajankohta määritetään ratkaisemalla

$$\max_{\tau \in \Gamma} \min_{P \in \mathcal{P}_0} \mathbb{E}_P(X_\tau - c\tau). \quad (3.1)$$

Määrittelemme tätä vastaavan arvofunktion V , joka esitetään myöhemmin signaalin funktiona. Oletamme jatkossa, että priorijoukko \mathcal{M}_0 voidaan parametrizoida kahden mahdollisen priorin välillä $m_0 \leq m \leq \bar{m}_0$ seuraavasti.

Oletamme, että tuntematon parametri θ voi saada arvot θ_0 tai θ_1 , jolloin parametrin tila-avaruus on $\Theta = \{\theta_0, \theta_1\}$. Merkitään δ_θ Diracin mittaa, joka keskittyy pisteeseen $\theta \in \Theta$.

$$\mathcal{M}_0 = \{\mu^m = (1 - m)\delta_{\theta_0} + m\delta_{\theta_1} : m_0 \leq m \leq \bar{m}_0\} \quad (3.2)$$

Kuten (3.2) osoittaa, priorijoukko voidaan parametrizoida m arvoihin, jotka kertovat, kuinka todennäköisenä pidetään suurempaa parametria θ_1 .

Kun signaaleja on havainnoitu aikaan t asti, päivittyy posteriorijoukko Bayesin säännön mukaisesti, jolloin

$$\mathcal{M}_t = \{(1 - m_t)\delta_{\theta_0} + m_t\delta_{\theta_1} : m_t \in [m_t, \bar{m}_t]\}. \quad (3.3)$$

Yhtälö (3.3) siis karakterisoi epävarmuuden ajanhetkellä t .

Päivityssääntö seuraa Bayesin kaavasta, kun signaaliprosessin todennäköisyys-tiheyksien suhdetta kuvataan funktiolla $\varphi(t, Z_t)$ jossa $\varphi(t, z)$ on signaaliprosessin generoima uskottavuusosamäärä, jonka tarkempi muoto johdetaan jäljempänä. Posteriorin alempi raja päivittyy muodossa

$$m_t = \frac{\frac{m_0}{1-m_0} \varphi(t, Z_t)}{1 + \frac{m_0}{1-m_0} \varphi(t, Z_t)}, \quad (3.4)$$

ja ylempi raja analogisesti

$$\bar{m}_t = \frac{\frac{\bar{m}_0}{1-\bar{m}_0} \varphi(t, Z_t)}{1 + \frac{\bar{m}_0}{1-\bar{m}_0} \varphi(t, Z_t)}. \quad (3.5)$$

Siten epävarmuus väliltä $[m_0, \bar{m}_0]$ päivittyy havainnon Z_t mukana välille $[m_t, \bar{m}_t]$.
Päivityskerroin $\varphi(t, z)$ on

$$\varphi(t, z) = \exp\left\{\frac{\theta_1 - \theta_0}{\sigma^2} z - \frac{(\theta_1^2 - \theta_0^2)}{2\sigma^2} t\right\}. \quad (3.6)$$

Yhtälö (3.6) kuvaa, kuinka havaintoprosessin realisaatio Z_t muuttaa posterioriuskoa ajautuman arvosta θ_0 kohti θ_1 .

Lipsterin ja Shiryaevin (kappale 9.4, esimerkki 1)[18] teoksessa esitetään samankaltainen todistus yhtälöille (3.4) ja (3.5). Johdetaan ne kuitenkin myös tässä, jotta niiden alkuperä ja rakenne käyvät lukijalle selviksi.

Todistus. Osoitetaan, että posterioritodennäköisyys havaintomallissa (3.3) saa muodon (3.4). Tarkastellaan havaintomallia

$$dZ_t = \theta dt + \sigma dB_t, \quad \theta \in \{\theta_0, \theta_1\}.$$

Alkupriorit ovat

$$\mathbb{P}(\theta = \theta_1) = m_0, \quad \mathbb{P}(\theta = \theta_0) = 1 - m_0.$$

Merkitään havaintorealisaatiota $Z_t = z$. Girsanovin lauseen avulla kahden mahdollisen ajautuman tuottama uskottavuuskerroin on yhtälön (3.6) mukaisesti

$$\frac{dP_{\theta_1}}{dP_{\theta_0}} \Big|_{\mathcal{F}_t} = \exp\left(\frac{\theta_1 - \theta_0}{\sigma^2} z - \frac{\theta_1^2 - \theta_0^2}{2\sigma^2} t\right).$$

Bayesin lauseen mukaan posteriori todennäköisyys on

$$\mathbb{P}(\theta = \theta_1 \mid Z_t = z) = \frac{m_0 \varphi(t, z)}{(1 - m_0) + m_0 \varphi(t, z)}.$$

Tällöin posteriorimitta on $(1 - m_t)\delta_{\theta_0} + m_t\delta_{\theta_1}$, mikä on \mathcal{M}_t :n rakenteen (3.3) mukainen.

Posteriori voidaan kirjoittaa muodossa

$$\frac{m_t}{1 - m_t} = \frac{m_0}{1 - m_0} \varphi(t, Z_t).$$

Ratkaisemalla tästä m_t saadaan

$$m_t = \frac{\frac{m_0}{1-m_0} \varphi(t, Z_t)}{1 + \frac{m_0}{1-m_0} \varphi(t, Z_t)}.$$

Tämä vastaa (3.4) muotoa. Yläraja käsitellään vastaavalla tavalla. \square

Oletetaan, että toteutuva parametri on $\theta \in \{\theta_0, \theta_1\}$. Jos päätöksentekijä valitsee ajanhetkellä t toimenpiteen a_i , sen tuotto tunnetulla parametrilla θ_j merkitään $u(a_i, \theta_j)$, ja oletetaan aina ei-negatiiviseksi. Funktio $u(a_i, \theta_j)$ kuvaa toimenpiteen a_i tuottoa tilassa θ_j . Mahdollinen riskinkarttaminen on sisällytetty tähän hyötyfunktioon, joten kun θ_j on tunnettu, $u(a_i, \theta_j)$ on deterministinen suure.

Oletetaan lisäksi, että jokaiselle $i, j = 0, 1, i \neq j$, pätee

$$u(a_j, \theta_j) = u(a_i, \theta_i) > u(a_j, \theta_i). \quad (3.7)$$

Tämä tarkoittaa, että a_0 on parempi valinta, mikäli $\theta = \theta_0$, ja a_1 on parempi valinta, mikäli $\theta = \theta_1$. Toisin sanoen, kumpikin toimenpide on paras vain oman parametrinsa toteutuessa.

Kolmas toimenpide a_2 ei riipu parametrin arvosta θ , ja sitä voidaan tulkita ns. varavaihtoehdoksi. Sen tuotto ei ole epävarma, koska mallissa epävarmuus koskee ainoastaan parametria θ . Merkitään tällöin

$$u(a_2, \theta) = u_2 \quad \forall \theta \in \{\theta_0, \theta_1\}.$$

Toimenpide a_2 toimii vertailukohtana, joka määrää päätöksentekijän hyötytason alarajan. Jotta se ei olisi triviaalisti optimaalinen, oletetaan lisäksi, että sen tuotto on molemmissa parametrialivoissa pienempi kuin parhaalla tilannesidonnaisella toimenpiteellä, eli

$$u_2 < u(a_i, \theta_i), \quad i = 0, 1.$$

Kun aika t etenee ja havaintoihin perustuva posteriorijoukko on \mathcal{M}_t , päätöksentekijän Gilboa-Schmeidler tyyppinen[13] a_i :n hyöty on

$$\min_{\mu \in \mathcal{M}_t} \int u(a_i, \theta) d\mu.$$

Jos päätöksentekijä valitsee ajanhetkellä t parhaan toimenpiteen, hänen hetkellinen tuottoprofilinsa on

$$X_t = \max \left\{ \min_{\mu \in \mathcal{M}_t} \int u(a_0, \theta) d\mu, \min_{\mu \in \mathcal{M}_t} \int u(a_1, \theta) d\mu, u_2 \right\}. \quad (3.9)$$

Tarkastellaan seuraavia kahta lisäoletusta:

- **Tuottosymmetria:** $u(a_0, \theta_1) = u(a_1, \theta_0)$.
- **Ei riskitöntä vaihtoehtoa:** $u_2 \leq u(a_i, \theta_j)$ kaikille $i \neq j, i, j \in \{0, 1\}$.

Ensimmäinen oletus lisää symmetriaa verrattuna yhtälöön (3.6). Toinen tarkoittaa, että a_2 on (heikosti) pienempi kuin a_0 ja a_1 kummassakin parametrialvossa, jolloin sitä ei koskaan valita optimaalisesti ja se voidaan jättää huomiotta. Tällöin ongelma supistuu kahden toimenpiteen valintaan.

Seuraavaksi esitellään mallit, mitä varten nämä oletukset tehtiin. Ellsbergin urnat[5] toteuttavat tuottosymmetriaa koskevan vaatimuksen, kun taas hypoteesientestausasetelmassa päätöksentekijällä ei ole käytettävissään varmaa vaihtoehtoa, vaan hänen on valittava vaihtoehtojen a_0 ja a_1 välillä.

3.2 Oppiminen ja Ellsbergin urnat

Tarkastellaan kahta urnaa, joista kumpikin sisältää punaisia ja sinisiä palloja. Ensimmäinen on riskillinen urna, jossa punaisten pallojen osuus on tunnetusti $\frac{1}{2}$. Toinen on epäselvä urna, jonka värikoostumus on tuntematon. Merkitään punaisten pallojen tuntematonta osuutta $\frac{1}{2} + \theta$, missä θ kuvaa poikkeamaa tasajaosta. Näin ollen $\theta > 0$ tarkoittaa, että punaisia palloja on enemmän, $\theta < 0$ että vähemmän, ja $\theta = 0$ vastaa riskillistä urnaa.

Päätöksentekijä voi valita, lyökö vetoa punaisen tai sinisen arvonnasta riskillisestä vai epäselvästä urnasta. Ellsbergin paradoksin[5] mukaisesti ilman oppimista epävarmuutta karttava päätöksentekijä valitsee vedon riskillisestä urnasta riippumatta väristä.

Mallin mukaisesti asetetaan θ -arvot

$$\theta_0 = -\alpha, \quad \theta_1 = \alpha, \quad (3.10)$$

missä $\alpha \in (0, \frac{1}{2})$ mittaa mahdollista harhaa punaisten hyväksi tai niitä vastaan.

Merkitään a_2 vedoksi riskillisestä urnasta, ja a_1 (a_0) vedoksi punaisesta (sinisestä) pallosta epäselvästä urnasta. Mahdolliset harhat ovat siten $\frac{1}{2} - \alpha$ ja $\frac{1}{2} + \alpha$, mutta ei tiedetä kumpi on todennäköisempi. Epävarmuus kuvataan priorijoukolla \mathcal{M}_0 , jonka muoto on kuten yhtälössä (3.3). Lisäksi oletetaan, että todennäköisyysväli parametrille α (harha punaisen hyväksi) on

$$m_0 + \bar{m}_0 = 1,$$

tai vastaavasti

$$m_0 = \frac{1 - \varepsilon}{2}, \quad \bar{m}_0 = \frac{1 + \varepsilon}{2}, \quad (3.11)$$

missä $\varepsilon \in (0, 1)$ kuvaa epävarmuuden astetta.

Tällöin pienin todennäköisyys punaisen hyväksi harhalle on sama kuin sinisen hyväksi harhalle, mikä tarkoittaa välinpitämättömyyttä punaiseen ja siniseen vetoamisen välillä ajanhetkellä $t = 0$. Parametri ε voidaan tulkita epävarmuuden mittana: mitä suurempi ε , sitä laajempi prioriväli ja suurempi epävarmuus harhan suunnasta. Erityistapauksessa $\varepsilon = 0$ priorijoukko \mathcal{M}_0 supistuu yksittäiseksi pisteeksi, jolloin päätöksentekijä on puhtaasti bayesilainen eikä kohtaa malliepävarmuutta, vaan pelkästään satunnaisuutta.

Toisaalta α mittaa epävarmuutta itse harhan suuruudesta eli priorivarianssia: kun $\alpha = 0$, epäselvän urnan koostumus on identtinen riskillisen kanssa.

Lopuksi määritetään tuotot u . Kaikki vedot on normitettu siten, että voitto vastaa arvoa 1 ja tappio arvoa 0. Koska epäselvän urnan koostumus on $\frac{1}{2} \pm \alpha$, odotetut tuotot ovat

$$u(a_0, -\alpha) = u(a_1, \alpha) = \alpha + \frac{1}{2}, \quad u(a_0, \alpha) = u(a_1, -\alpha) = \alpha - \frac{1}{2}, \quad u_2 = \frac{1}{2}. \quad (3.12)$$

Yhtälöiden (3.10)–(3.12) oletukset täyttävät kappaleessa 3.1 esitetyn mallin ehdot. Tämä voidaan nähdä seuraavasti. Yleisen mallin kappaleessa 3.1 päätöksentekijällä on kaksi mahdollista parametria $\theta_0 < \theta_1$, priorijoukko (3.2) sekä kolme mahdollista toimenpidettä a_0, a_1, a_2 . Hyötyfunktioilta $u(a_i, \theta_j)$ edellytettiin, että kukin a_i on

optimaalinen vain oman parametrinsa toteutuessa ja että kolmas vaihtoehto a_2 on vakiohyötyinen ja tuottaa vähemmän kuin kumpikaan ensimmäisistä vaihtoehdoista.

Ellsbergin urnien tapauksessa yhtälöissä (3.10)–(3.12) annetut oletukset muodostavat juuri tämän rakenteen erityistapauksen. Parametrit $\theta_0 = -\alpha$ ja $\theta_1 = \alpha$ vastaavat kahta mahdollista harhan suuntaa. Priorijoukko $[m_0, \bar{m}_0] = [\frac{1-\varepsilon}{2}, \frac{1+\varepsilon}{2}]$ on muodoltaan sama kuin yleisessä mallissa määritelty \mathcal{M}_0 ja hyötyehdot

$$u(a_i, \theta_i) > u(a_j, \theta_i), \quad i \neq j, \quad \text{sekä} \quad u_2 < u(a_i, \theta_i),$$

toteutuvat eksplisiittisesti yhtälöstä (3.12). Näin ollen Ellsbergin asetelma on täsmällinen spesifikaatio yleisestä mallista, jossa epävarmuus koskee parametrin θ mahdollista harhaa, ja päätöksentekijän tehtävä on ratkaista optimaalinen päätös tämän epävarmuuden ja oppimisen suhteen.

Seuraavaksi esitetään tässä ehdollinen tuottolauseke $X_t = X(Z_t)$, joka seuraa edellä määritellystä mallista:

$$X(Z_t) = \begin{cases} \left(\frac{1}{2} + \alpha\right) - \frac{2\alpha}{1 + \frac{1-\varepsilon}{1+\varepsilon}\varphi(Z_t)}, & \text{jos } Z_t > \frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right), \\ \left(\frac{1}{2} - \alpha\right) + \frac{2\alpha}{1 + \frac{1+\varepsilon}{1-\varepsilon}\varphi(Z_t)}, & \text{jos } Z_t < -\frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right), \\ \frac{1}{2}, & \text{muussa tapauksessa.} \end{cases} \quad (3.13)$$

missä $\varphi(z) = \exp(2\alpha z/\sigma^2)$. Jos Z_t saa suuren positiivisen arvon, optimaalinen veto on punainen (epävarmasta urnasta), ja jos Z_t on suuri negatiivinen, veto on sininen. Väliarvoilla Z_t :n perusteella ei ole vielä riittävästi näyttöä harhan suunnasta, ja tällöin riskillinen urna (vakiohyöty u_2) on paras valinta. Erityisesti alkuhetkellä $Z_0 = 0$, jolloin päätöksentekijä toimii kuten Ellsbergin klassisen kokeen osallistuja: hän valitsee riskillisen urnan, koska oppimista ei vielä ole tapahtunut.

Perustellaan myös hiukan, mistä Epstein tämän muodon on johtanut. Määritelmän (3.9) mukaan

$$X_t = \max \left\{ \min_{\mu \in \mathcal{M}_t} \int u(a_0, \theta) d\mu, \min_{\mu \in \mathcal{M}_t} \int u(a_1, \theta) d\mu, u_2 \right\}.$$

Ellsbergin spesifikaatiolla (3.12) $u(a_1, \theta_1) = u(a_0, \theta_0) = \frac{1}{2} + \alpha$, $u(a_1, \theta_0) = u(a_0, \theta_1) = \frac{1}{2} - \alpha$, $u_2 = \frac{1}{2}$, jolloin

$$\min_{\mu \in \mathcal{M}_t} \int u(a_1, \theta) d\mu = \min_{m \in [m_t, \bar{m}_t]} \left(\left(\frac{1}{2} - \alpha\right) + 2\alpha m \right) = \left(\frac{1}{2} - \alpha\right) + 2\alpha m_t,$$

$$\min_{\mu \in \mathcal{M}_t} \int u(a_0, \theta) d\mu = \min_{m \in [m_t, \bar{m}_t]} \left(\left(\frac{1}{2} + \alpha\right) - 2\alpha m \right) = \left(\frac{1}{2} + \alpha\right) - 2\alpha \bar{m}_t,$$

koska ensimmäinen on kasvava ja jälkimmäinen vähenevä muuttujan m suhteen.

Posteriorien suljetut muodot: Yhtälöistä (3.4)–(3.6) sekä (3.11) (eli $m_0 = \frac{1-\varepsilon}{2}$, $\bar{m}_0 = \frac{1+\varepsilon}{2}$, $\theta_0 = -\alpha$, $\theta_1 = \alpha$) saadaan

$$\varphi(Z_t) = \exp\left(\frac{2\alpha}{\sigma^2} Z_t\right), \quad m_t = \frac{\frac{1-\varepsilon}{1+\varepsilon} \varphi(Z_t)}{1 + \frac{1-\varepsilon}{1+\varepsilon} \varphi(Z_t)}, \quad \bar{m}_t = \frac{\frac{1+\varepsilon}{1-\varepsilon} \varphi(Z_t)}{1 + \frac{1+\varepsilon}{1-\varepsilon} \varphi(Z_t)}.$$

Sijoittamalla nämä saadaan kaksi ehdokasta

$$V_1(Z_t) = \left(\frac{1}{2} - \alpha\right) + 2\alpha m_t, \quad V_0(Z_t) = \left(\frac{1}{2} + \alpha\right) - 2\alpha \bar{m}_t.$$

Tässä $V_i(Z_t)$ kuvaa toiminnon i posterioriin perustuvaa odotettua tuottoa, ei prosessin realisaatiota $X(Z_t)$. Tehdään muunnos ja saadaan ekvivalentit muodot

$$V_1(Z_t) = \left(\frac{1}{2} - \alpha\right) + \frac{2\alpha}{1 + \frac{1+\varepsilon}{1-\varepsilon} \varphi(Z_t)^{-1}} = \left(\frac{1}{2} + \alpha\right) - \frac{2\alpha}{1 + \frac{1-\varepsilon}{1+\varepsilon} \varphi(Z_t)},$$

$$V_0(Z_t) = \left(\frac{1}{2} + \alpha\right) - \frac{2\alpha}{1 + \frac{1+\varepsilon}{1-\varepsilon} \varphi(Z_t)}.$$

Missä kohtaa päätöksentekijä valitsee minkäkin ehdokkaan? Vertaamalla $V_1(Z_t)$:a ja $u_2 = \frac{1}{2}$ saadaan

$$V_1(Z_t) \geq \frac{1}{2} \iff Z_t \geq \frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right).$$

Vastaavasti $V_0(Z_t) \geq \frac{1}{2}$ täsmälleen kun

$$Z_t \leq -\frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right).$$

Muussa tapauksessa $\frac{1}{2}$ dominoi.

Edellisen perusteella siis saadaan

$$X_t = X(Z_t) = \begin{cases} \left(\frac{1}{2} + \alpha\right) - \frac{2\alpha}{1 + \frac{1+\varepsilon}{1-\varepsilon} \varphi(Z_t)}, & \text{jos } Z_t > \frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right), \\ \left(\frac{1}{2} - \alpha\right) + \frac{2\alpha}{1 + \frac{1-\varepsilon}{1+\varepsilon} \varphi(Z_t)}, & \text{jos } Z_t < -\frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right), \\ \frac{1}{2}, & \text{muussa tapauksessa,} \end{cases}$$

mikä on täsmälleen (3.13).

Seuraavaksi johdetaan eksplisiittinen ratkaisu optimaalisen pysäytysajan ongelmaan, joka täyttää yhtälöt (3.10)-(3.12) ja perustuu yleiseen malliin (3.2). Määritellään ensin apufunktio

$$l(r) = 2 \log\left(\frac{r}{1-r}\right) - \frac{1}{r} + \frac{1}{1-r}, \quad r \in (0, 1), \quad (3.14)$$

ja olkoon

$$l(\hat{r}) = \frac{2\alpha^3}{c\sigma^2}. \quad (3.15)$$

Siten \hat{r} on yksikäsitteisesti määritelty, sillä

$$l(0) = -\infty, \quad l\left(\frac{1}{2}\right) = 0, \quad l(1) = \infty$$

ja $l(r)$ on monotonisesti kasvava.

Lause 2. 1. Optimaalinen pysäytysaika on $\tau^* = 0$ täsmälleen silloin, kun $\frac{1+\varepsilon}{2} \geq \hat{r}$, jolloin $X_{\tau^*} = X_0 = \frac{1}{2}$.

2. Jos $\frac{1+\varepsilon}{2} < \hat{r}$, niin optimaalinen pysäytysaika $\tau^* > 0$ on määritelty ehdolla

$$\tau^* = \min\{t \geq 0 : |Z_t| \geq \bar{z}\},$$

missä

$$\bar{z} = \frac{\sigma^2}{2\alpha} \left[\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) + \log\left(\frac{\bar{r}}{1-\bar{r}}\right) \right] > 0, \quad (3.16)$$

ja \bar{r} , joka toteuttaa epäyhtälön $\hat{r} < \bar{r} < 1$, on ainoa ratkaisu yhtälölle

$$l(r) + l\left(\frac{1+\varepsilon}{2}\right) = \frac{4\alpha^3}{c\sigma^2}. \quad (3.17)$$

Pysäytyksen tapahtuessa vedon valinta määräytyy havainnon perusteella: jos $Z_{\tau^} \geq \bar{z}$, valitaan veto punaisesta (epävarmasta) urnasta, ja jos $Z_{\tau^*} \leq -\bar{z}$, valitaan veto sinisestä. Riskillistä urnaa ei tällöin enää valita, sillä oppiminen on tuottanut riittävästi näyttöä harhan suunnasta.*

Lisäksi, jos $\varepsilon < \varepsilon' < 2\hat{r} - 1$, ja $\tau^{'}$ on tätä suurempaa epävarmuutta vastaava optimaalinen pysäytysaika, niin $\tau^{*'} \geq \tau^*$. Toisin sanoen, mitä suurempi epävarmuus (leveämpi prioriväli $[m_0, \bar{m}_0]$), sitä pidempään päätöksentekijä odottaa ennen pysäyttämistä ja päätöksen tekemistä.*

Lauseen todistus esitetään kappaleessa 5.3.

Edellä esitetyt kaksi tapausta määräytyvät epävarmuusparametrin ε ja raja-arvon \hat{r} suhteellisen suuruuden mukaan. Parametri ε mittaa epävarmuuden astetta, kun taas \hat{r} kasvaa suhteen $\alpha^3/(c\sigma^2)$ myötä. Erityisesti α vaikuttaa positiivisesti siihen, kuinka arvokasta on tietää todellisen harhan suunta. Suurempi α kasvattaa tuottoeroa parametriarvojen θ_0 ja θ_1 välillä, jolloin virheellisen toimenpiteen kustannus kasvaa ja informaation arvo lisääntyy.

Tapaus 1) vastaa tilannetta, jossa epävarmuus on suuri suhteessa odotettuihin tuottoihin (ottaen huomioon myös näytteenoton kustannuksen ja signaalin varianssin). Tällöin oppiminen ei ole kannattavaa, ja päätöksentekijä valitsee heti riskillisen urnan ilman lisähavaintoja.

Sen sijaan tapauksessa 2), jossa epävarmuus on riittävän pieni, oppiminen on ainakin osittain optimaalista. Tämä sisältää myös rajatapauksen $\varepsilon = 0$, jossa epävarmuutta ei ole ja malli redusoituu tavanomaiseen bayesilaiseen asetelmaan. Näin ollen voidaan todeta, että oppimisen hylkääminen on optimaalista vain silloin, kun epävarmuus (parametrin ε mittaamana) on riittävän suuri.

Tapaus 2) siis tarkoittaa, että havaintoja kannattaa kerätä niin kauan kuin signaaliprosessi Z_t pysyy jatkoalueella $(-\bar{z}, \bar{z})$. Tällä oppimisalueella on kaksi huomionarvoista piirrettä. Ensinnäkin, kun Z_t saavuttaa välin jonkin rajan, oppiminen päättyy ja päätöksentekijä tekee vedon epävarmasta urnasta. Toisin sanoen, riskillinen urna valitaan täsmälleen silloin, kun oppiminen ei ole enää optimaalista. Toiseksi, otantaa jatketaan pidempään (eli oppimisen kesto kasvaa) epävarmuuden lisääntyessä, mikäli $\varepsilon < 2\hat{r} - 1$. Kun epävarmuus saavuttaa tämän tason, 1) tapaus taas realisoituu ja oppiminen ei ole enää lainkaan kannattavaa.

Intuitiivisesti epävarmuuden (parametrin ϵ) kasvu vähentää päätöksentekijän hakkuutta oppia. PT:n alkuperäinen uskomus sallii vain kaksi mahdollista hypoteesia signaalin harhasta, α ja $-\alpha$. Hän voi oppimalla selvittää, kumpi näistä on tosi, mutta oppiminen on kustannuksellista. Koska päätöksentekijä on varovainen ja maksimoi pahimman tapauksen odotetun hyödyn, hän arvioi vaihtoehtoja epäsuotuisimman todennäköisyysmitan suhteen. Tämä vastaa tilannetta, jossa signaaliprosessi on mahdollisimman vähän informatiivinen päätöksenteon kannalta.

Tämä johtaa kahteen vaikutukseen. Ensinnäkin signaalien informaatioisältö pienenee, jolloin oppiminen on vähemmän tehokasta ja kestää kauemmin. Toiseksi, koska epävarmuus pienentää odotettua etukäteishyötyä epävarman vaihtoehdon valinnasta, signaalit eivät enää muuta päätöksentekijän preferenssiä yhtä helposti. Näin ollen epävarmuuden kasvaessa houkutus jatkaa oppimista pienenee. Näiden seurauksena oppiminen on kannattavaa vain rajallisen ajan, ja mallissa esiintyy kynnsarvo ϵ^* , jota suuremmalla epävarmuudella lisäoppiminen ei enää tuota riittävää hyötyä suhteessa sen kustannukseen.

Pienemmillä epävarmuuden arvoilla nousee luonnollisesti esille toinen kysymys: miksi ei ole koskaan optimaalista oppia hetken aikaa ja sitten pysähtyä ja valita riskinen vaihtoehto? Tämä liittyy mallin perusoletukseen priorijoukosta \mathcal{M}_0 . Koska \mathcal{M}_0 :ssa uskotaan harhan olevan varmasti joko α tai $-\alpha$, tilanteet, joissa havaittu signaali on lähellä nollaa, tulkitaan puhtaasti kohinaksi. Näin ollen, jos olisi optimaalista valita riskinen vaihtoehto jossakin hetkessä t , se olisi ollut optimaalista jo alussa.

Jos kuitenkin \mathcal{M}_0 :aa yleistetään siten, että sen prioreilla on yhteinen äärellinen määrittelyjoukko, joka sisältää myös nollaharhan, tämä tulos voisi muuttua. Tällöin oppiminen ja riskillisen vaihtoehdon valinta voisivat olla yhteensopivia päätöksiä, koska nollaharha olisi mahdollinen ja signaaleilla olisi siten erilainen tulkinta.

Seuraavaksi voidaan hyödyntää edellä johdettua suljetun muodon ratkaisua arvioidaksemme epävarmuuden vaikutusta optimaaliseen oppimiseen. Rajoitutaan tapaukseen, jossa epävarmuuden aste ϵ on välillä $[0, 2\bar{r} - 1]$, jolloin ainakin jonkin verran oppimista on optimaalista. Merkitään P^θ :lla stokastisen prosessin (Z_t) todennäköisyysjakaumaa, kun todellinen harha on θ . Tällöin keskimääräinen otanta-aika (tai näytteen pituus) τ^* määräytyy erikseen. Se on odotusaika, joka kuluu ennen kuin prosessi saavuttaa rajan \hat{z} . Koska (Z_t) käyttäytyy tässä ajautumallisen Brownin liikkeen tavoin, voidaan hyödyntää tunnettuja tuloksia ensimmäisen pysäytysajan odotusarvosta (ks. esim. Borodin ja Salminen)[4]. Näiden tulosten mukaan keskimääräinen otanta-aika, eli $\mathbb{E}^\theta[\tau^*]$, poistumiselle alkuasemasta $x = 0$ välille $(-\bar{z}, \bar{z})$ ajautumaparametrilla θ ja volatilitteetilla σ on

$$\mathbb{E}^\theta[\tau^*] = \begin{cases} \left(\frac{\bar{z}}{\sigma}\right)^2 \frac{\tanh\left(\frac{\theta\bar{z}}{\sigma^2}\right)}{\frac{\theta\bar{z}}{\sigma^2}}, & \theta \neq 0, \\ \left(\frac{\bar{z}}{\sigma}\right)^2, & \theta = 0. \end{cases} \quad (3.18)$$

Jotta tulee myös selväksi, mistä Epstein tämän muodon on saanut, niin tulos voidaan johtaa ratkaisemalla vastaava reuna-arvotehtävä diffuusion generaattorille, kuten seuraavassa esitetään.

Olkoon $X_t = x + \theta t + \sigma W_t$ ja

$$\tau_{\bar{z}} := \inf\{t \geq 0 : |X_t| = \bar{z}\}, \quad \bar{z} > 0.$$

Tällöin, alkuasemasta $x = 0$,

$$\mathbb{E}^\theta[\tau^*] = \begin{cases} \left(\frac{\bar{z}}{\sigma}\right)^2 \frac{\tanh\left(\frac{\theta\bar{z}}{\sigma^2}\right)}{\frac{\theta\bar{z}}{\sigma^2}}, & \theta \neq 0, \\ \left(\frac{\bar{z}}{\sigma}\right)^2, & \theta = 0. \end{cases} \quad (*)$$

Todistus. Asetetaan $u(x) = \mathbb{E}^x[\tau^*]$ kun $x \in (-\bar{z}, \bar{z})$. Diffuusion generaattori on tässä tapauksessa

$$L = \theta \frac{d}{dx} + \frac{\sigma^2}{2} \frac{d^2}{dx^2}.$$

Dynkinin kaavaa³ soveltamalla saadaan, että odotettu pysäytysaika määräytyy seuraavasta reuna-arvotetävästä

$$Lu(x) = -1, \quad x \in (-\bar{z}, \bar{z}), \quad u(\pm\bar{z}) = 0.$$

Kirjoitetaan $u'' + \beta u' = -\gamma$ missä $\beta = \frac{2\theta}{\sigma^2}$ ja $\gamma = \frac{2}{\sigma^2}$. Asettamalla $v = u'$ saadaan $v' + \beta v = -\gamma$, siis $v(x) = Ce^{-\beta x} - \gamma/\beta = Ce^{-\beta x} - 1/\theta$. Integroimalla:

$$u(x) = K - \frac{x}{\theta} - \frac{C}{\beta} e^{-\beta x}.$$

Reunaehtoa $u(\pm\bar{z}) = 0$ soveltamalla saadaan ratkaisuksi

$$C = -\frac{\bar{z}}{\theta} \frac{\beta}{\sinh(\beta\bar{z})}, \quad K = \frac{\bar{z}}{\theta} + \frac{\bar{z}}{\theta} \frac{e^{-\beta\bar{z}}}{\sinh(\beta\bar{z})}.$$

Sijoittamalla $x = 0$ saadaan

$$u(0) = \frac{\bar{z}}{\theta} \tanh\left(\frac{\theta\bar{z}}{\sigma^2}\right) = \left(\frac{\bar{z}}{\sigma}\right)^2 \frac{\tanh\left(\frac{\theta\bar{z}}{\sigma^2}\right)}{\frac{\theta\bar{z}}{\sigma^2}},$$

mikä vastaa yhtälön (*) ensimmäistä riviä. Ottamalla raja $\theta \rightarrow 0$ saadaan yhtälön (*) toinen rivi. \square

Huomataan (yhtälöstä (3.18)), että suurempi epävarmuus ϵ kasvattaa odotettua otantajaksoa. Lisäksi pätee, että $\theta Z_{\tau^*} > 0$ täsmälleen silloin, kun päätöksentekijä valitsee lopetettaessa punaisen vaihtoehdon ($\theta > 0$) tai sinisen vaihtoehdon ($\theta < 0$).

³Dynkinin kaava on keskeinen työkalu diffuusioprosessien analyysissä. Se kertoo, että jos u on riittävän säännöllinen funktio ja τ on pysäytysaika, niin

$$\mathbb{E}^x[u(X_\tau)] = u(x) + \mathbb{E}^x\left[\int_0^\tau Lu(X_s) ds\right],$$

missä L on prosessin generaattori.

Täten, jos $\theta \neq 0$ on todellinen harha, on todennäköisyys sille, että päätöksentekijä tekee ”oikean” valinnan lopetettaessa,

$$P^\theta(\{\theta Z_{\tau^*} > 0\}) = \frac{1}{1 + \exp\left(-\frac{2|\theta|\bar{z}}{\sigma^2}\right)}, \quad \theta \neq 0,$$

mikä kasvaa epävarmuuden ϵ myötä.

Todistus. Tuloksen (”oikean” valinnan todennäköisyys) todistus. Olkoon $Z_t = \theta t + \sigma W_t$ mitan P^θ alla ja pysäytysaika $\tau^* = \inf\{t \geq 0 : |Z_t| = \bar{z}\}$. Tarkastellaan eksponentiaalista martingaalia

$$M_t = \exp\left(-\frac{2\theta}{\sigma^2} Z_t\right), \quad t \geq 0.$$

Itön lemman mukaan M on P^θ -martingaali. Asetetaan $\tau_n := \tau^* \wedge n$. Doobin valinnaisen otannan lause antaa

$$\mathbb{E}^\theta[M_{\tau_n}] = \mathbb{E}^\theta[M_0] = 1 \quad \text{kaikilla } n.$$

Koska $Z_{\tau^*} \in \{-\bar{z}, \bar{z}\}$, on $M_{\tau^*} \in \{e^{-2\theta\bar{z}/\sigma^2}, e^{2\theta\bar{z}/\sigma^2}\}$ ja näin $|M_{\tau^*}| \leq e^{2|\theta|\bar{z}/\sigma^2}$. Tästä seuraa, että $M_{\tau_n} \rightarrow M_{\tau^*}$ m.v. ja $|M_{\tau_n}| \leq e^{2|\theta|\bar{z}/\sigma^2}$, jolloin dominoidun konvergenssin nojalla

$$\mathbb{E}^\theta[M_{\tau^*}] = \lim_{n \rightarrow \infty} \mathbb{E}^\theta[M_{\tau_n}] = 1.$$

Merkitään $p = P^\theta(Z_{\tau^*} = \bar{z})$. Tällöin

$$1 = \mathbb{E}^\theta[M_{\tau^*}] = p e^{-2\theta\bar{z}/\sigma^2} + (1-p) e^{2\theta\bar{z}/\sigma^2}.$$

Ratkaisemalla p saadaan, kun $a = e^{-2\theta\bar{z}/\sigma^2}$,

$$1 = p a + (1-p) a^{-1} \implies p = \frac{1}{1+a} = \frac{1}{1 + \exp\left(-\frac{2\theta\bar{z}}{\sigma^2}\right)}.$$

Koska lauseessa kiinnostuksen kohteena on tapahtuma $\{\theta Z_{\tau^*} > 0\}$, saadaan lopulta symmetrian nojalla

$$P^\theta(\{\theta Z_{\tau^*} > 0\}) = \frac{1}{1 + \exp\left(-\frac{2|\theta|\bar{z}}{\sigma^2}\right)}.$$

□

Lauseen 2 todistus antaa suljetun muodon optimaalisen pysäytysongelman arvo-funktiolle. Erityisesti ajanhetken 0 arvo v_0 toteuttaa

$$v_0 - \frac{1}{2} = \begin{cases} 0, & \text{jos } \frac{1+\epsilon}{2} \geq \hat{r}, \\ \frac{\alpha^2 \sigma^2}{4\bar{z}^2} \left[\frac{1}{\bar{r}(1-\bar{r})} - \frac{4}{(1+\epsilon)(1-\epsilon)} \right], & \text{jos } \frac{1+\epsilon}{2} < \hat{r}, \end{cases} \quad (3.19)$$

missä $\frac{1}{2}$ on suurin saavutettavissa oleva tuotto ilman oppimista. Täten $v_0 - \frac{1}{2}$ kuvaa oppimisvaihtoehdon arvoa. Pienillä epävarmuuden arvoilla $\epsilon < 2\hat{r} - 1$ tämä arvo on positiivinen, mutta pienenee jatkuvasti ja lähestyy nollaa, kun ϵ kasvaa kohti vaihtopistettä. Huomataan, että ehto $\frac{1+\epsilon}{2} = \hat{r}$ vastaa juuri tilannetta, jossa molemmat tapaukset yhtyvät arvoon \bar{r} , minkä vuoksi v_0 on jatkuva kohdassa $\epsilon = 2\hat{r} - 1$. Tämä on myös yhdenmukaista aiemmin esitetyn intuition kanssa: kun epävarmuus kasvaa, oppimisen kannattavuus vähenee, ja sen arvo lähestyy nollaa.

Numeerisena esimerkkinä tarkastellaan tapausta, jossa $(c, \sigma, \alpha) = (0.02, 0.8, 0.1)$. Tällöin epävarmuuden kynnyksarvoksi saadaan $\epsilon_{\text{cut}} \approx 0.0195$. Oppiminen hyväksytään, jos $\epsilon = 0.01$, mutta hylätään, kun $\epsilon = 0.03$. Kun $\epsilon = 0.01$, odotettu otoskoko $P^{\theta=0}$ -mitan alla on $\mathbb{E}^{\theta=0}[\tau^*] \approx 0.098$, ja vastaava todennäköisyys valita ”oikea” vaihtoehto $P^{\theta=\alpha}$ -tapauksessa on noin 0.52.

Vertauksena tarkastellaan binääristä valintaa ”panos punaiselle”, jonka voittotodennäköisyys on tuntematon mutta rajoissa $[(1-\epsilon)/2, (1+\epsilon)/2]$. Symmetrian vuoksi robusti vastapuoli valitsee ”punaiselle” epäsuotuisimman reunan $(1-\epsilon)/2$, kun taas siniselle vastaavasti $(1+\epsilon)/2$. Yhden periodin asetelmassa epävarman panoksen arvo supistuu lineaarisesti kohti keskipistettä, jolloin todennäköisyysvastine saadaan puolittamalla etäisyys neutraalista arvosta $1/2$:

$$p_{\text{Ell}}(\epsilon) = \frac{1}{2} - \frac{1}{4}\epsilon.$$

Täten $\epsilon = 0.03$ ja $\epsilon = 0.01$ tuottavat vastaavasti 0.4925 ja 0.4975. Huomataan siis, että tällainen klassinen Ellsbergin asetelma (ilman oppimista) ei eroa dramaattisesti aikaisemmasta tuloksesta.

3.3 Robusti peräkkäinen hypoteesin testaus

PT suorittaa otannan signaaliprosessista Z päämääränään suorittaa valinta kahden hypoteesin väliltä

$$H_0 : \theta = 0 \text{ ja } H_1 : \theta = \beta,$$

missä $\beta > 0$. Tässäkin on tärkeää muistaa, että θ :ssa on epävarmuutta ja halutaan robusti päätössääntö.

Epstein ja Ji tarkastelevat seuraavaa yleisen mallin erikoistapausta. Olkoon $\Theta = \{0, \beta\}$. Valinnat a_0 ja a_1 tarkoittavat hypoteesin H_0 ja H_1 hyväksymistä. Kolmatta valintaa ei ole mukana, koska ulkopuolista vaihtoehtoa ei ole. PT:n on valittava jompikumpi hypoteeseista.

Priorijoukko \mathcal{M}_0 on sama kuin yhtälössä (3.3), eli todennäköisyys $\theta = \beta$ kuuluu välille $[m_0, \bar{m}_0]$. Tuottofunktiot määritellään seuraavasti:

$$u(a_0, 0) = u(a_1, \beta) = a + b, \quad u(a_0, \beta) = b, \quad u(a_1, 0) = a,$$

missä $a, b > 0$.

Tässä spesifikaatiossa on kaksi eroa Ellsbergin asetelmaan verrattuna. Ensiksi hypoteesivalinnassa ei ole riskitöntä verrokkia, kuten urnassa, josta osa todennäköisyyksistä on tunnettu. Toiseksi värien symmetria on Ellsbergissä luontevaa, mutta

hypoteesien symmetria ei ole, siksi parametreja a ja b ei oleteta yhtäsuuriksi eikä todennäköisyysvälin $[m_0, \bar{m}_0]$ tarvitse olla symmetrinen pisteen $\frac{1}{2}$ suhteen.

Optimaalisen pysäytysongelman (3.1) ratkaisu saadaan tässä tapauksessa suljetussa muodossa. Tarkastellaan aluksi erityistä bayesilaista tapausta, jossa $M_0 = \{\mu\}$ ja kaikilla ajanhetkellä t pätee $M_t = \{\mu_t\}$, missä μ_t on posteriorijakauma parametrille θ . Koska $\Theta = \{0, \beta\}$, posteriorijakauma määräytyy yksikäsitteisesti luvun $m_t = \mu_t(\beta)$ perusteella, joka on todennäköisyys sille, että $\theta = \beta$.

Merkitään $\tilde{r}_B^\ell < \tilde{r}_B^R$ pysäytysrajoja. Näiden rajojen yleinen karakterisointi esitetään luvussa 4. Tässä erityistapauksessa ne voidaan kuitenkin laskea eksplisiittisesti ja ne saavat seuraavan muodon:

$$l(\tilde{r}_B^R) - l(\tilde{r}_B^\ell) = \frac{a + b}{\hat{c}}, \quad \frac{\tilde{r}_B^\ell}{1 - \tilde{r}_B^\ell} = \frac{b - a}{\hat{c}}, \quad (3.20)$$

missä

$$\hat{c} = \frac{2\sigma^2}{\beta^2}c.$$

Tässä $c > 0$ on alkuperäinen oppimiskustannus aikayksikköä kohden, σ on signaaliprosessin volatilitteetti ja β on hypoteesin H_1 mukainen ajautuma. Siten \hat{c} on normalisoitu oppimiskustannus, joka syntyy, kun pysäytysongelma kirjoitetaan posterioritodennäköisyyden m_t funktiona.

Tämän perusteella saadaan seuraava klassinen tulos.

Lause 3. (*Peskir and Shiryaev 2006*). *Bayesilaisessa tapauksessa, mille tahansa alkupriorille m_0 , on optimaalista jatkaa oppimista hetkellä t täsmälleen silloin kun*

$$\tilde{r}_B^\ell < m_t < \tilde{r}_B^R. \quad (3.21)$$

Muussa tapauksessa on optimaalista hyväksyä hypoteesi H_1 kun $m_t \geq \tilde{r}_B^R$ ja hypoteesi H_0 kun $m_t \leq \tilde{r}_B^\ell$.

Mallissa, jossa päätöksentekijä kohtaa epävarmuutta prioreista, pysäytysrajat muuttuvat arvoiksi \tilde{r}^ℓ ja \tilde{r}^R siten, että $\tilde{r}^\ell < \tilde{r}^R$. Seuraava tulos karakterisoi optimaalisen pysäytysstrategian tässä tapauksessa.

Lause 4. *Epävarmuuden vallitessa on optimaalista pysäyttää ja hyväksyä H_1 tai H_0 sen mukaan, onko $m_t \geq \tilde{r}^R$ tai $m_t \leq \tilde{r}^\ell$. Muussa tapauksessa on optimaalista jatkaa.*

Lisäksi, jos $a = b$, niin silloin pätee

$$\tilde{r}_B^\ell < \tilde{r}^\ell \quad \text{ja} \quad \tilde{r}^R < \tilde{r}_B^R. \quad (3.22)$$

Lauseen todistus löytyy kappaleesta 5.4.

Oletetaan, että tappiorakenne on symmetrinen ($a = b$). Tällöin lauseella on merkittäviä seurauksia sen suhteen, miten Bayesilaisen ja robustia päätöksentekoa tekevän henkilön optimaalinen pysäytysstrategia vertautuvat toisiinsa.

Jos $m_0 \in [\underline{m}_0, \bar{m}_0]$ viittaa yhteensopivaan bayesilaiseen, lauseen tulokset merkitsevät seuraavaa:

1. Jos kaikki yhteensopivat bayesilaiset lopettavat ja valitsevat jonkin vaihtoehdon α_i , niin myös robusti päätöksentekijä lopettaa ja valitsee saman vaihtoehdon, $i = 1, 2$.

2. Jos kaikki yhteensopivat bayesilaiset jatkavat havainnointia, voi silti olla robustille päätöksentekijälle optimaalista lopettaa.

Toisin sanoen robustin päätöksentekijän tulisi hyväksyä yksimielinen suositus lopettamisesta, mutta yksimielinen suositus jatkamisesta ei riitä takaamaan samaa. Tässä mielessä oppimisen hyötyä voidaan yliarvioida robustin päätöksentekijän tapauksessa.

Intuitiivisesti syy on seuraava. Prioriepävarmuus saa robustin päätöksentekijän näkemään signaalirakenteen vähemmän informatiivisena, koska hän tarkastelee sitä pahimman tapauksen todennäköisyysmitan P^* näkökulmasta. Signaaliprosessi itsessään ei ole epäselvä, mutta robusti päätöksentekijä epäilee taustalla olevaa mallia. Bayesilaisilla ei ole tällaista vastaavaa ongelmaa: jokainen bayesilainen luottaa omilla priori-uskomuksillaan malliin ja pitää signaalin lisäyksiä ehdollisesti iid-muotoisina. Vain robusti päätöksentekijä sisällyttää päätöksiinsä epävarmuuden mallin oikeellisuudesta ja siten vähentää oppimisen hyötyä.

4 Yleisempi tulos

Seuraavaksi Epstein laajentaa teoriaansa ja siirtyy tarkastelemaan optimaalisen pysäyttämisen ongelmaa yleisemmässä muodossa. Notaatioiden tiivistämiseksi kirjoitetaan u_{ij} lyhenteenä hyötyfunktioista $u(a_i, \theta_j)$, missä $i, j = 0, 1$.

Lause 5 kuvaa optimaalisen pysäytyksen ratkaisun kappeleessa 3.1 esitetyssä ongelmassa oletuksella, että vallitsee joko tuottojen symmetria ($u_{01} = u_{10}$) tai ettei päätöksentekijällä ole käytettävissään riskitöntä vaihtoehtoa ($u_2 \leq \min\{u_{10}, u_{01}\}$).

Tuottosymmetria toteutuu Lauseessa 2, mutta siinä oletetaan lisäksi enemmän, erityisesti etukäteinen indifferenssi vaihtoehtojen a_0 ja a_1 välillä ($m_0 = \bar{m}_0 = 1$) sekä $u_2 = \frac{1}{2}(u_{00} + u_{10})$. Nämä oletukset yleistetään myöhemmin kohdassa Lause 5(a).

Oletus riskittömän vaihtoehdon puuttumisesta on keskeinen hypoteesintestausmerkissä, ja vastaava optimaalisen pysäytyksen ongelma on rakenteeltaan isomorfinen Lauseen 5 osan (b) kanssa.

Sekä m_t että \bar{m}_t , jotka on määritelty yhtälössä (3.4)-(3.5), ovat kasvavia funktioita suureesta $\varphi(t, z_t)$. Tästä seuraa, että on olemassa yksikäsitteinen todennäköisyyspari $(\pi, \bar{\pi})$ sekä yksikäsitteinen (deterministinen) signaalirealisaation kulku (\tilde{z}_t) siten, että jokaisella hetkellä t pätee

$$\pi = m_t(\tilde{z}_t), \quad \bar{\pi} = \bar{m}_t(\tilde{z}_t),$$

ja lisäksi

$$\pi u_{11} + (1 - \pi)u_{10} = \bar{\pi} u_{01} + (1 - \bar{\pi}) u_{00}.$$

Esimerkiksi $\tilde{z}_0 = 0$, $\pi = m_0$ ja $\bar{\pi} = \bar{m}_0$ täyttyvät täsmälleen silloin, kun a_0 ja a_1 ovat etukäteen indifferenssissä. Yleisemmin valinnat a_0 ja a_1 ovat yhdentekeviä signaalin \tilde{z}_t kohdalla ajanhetkellä t , ja päätöksentekijä suosii a_0 -valintaa, jos $Z_t < \tilde{z}_t$, ja a_1 -valinta, jos $Z_t > \tilde{z}_t$.

Normalisoidaan oppimisen kustannus

$$\hat{c} = \frac{2c\sigma^2}{(\theta_1 - \theta_0)^2}.$$

Optimaaliset pysäytysstrategiat määritellään useiden kriittisten arvojen avulla, jotka puolestaan määritellään funktioiden l ja \tilde{l} avulla. Kaikille $r \in (0, 1)$ asetetaan

$$l(r) = 2 \log\left(\frac{r}{1-r}\right) - \frac{1}{r} + \frac{1}{1-r},$$

$$\tilde{l}(r) = \log\left(\frac{r}{1-r}\right) + \frac{r}{1-r}.$$

Olkoon kriittiset pisteet (r_1^R, r_2^R) , (r_1^l, r_2^l) , (r^R, r^l) ja $(\tilde{r}^R, \tilde{r}^l)$ sellaiset, että ne ratkaisevat vastaavat yhtälöparit. Esimerkiksi pisteet (r_1^R, r_2^R) määräytyvät ratkaisemalla

$$\begin{aligned} l(r_2^R) - l(r_1^R) &= \frac{u_{11} - u_{10}}{\hat{c}}, \\ \tilde{l}(r_2^R) - \tilde{l}(r_1^R) &= \frac{u_2 - u_{10}}{\hat{c}}, \end{aligned} \tag{4.1}$$

missä u_{ij} on utiliteetti päätöksestä a_i parametrin θ_j ollessa tosi. Nämä kriittiset arvot määrittävät optimaalisen oppimisen ja pysäytyksen rajat mallissa.

Arvot (r_1^l, r_2^l) määritellään ratkaisemalla

$$l(r_2^l) - l(r_1^l) = -\frac{u_{00} - u_{01}}{\hat{c}}, \quad (2)$$

$$\tilde{l}(r_2^l) - \tilde{l}(r_1^l) = \frac{u_2 - u_{00}}{\hat{c}}. \quad (4.2)$$

ja pisteet r^R ja r^l ratkaisevat

$$l(r^R) - l(\pi) = \frac{u_{11} - u_{10}}{\hat{c}}, \quad (3)$$

$$l(r^l) - l(\pi) = -\frac{u_{00} - u_{01}}{\hat{c}}. \quad (4.3)$$

Lisäksi arvot \tilde{r}^R ja \tilde{r}^l määräytyvät yhtälöistä

$$l(\tilde{r}^R) - l(\pi) = l(\tilde{r}^l) - l(\pi) + \frac{u_{11} - u_{10} + u_{00} - u_{01}}{\hat{c}}, \quad (4)$$

$$\tilde{l}(\tilde{r}^R) - \tilde{l}(\pi) = \pi(l(\tilde{r}^R) - l(\pi)), \quad (5)$$

$$\tilde{l}(\tilde{r}^l) - \tilde{l}(\pi) = -\pi(l(\tilde{r}^l) - l(\pi)). \quad (4.4)$$

Jos palkkiosymmetria pätee, yhtälöryhmä (4.4) supistuu muotoon (4.3).

Määritellään

$$u_2^{**} = \frac{\hat{c}}{2} \left[\frac{1}{r^l(1 - r^l)} - \frac{1}{\pi(1 - \pi)} \right] + \frac{u_{00} - u_{01}}{2}. \quad (4.5)$$

(4.5) toimii jatkossa raja-arvona sille, onko a_2 ikinä optimaalinen. Seuraava lemma osoittaa näiden kriittisten arvojen olemassaolon ja yksikäsitteisyyden sekä joukon ominaisuuksia, jotka ovat tärkeitä optimaalisen pysäytyssäännön johtamisessa.

Lemma 2. *Tässä lemmassa tarkastellaan eri raja-arvojen olemassaoloa ja niiden sijoittumista posteriorin π ympärille. Ensimmäiseksi todetaan, että yhtälöillä (4.3) ja (4.4) on yksikäsitteiset ratkaisut. Näille ratkaisuille pätee*

$$\tilde{r}^l < \pi, \quad \tilde{r}^R > \pi. \quad (4.6)$$

*Jos turvallisen vaihtoehdon tuotto täyttää ehdon $u_2 \geq u_2^{**}$, niin myös yhtälöillä (4.1) ja (4.2) on yksikäsitteiset ratkaisut. Näiden ratkaisujen järjestys on*

$$r_2^l < r_1^l < r_1^R < r_2^R, \quad \pi < r^R, \quad r^l < \pi.$$

Jos lisäksi palkkiosymmetria pätee, saadaan kaksi lisäehtoa:

$$\pi + \bar{\pi} = 1 = r^l + r^R, \quad (4.7)$$

sekä keskeinen ekvivalenssi

$$r_1^l \leq \pi \iff r_1^R \geq \pi \iff u_2 \geq u_2^{**}. \quad (4.8)$$

Seuraavaksi määritellään funktiot, jotka ilmaisevat sen signaalin Z_t arvon, jolla posteriori on täsmälleen r . Ne ovat:

$$\begin{aligned}\bar{f}(t, r) &= \frac{\theta_1 + \theta_0}{2} t + \frac{\sigma^2}{\theta_1 - \theta_0} \log\left(\frac{1 - \bar{m}_0}{\bar{m}_0} \frac{r}{1 - r}\right), \\ f(t, r) &= \frac{\theta_1 + \theta_0}{2} t + \frac{\sigma^2}{\theta_1 - \theta_0} \log\left(\frac{1 - m_0}{m_0} \frac{r}{1 - r}\right).\end{aligned}$$

Jälkimmäinen kuvastaa alarajan prioripisteen funktiota. Näille funktioille pätee:

$$m_t(\bar{f}(t, r)) = r = \bar{m}_t(f(t, r)).$$

Lisäksi, mille tahansa pisteparille r_1 ja r_2 pätee:

$$\bar{f}(t, r_1) \leq \tilde{z}_t \iff r_1 \leq \bar{\pi}, \quad f(t, r_2) \geq \tilde{z}_t \iff r_2 \geq \pi. \quad (4.9)$$

Lopuksi määritellään kolme pysäytysaikaa, jotka vastaavat kolmea mahdollista tapaa, joilla oppimisprosessi voi päättyä:

$$\tau_0 \equiv \min\{t \geq 0 : Z_t \leq \bar{f}(t, r_2^l)\} = \min\{t \geq 0 : m_t \leq r_2^l\},$$

$$\tau_1 \equiv \min\{t \geq 0 : Z_t \geq f(t, r_2^R)\} = \min\{t \geq 0 : m_t \geq r_2^R\},$$

$$\tau_2 \equiv \min\{t \geq 0 : \bar{f}(t, r_1^l) \leq Z_t \leq f(t, r_1^R)\} = \min\{t \geq 0 : m_t \geq r_1^l \text{ ja } m_t \leq r_1^R\}.$$

Lause 5. (a) Oletetaan palkkiosymmetria ($u_{01} = u_{10}$).

(a.i) Jos $r_1^l \leq \pi$, niin optimaalinen pysäytysaika τ^* on

$$\tau^* = \min\{\tau_i : i = 0, 1, 2\}.$$

Lisäksi, jos $\tau^* = \tau_i$, niin valinta a_i on optimaalinen kyseisellä pysäytyshetkellä. Erityisesti, jos lähtötilanteessa a_0 ja a_1 ovat etukäteen samantarvoisia ($\pi = m_0$ ja $\bar{\pi} = \bar{m}_0$), niin $\tau^* = 0$ ja valitaan a_2 .

(a.ii) Jos $r_1^l > \pi$, silloin

$$\begin{aligned}\tau^* &= \min\{t \geq 0 : Z_t \leq \bar{f}(t, r^l) \text{ tai } Z_t \geq f(t, r^R)\} \\ &= \min\{t \geq 0 : m_t \leq r^l \text{ tai } m_t \geq r^R\}.\end{aligned}$$

Lisäksi valinta a_0 on optimaalinen pysäytyksessä, jos $Z_{\tau^*} \leq \bar{f}(\tau^*, r^l)$ (vastaavasti jos $m_{\tau^*} \leq r^l$). Päätös a_1 on optimaalinen, jos $Z_{\tau^*} \geq f(\tau^*, r^R)$ (vastaavasti jos $m_{\tau^*} \geq r^R$). Tässä tapauksessa a_2 ei ole koskaan optimaalinen.

(b) Oletetaan $u_2 \leq \min\{u_0, u_{01}\}$. Tällöin

$$\begin{aligned}\tau^* &= \min\{t \geq 0 : Z_t \leq \bar{f}(t, r_2^l) \text{ tai } Z_t \geq f(t, r_2^R)\} \\ &= \min\{t \geq 0 : m_t \leq r_2^l \text{ tai } m_t \geq r_2^R\}.\end{aligned}$$

Lisäksi valinta a_0 on optimaalinen, jos $Z_{\tau^*} \leq \bar{f}(\tau^*, r_2^l)$ (vastaavasti $m_{\tau^*} \leq r_2^l$). Valinta a_1 on optimaalinen, jos $Z_{\tau^*} \geq f(\tau^*, r_2^R)$ (vastaavasti $m_{\tau^*} \geq r_2^R$). Myös tässä tapauksessa a_2 ei ole koskaan optimaalinen.

Tapauksessa (a), eli kun palkkiosymmetria pätee, jaamme tilanteen edelleen kahteen alatapaukseen riippuen arvojen r_1^l ja $\bar{\pi}$ keskinäisestä suuruussuhteesta. Yhtälöstä (4.2) seuraa, että arvo r_1^l pienenee, kun turvallisen vaihtoehdon tuotto u_2 kasvaa, kun taas $\bar{\pi}$ ei riipu u_2 :sta lainkaan. Siksi (a.i) toteutuu silloin, kun turvallisen (ei epävarmuutta sisältävän) vaihtoehdon a_2 tuotto on riittävän suuri.

Toinen tekijä, joka johtaa tilanteeseen (a.i), on suuri $\bar{\pi}$ eli vastaavasti pieni π . Yhtälön (4.7) mukaan $\bar{\pi}$ on suuri silloin, kun yläposteriorin raja \bar{m}_0 on suuri ja alaraja m_0 pieni. Toisin sanoen tilanne (a.i) saa tukea myös tapauksista, joissa lähötöjakauman epävarmuus on suuri (eli priorin vaihteluväli on suuri).

Alatapauksessa (a.i) optimaalinen pysäytysaika on $\tau^* = 0$, jos jokin seuraavista ehdoista täyttyy:

- $\bar{m}_0 \leq r_1^l$: lähtöuskomukset ovat vahvasti painottuneet kohti parametria θ_0 , jolloin valitaan heti a_0 ;
- $m_0 \geq r_2^R$: lähtöuskomukset ovat voimakkaasti painottuneet kohti parametria θ_1 , jolloin valitaan heti a_1 ;
- $m_0 \geq r_1^l$ ja $m_0 \leq r_1^R$: pahimman tapauksen todennäköisyydet molemmille tiloille ovat niin pienet, etteivät a_0 tai a_1 ole riittävän houkuttelevia vaihtoehtoja, jolloin näyttteenoton kustannusta ei ole perusteltua maksaa ja valitaan turvallinen a_2 .

Näistä seuraa, että oppimisen jatkaminen ($\tau^* > 0$) on optimaalista vain, jos alkuperäiset uskomukset ovat tiettyssä mielessä "keskitasoisia". Tämä tarkoittaa, että posterioriväli asettuu sopivasti kriittisten rajojen väliin. Jatkaminen on optimaalista täsmälleen silloin, kun joko

$$[r_2^l < \bar{m}_0 < r_1^l] \quad \text{ja} \quad m_0 < r_2^R,$$

tai

$$[r_1^R < m_0 < r_2^R] \quad \text{ja} \quad \bar{m}_0 > r_1^l.$$

Nämä ehdot määrittelevät sen posteriorialueen, jolla lisähavainnointi on optimaalista.

Tämä jatkamisalue voi olla tyhjä. Koska oppiminen koskee vain valintojen a_0 ja a_1 tuottojen erottelua, on lähtöhetken kannalta oppimisen kannalta epäsuotuisin tilanne se, jossa nämä kaksi vaihtoehtoa ovat etukäteen samanarvoisia. Tällöin oppiminen olisi hidasta ja kallista, koska tarvittaisiin pitkä näyttteenottojakso, jotta erot alkaisivat näkyä. Tässä tilanteessa on siis optimaalista hylätä oppiminen ja valita suoraan a_2 , kuten lauseessa 2 todettiin.

Jos taas esimerkiksi a_1 on selvästi parempi jo lähtötilanteessa, voi suhteellisen lyhytkin näyttteenotto riittää erottamaan a_1 :n ja a_2 :n. Lisäksi, jos lähtöposteriorin yläraja \bar{m}_0 on riittävän suuri (esimerkiksi lähellä arvoa 1), lähes varma tieto siitä, että $\theta = \theta_1$, voi johtaa siihen, että oppiminen hylätään kokonaan ja valitaan heti a_1 pikemminkin kuin a_2 , samoin kuin Ellsbergin kontekstissa.

Tapauksessa (a.ii) optimaalinen pysäytysaika on $\tau^* = 0$ täsmälleen silloin, kun lähtöposteriorin väli $[m_0, \bar{m}_0]$ ei leikkaa väliä (r^l, r^R) . Tällöin valintaa a_2 ei koskaan

tehdä, riippumatta siitä, aloitetaanko oppiminen vai ei. Sen tuotto on yksinkertaisesti liian alhainen (yhtälöstä (4.8), koska $u_2 < u_2^{**}$) verrattuna siihen, mitä voidaan odottaa joko a_0 :sta tai a_1 :stä, mahdollisesti lyhyen oppimisen jälkeen.

Lisäksi, vaikka jonkin verran oppimista tapahtuisi, ei ole missään vaiheessa optimaalista valita a_2 , koska toteutunut näyte ei tee siitä kannattavaa, kuten lauseen 2 yhteydessä todettiin. Lemma 2 antaa lisäksi $\tau^* > 0$ tapauksessa (a.ii). Yhdistämällä tämä tapaukseen (a.i) saadaan, että jos lähtötilanteessa a_0 ja a_1 ovat etukäteen samanarvoisia, niin a_2 valitaan täsmälleen silloin ja vain silloin, kun oppimista ei lainkaan aloiteta. Tämä yleistää Ellsbergin mallin tuloksen (missä oletettiin myös $u_2 = \frac{1}{2}(u_{00} + u_{10})$, mutta tätä oletusta ei tässä tarvita).

Tarkastellaan lopuksi tapausta (b), jossa varman vaihtoehdon tuotto on niin alhainen, ettei sitä valittaisi koskaan, riippumatta lähtöuskomuksista tai mahdollisuudesta oppia. Optimaalinen strategia on tällöin sisällöllisesti sama kuin tapauksessa (a.ii); vain kriittiset raja-arvot muuttuvat, koska tuotto-oletukset ovat erilaiset.

Lisäksi huomataan, että kun $\bar{m}_0 = m_0$, myös $\bar{\pi} = \pi$, ja yhtälöt (4.4) määräävät kriittiset pisteet \tilde{r}^R ja \tilde{r}^l seuraavasti:

$$l(\tilde{r}^R) - l(\tilde{r}^l) = \frac{u_{11} - u_{10} + u_{00} - u_{01}}{\hat{c}},$$

$$\tilde{l}(\tilde{r}^R) - \tilde{l}(\tilde{r}^l) = \frac{u_{00} - u_{10}}{\hat{c}},$$

jotka vastaavat yhtälöitä (21.1.14) ja (21.1.15) teoksessa Peskir & Shiryaev (2006).

Huomio Epstein & Ji:n viittauksesta Peskir–Shiryaeviin. Epstein ja Ji viittaavat yhtälöiden (4.4) yhteydessä Peskirin ja Shiryaevin *Optimal Stopping and Free-Boundary Problems* -teoksen yhtälöihin (21.1.14) ja (21.1.15). Epstein ja Ji korostavat rakenteellista samankaltaisuutta: molemmissa tapauksissa kriittiset raja-arvot määräytyvät kahden riippumattoman rajaehdon kautta ja nämä ehdot yhdessä yksikäsitteistävät optimaalisen pysäytysalueen.

Tässä mielessä heidän mallinsa voidaan nähdä yleistyksenä Peskir–Shiryaevin käsittelemästä kahden rajan pysäytysongelmasta tapaukseen, jossa rajat eivät synny arvofunktion differentiaaliyhtälöstä vaan posteriorin logaritmisista muunnoksista. Aikaisempi analyysi osoittaa, että Epstein & Ji:n yhtälöpari (4.4) täyttää täsmälleen saman teoreettisen roolin: se antaa kaksi toisistaan riippumatonta ehtoa, joiden avulla kriittiset rajat \tilde{r}^l ja \tilde{r}^R voidaan ratkaista yksikäsitteisesti.

5 Todistukset ja muut tulokset

5.1 Lauseen 5 todistus

Lähdetään liikkeelle todistamalla lause 5, sillä tuloksena se on tutkielman tärkein. Epsteinin todistusstrategia on seuraavanlainen:

1. Arvataan mittojen joukosta \mathcal{P}_0 priori P^* , joka edustaa päätöksentekijän kannalta pahinta mahdollista tapausta.
2. Ratkaistaan tavallinen optimaalisen pysäytyksen ongelma olettaen, että priori on yksiselitteisesti juuri tämä P^* .
3. Näytetään, että kohdassa (2) johdettu arvofunktiio on samalla arvofunktiio myös alkuperäiselle robustille ongelmalle (3.1).
4. Määritetään optimaalinen pysäytysaika τ^* tämän arvofunktion avulla.

Todistus tulee noudattamaan paljolti Epsteinin alkuperäistä muotoilua, mutta siinä osoitetaan todeksi muutamia kohtia, mitkä eivät tule esille Epsteinin todistuksessa. Pohjustetaan myös tarkemmin tiettyjä tärkeitä termejä, mitä todistuksessa esiintyy. Hamilton-Jacobi-Bellman-(HJB)-yhtälö tulee esiin useasti todistuksessa. Optimaalisen pysäytyksen yhteydessä se seuraa dynaamisen ohjelmoinnin periaatteesta ja saa ns. variaatioepäyhtälön muodon

$$\max \left\{ X(t, z) - v(t, z), -c + v_t(t, z) + \mathcal{L}v(t, z) \right\} = 0,$$

missä \mathcal{L} on prosessin generaattori annetun mitta-oletuksen alla. Yhtälö kuvaa valintaa pysäyttämisen ja jatkamisen välillä: pysäytettäessä saadaan hyöty $X(t, z)$, kun taas jatkaminen tuottaa odotetun arvon dynamiikan $-c + v_t + \mathcal{L}v$ kautta. Jatkoalueella ($v > X$) yhtälö supistuu HJB-muotoon

$$-c + v_t(t, z) + \frac{1}{2}\sigma^2 v_{zz}(t, z) + f(t, z)v_z(t, z) = 0.$$

Pysäytysalueella päätöksentekijä pysäyttää prosessin välittömästi, jolloin arvofunktiio on yhtä suuri kuin välitön tuotto eli

$$v(t, z) = X(t, z).$$

Yhdessä näistä muodostuu vapaan reunan ongelma, jossa tuntemattomat reunat määritellään riittävillä sileys-ehdoilla ("smooth fit").

Tässä mallissa generaattorin ajautuma-termissä esiintyvä funktio $f(t, z)$ riippuu pahimman tapauksen mitasta P^* , joka valitsee posteriorin m_t tai \tilde{m}_t sen mukaan, onko Z_t pisteen \tilde{z}_t ylä- vai alapuolella. Tämän vuoksi arvofunktion muoto ja HJB-yhtälö eroavat kahdella alueella, mikä tekee yhtälön ratkaisemisesta vapaan reunan ongelman.

Seuraavassa vaiheessa ratkaistaan tämä HJB-yhtälö sekä siihen liittyvät sileys-ehdot ja vahvistetaan, että saatu funktio on todellakin optimaalisen pysäytyksen arvofunktiio.

Intuitio ehdotetulle pahimman tapauksen mitalle P^* annettiin jo kappaleessa 3.2 Ellsbergin esimerkin yhteydessä. Tässä yleisemmässä asetelmassa ajatellaan, että P^* valitsee ajautuman siten, että todennäköisyysmassaa siirtyä ylöspäin (kun $Z_t \geq \tilde{z}_t$) tai alaspäin (kun $Z_t < \tilde{z}_t$) tehdään päätöksentekijän kannalta mahdollisimman epäedulliseksi: mitta käyttää posterioria m_t kun $Z_t > \tilde{z}_t$ ja \tilde{m}_t kun $Z_t < \tilde{z}_t$. (Täsmällinen määritelmä esitetään kaavassa (5.3).)

Arvofunktion v ratkaisemiseksi tarkastellaan jatkoaluetta (tässä kohtaa Epstein viittaa Peskir ja Shirayev[20] teokseen taas, mutta pohjustetaan tulosta itse), jossa pysäytystä ei ole vielä tapahtunut. Tällä alueella arvofunktio toteuttaa edellä huomautun mukaisesti Hamilton-Jacobi-Bellman-yhtälön

$$-c + v_t(t, z) + \frac{1}{2}\sigma^2 v_{zz}(t, z) + f(t, z) v_z(t, z) = 0,$$

missä $f(t, z)$ on pahimman tapauksen mitan P^* aiheuttama ajautuma. Koska päätöksenteko riippuu ainoastaan posteriorista $y = m_t \in (0, 1)$ eikä arvofunktio riipu eksplisiittisesti ajasta, voidaan kirjoittaa $v(t, z) = V(y)$, jolloin HJB-yhtälö supistuu yksiulotteiseksi differentiaaliyhtälöksi

$$-c + \frac{1}{2}\hat{c}^2 y^2 (1-y)^2 V''(y) = 0.$$

Tämä antaa eksplisiittisen muodon

$$V''(y) = \frac{2c}{\hat{c}^2 y^2 (1-y)^2},$$

josta integrointi tuottaa tulokseksi

$$V'(y) = \frac{\hat{c}}{y(1-y)} + C_1, \quad V(y) = \hat{c}(2y-1) \log\left(\frac{y}{1-y}\right) + C_1 y + C_2.$$

Perustelu HJB-yhtälön supistumiselle: Olkoon havaintoprosessi

$$dZ_t = \theta dt + \sigma dB_t, \quad \sigma > 0, \quad \theta \in \{\theta_0, \theta_1\},$$

missä $(B_t)_{t \geq 0}$ on Brownin liike ja havaintojen filtraatio on $\mathcal{G}_t = \sigma(Z_s : 0 \leq s \leq t)$.

Määritellään posterioritodennäköisyys

$$m_t = \mathbb{P}(\theta = \theta_1 \mid \mathcal{G}_t) \in (0, 1).$$

Koska $m_t = \mathbb{E}[1_{\{\theta=\theta_1\}} \mid \mathcal{G}_t]$, prosessi $(m_t)_{t \geq 0}$ on \mathcal{G}_t -martingaali.

Merkitään $\mathbb{P}_i = \mathbb{P}(\cdot \mid \theta = \theta_i)$. Girsanovin lauseesta seuraa, että uskottavuuskerroin \mathcal{G}_t -filtraatiossa on

$$\Lambda_t = \frac{d\mathbb{P}_1}{d\mathbb{P}_0} \Big|_{\mathcal{G}_t} = \exp\left(\frac{\theta_1 - \theta_0}{\sigma^2} Z_t - \frac{\theta_1^2 - \theta_0^2}{2\sigma^2} t\right).$$

Bayesin kaavan mukaan posteriori voidaan kirjoittaa muodossa

$$m_t = \frac{m_0 \Lambda_t}{(1 - m_0) + m_0 \Lambda_t}.$$

Tästä seuraa logit-muunnokselle

$$L_t = \log \frac{m_t}{1 - m_t} = \log \frac{m_0}{1 - m_0} + \log \Lambda_t = \log \frac{m_0}{1 - m_0} + \frac{\theta_1 - \theta_0}{\sigma^2} Z_t - \frac{\theta_1^2 - \theta_0^2}{2\sigma^2} t.$$

Määritellään posteriorikeskiarvo

$$\hat{\theta}_t = \mathbb{E}[\theta \mid \mathcal{G}_t] = \theta_0(1 - m_t) + \theta_1 m_t.$$

Tällöin prosessi

$$W_t = \frac{1}{\sigma} \left(Z_t - \int_0^t \hat{\theta}_s ds \right)$$

on \mathcal{G}_t -Brownin liike. Siten havaintoprosessi voidaan kirjoittaa muotoon

$$dZ_t = \hat{\theta}_t dt + \sigma dW_t.$$

Asetetaan $\Delta\theta = \theta_1 - \theta_0$. Koska $\hat{\theta}_t = \theta_0 + \Delta\theta m_t$, saadaan

$$dL_t = \frac{\Delta\theta}{\sigma^2} dZ_t - \frac{\theta_1^2 - \theta_0^2}{2\sigma^2} dt = \frac{\Delta\theta}{\sigma^2} (\hat{\theta}_t dt + \sigma dW_t) - \frac{(\theta_1 - \theta_0)(\theta_1 + \theta_0)}{2\sigma^2} dt.$$

Yhdistämällä dt -termit ja käyttämällä $\hat{\theta}_t = \theta_0 + \Delta\theta m_t$ saadaan

$$\frac{\Delta\theta}{\sigma^2} \hat{\theta}_t - \frac{\Delta\theta(\theta_1 + \theta_0)}{2\sigma^2} = \frac{\Delta\theta}{\sigma^2} \left(\theta_0 + \Delta\theta m_t - \frac{\theta_1 + \theta_0}{2} \right) = \frac{(\Delta\theta)^2}{\sigma^2} \left(m_t - \frac{1}{2} \right).$$

Siis

$$dL_t = \frac{\Delta\theta}{\sigma} dW_t + \frac{(\Delta\theta)^2}{\sigma^2} \left(m_t - \frac{1}{2} \right) dt.$$

Koska $m_t = \ell(L_t)$, missä $\ell(x) = \frac{e^x}{1+e^x}$, pätevät derivaatat

$$\ell'(L) = m(1 - m), \quad \ell''(L) = m(1 - m)(1 - 2m).$$

Itön kaavalla

$$dm_t = \ell'(L_t) dL_t + \frac{1}{2} \ell''(L_t) d\langle L \rangle_t.$$

Koska $d\langle L \rangle_t = \left(\frac{\Delta\theta}{\sigma} \right)^2 dt$, saadaan

$$dm_t = m_t(1 - m_t) \left[\frac{\Delta\theta}{\sigma} dW_t + \frac{(\Delta\theta)^2}{\sigma^2} \left(m_t - \frac{1}{2} \right) dt \right] + \frac{1}{2} m_t(1 - m_t)(1 - 2m_t) \left(\frac{\Delta\theta}{\sigma} \right)^2 dt.$$

Huomataan, että dt -kertoimet kumoavat toisensa. Tämä ei ole sattumaa, vaan seurausta siitä, että posteriori

$$m_t = \mathbb{E}[1_{\{\theta=\theta_1\}} \mid \mathcal{G}_t]$$

on \mathcal{G}_t -martingaali. Martingaalilla ei voi olla ajautumatermiä, joten dm_t :n dt -osa häviää välttämättä.

Siten saadaan

$$dm_t = \frac{\Delta\theta}{\sigma} m_t(1 - m_t) dW_t.$$

Merkitään lyhyesti

$$\hat{c} = \frac{\Delta\theta}{\sigma},$$

jolloin

$$dm_t = \hat{c} m_t(1 - m_t) dW_t.$$

Koska päätöksentekijän informaatio hetkellä t on täsmälleen \mathcal{G}_t ja kaikki tulevia havaintoja koskeva ehdollinen jakauma määräytyy posteriorin m_t kautta, jatkoalueella voidaan esittää arvofunktio muodossa

$$v(t, z) = V(m(t, z)).$$

Prosessina tämä tarkoittaa $v(t, Z_t) = V(m_t)$.

Sovelletaan Itô'n kaavaa prosessiin $V(m_t)$. Koska $dm_t = \hat{c} m_t(1 - m_t) dW_t$, niin

$$d\langle m \rangle_t = \hat{c}^2 m_t^2 (1 - m_t)^2 dt.$$

Tästä seuraa

$$\begin{aligned} dV(m_t) &= V'(m_t) dm_t + \frac{1}{2} V''(m_t) d\langle m \rangle_t \\ &= V'(m_t) \hat{c} m_t(1 - m_t) dW_t + \frac{1}{2} \hat{c}^2 m_t^2 (1 - m_t)^2 V''(m_t) dt. \end{aligned}$$

Koska oppimisesta maksetaan kustannus c aikayksikköä kohti, arvofunktion on toteutettava

$$-c + \mathcal{L}V(m) = 0,$$

missä \mathcal{L} on prosessin m_t generaattori. Koska m_t toteuttaa SDE:n

$$dm_t = \hat{c} m_t(1 - m_t) dW_t,$$

se on yksiulotteinen diffuusio ilman ajautumaa ja volatilitteetilla $\sigma(m) = \hat{c} m(1 - m)$.

Tällaisen diffuusion generaattori on määritelmän mukaan

$$\mathcal{L}f(m) = \frac{1}{2} \sigma^2(m) f''(m) = \frac{1}{2} \hat{c}^2 m^2 (1 - m)^2 f''(m).$$

$$\mathcal{L}V(m) = \frac{1}{2} \hat{c}^2 m^2 (1 - m)^2 V''(m),$$

ja näin saadaan yksiulotteinen differentiaaliyhtälö

$$-c + \frac{1}{2} \hat{c}^2 m^2 (1 - m)^2 V''(m) = 0.$$

Toisin sanoen jatkoalueella alkuperäinen HJB-yhtälö supistuu tähän muotoon, koska posteriori m_t on yksiulotteinen diffuusio ilman ajautumaa, ja arvofunktio voidaan esittää posteriorin funktiona. \square

Koska pahimman tapauksen mitta vaihtaa käyttämäänsä posterioria pisteessä $\bar{\pi}$, arvofunktio saa eri vakioarvot alueilla $y \in (\bar{\pi}, r^R)$ ja $y \in (r^l, \bar{\pi})$. Jatkoalueella saatu differentiaaliyhtälö

$$V''(y) = \frac{\tilde{c}}{y^2(1 - y)^2}$$

toteutuu molemmilla alueilla, mutta vakiot määräytyvät pysäytysrajojen ehdoista. Pysäytyshetkellä arvofunktion tulee yhtyä välittömään palkkioon sekä täyttää ns. arvon jatkuvuusehdon (value matching), mikä oikeanpuoleisella rajalla r^R antaa ehdot $V^R(r^R) = X(r^R)$ ja $V_y^R(r^R) = X_y(r^R)$ ja vastaavat ehdot vasemmalla rajalla r^l . Lisäksi pisteessä $\bar{\pi}$ arvofunktion ja sen derivaatan tulee olla jatkuvia, mikä tuottaa ehdot (5.2). Näin muodostuu yhtälöryhmä (5.1).

$$\left\{ \begin{array}{l} V_{yy}^R(y) = \frac{\tilde{c}}{y^2(1-y)^2}, \quad y \in (\bar{\pi}, r^R), \\ V^R(r^R) = (u_{11} - u_{10})r^R + u_{10}, \\ V_y^R(r^R) = (u_{11} - u_{10}), \\ \\ V_{yy}^l(y) = \frac{\tilde{c}}{y^2(1-y)^2}, \quad y \in (r^l, \bar{\pi}), \\ V^l(r^l) = -(u_{00} - u_{01})r^l + u_{00}, \\ V_y^l(r^l) = -(u_{00} - u_{01}). \end{array} \right. \quad (5.1)$$

Aikaisemmin mainitut sileys-ehdot ovat:

$$\begin{aligned} V^R(\bar{\pi}) &= V^L(\bar{\pi}), \\ V_y^R(\bar{\pi}) &= V_y^L(\bar{\pi}). \end{aligned} \quad (5.2)$$

Tapauksessa (a.ii) oletetaan tuotto-symmetria, jolloin

$$V_y^R(\bar{\pi}) = V_y^l(\bar{\pi}) = 0,$$

ja yhtälö (4.3) hajoaa kahdeksi erilliseksi jatkoalueen yhtälöksi. Tässä tapauksessa pysäytyshetkellä saatava palkkio X_t , joka on määritelty kohdassa (3.9), määräytyy hetkellisen päätösongelman perusteella. PT valitsee päätöksen $a \in \{a_0, a_1, a_2\}$, joka maksimoi pahimman tapauksen odotushyödyn annetun posteriorijoukon \mathcal{M}_t yli. Tästä syystä X_t voidaan kirjoittaa muodossa

$$X_t = \max_{a \in \{a_0, a_1, a_2\}} \min_{\mu \in \mathcal{M}_t} \int u(a, \theta) d\mu(\theta).$$

Kuten luvussa 3 on määritelty, parametrin tila-avaruus on binäärinen,

$$\Theta = \{\theta_0, \theta_1\},$$

ja posteriorijoukko ajan t informaatiolla on

$$\mathcal{M}_t = \left\{ \mu : \mu = (1-m)\delta_{\theta_0} + m\delta_{\theta_1}, \underline{m}_t \leq m \leq \bar{m}_t \right\}.$$

Tällöin mille tahansa $\mu \in \mathcal{M}_t$ ja mille tahansa valinnalle a pätee

$$\int u(a, \theta) d\mu(\theta) = u(a, \theta_0)\mu(\theta_0) + u(a, \theta_1)\mu(\theta_1) = (1-m)u(a, \theta_0) + mu(a, \theta_1),$$

koska $\mu(\theta_0) = 1 - m$ ja $\mu(\theta_1) = m$.

Käyttämällä notaatiota $u_{ij} = u(a_i, \theta_j)$ saadaan erityisesti valinnalle a_0

$$\int u(a_0, \theta) d\mu(\theta) = (1 - m)u_{00} + mu_{01} = (u_{00} - u_{01})(1 - m) + u_{01}.$$

Koska oikea puoli on muotoa

$$f(m) = (u_{00} - u_{01})(1 - m) + u_{01},$$

eli lineaarinen funktio muuttujasta m . Koska $u_{00} > u_{01}$, funktio f on aidosti vähenävä, joten minimi välillä $[m_t, \bar{m}_t]$ saavutetaan kohdassa $m = \bar{m}_t$. Tästä saadaan seuraava muoto:

$$\min_{\mu \in \mathcal{M}_t} \int u(a_0, \theta) d\mu(\theta) = (u_{00} - u_{01})(1 - \bar{m}_t) + u_{01}.$$

Vastaavasti valinnalle a_1 pätee

$$\int u(a_1, \theta) d\mu(\theta) = (1 - m)u_{10} + mu_{11} = (u_{11} - u_{10})m + u_{10}.$$

Kun $u_{11} > u_{10}$, lauseke on kasvava m :n suhteen, joten minimointi $\mu \in \mathcal{M}_t$ toteutuu kohdassa $m = \underline{m}_t$. Siis

$$\min_{\mu \in \mathcal{M}_t} \int u(a_1, \theta) d\mu(\theta) = (u_{11} - u_{10})\underline{m}_t + u_{10}.$$

Edellä oleva lasku kattaa lauseen 5 tapauksen (a.ii), jossa pysäytyshetkellä relevantit valinnat ovat a_0 ja a_1 . Tällöin pysäytyspalkkio X_t määräytyy näiden kahden valinnan pahimman tapauksen tuottojen maksimin perusteella. Tarkastellaan siksi ensin, millä posteriorin arvoilla a_0 ja a_1 tuottavat saman robustin pysäytyspalkkion.

Edellä johdetut lausekkeet ovat

$$\min_{\mu \in \mathcal{M}_t} \int u(a_0, \theta) d\mu(\theta) = (u_{00} - u_{01})(1 - \bar{m}_t) + u_{01},$$

ja

$$\min_{\mu \in \mathcal{M}_t} \int u(a_1, \theta) d\mu(\theta) = (u_{11} - u_{10})\underline{m}_t + u_{10}.$$

Tuotto-symmetrian tapauksessa (a.ii) näitä voidaan vertailla saman posterioriarvon m funktiona. Asettamalla edellä olevat lausekkeet yhtäsuuriksi saadaan kriittinen posterioritaso $\bar{\pi}$, joka ratkaisee yhtälön

$$(u_{00} - u_{01})(1 - m) + u_{01} = (u_{11} - u_{10})m + u_{10}.$$

Ratkaisemalla m saadaan

$$\bar{\pi} = \frac{u_{00} - u_{10}}{u_{00} + u_{11} - u_{01} - u_{10}}.$$

Koska oletuksilla $u_{00} > u_{01}$ ja $u_{11} > u_{10}$ edellinen a_0 :n tuotto on vähenevä ja a_1 :n tuotto kasvava m :n suhteen, seuraa että pysäytyspalkkio saa paloittaisen muodon

$$X_t = \begin{cases} (u_{00} - u_{01})(1 - \bar{m}_t) + u_{01}, & \text{kun } \bar{m}_t < \bar{\pi}, \\ (u_{11} - u_{10})\underline{m}_t + u_{10}, & \text{kun } \underline{m}_t \geq \bar{\pi}. \end{cases}$$

Edellä on tunnistettu kriittinen posterioritaso $\bar{\pi}$, jolla valintojen a_0 ja a_1 robustit tuotot ovat yhtä suuret. Seuraavaksi tarkastellaan, milloin vakiohyödyn u_2 tuottava valinta a_2 voi tulla optimaaliseksi pysäytyshetkellä.

Koska pysäytyspalkkio on muotoa

$$X_t = \max\{g_0(\cdot), g_1(\cdot), u_2\},$$

missä g_0 ja g_1 ovat edellä määritellyt valintojen a_0 ja a_1 robustit tuotot, valinta a_2 ei ole optimaalinen millään posteriorilla täsmälleen silloin, kun $u_2(m) \leq \max\{g_0(m), g_1(m)\}$ kaikilla m .

Koska funktio g_0 on vähenevä ja g_1 kasvava, niiden pisteittäinen maksimi saavuttaa minimiarvonsa kohdassa $m = \bar{\pi}$. Tämän perusteella määritellään kriittinen vakiohyöty

$$u_2^* = g_0(\bar{\pi}) = g_1(\bar{\pi}) = \frac{u_{11}u_{00} - u_{10}u_{01}}{u_{00} + u_{11} - u_{01} - u_{10}}.$$

Jos taas $u_2 \leq u_2^*$, päätös a_2 ei ole optimaalinen millään posteriorilla, jolloin pysäytyspalkkio säilyy edellä saadussa muodossa.

Jos sen sijaan $u_2 > u_2^*$, vakiohyödyn u_2 tuottava valinta a_2 on optimaalinen niillä posteriorin arvoilla, joilla se ylittää molemmat valinnat a_0 ja a_1 . Tämä johtaa kolmeen erilliseen alueeseen, joissa pysäytyshetkellä optimaalinen valinta on vastaavasti a_0 , a_2 ja a_1 , ja pysäytyspalkkio saa kolmiosaisen muodon

$$X_t = \begin{cases} (u_{00} - u_{01})(1 - \bar{m}_t) + u_{01}, & \text{kun } \bar{m}_t < \frac{u_{00} - u_2}{u_{00} - u_{01}}, \\ (u_{11} - u_{10})\underline{m}_t + u_{10}, & \text{kun } \underline{m}_t \geq \frac{u_2 - u_{10}}{u_{11} - u_{10}}, \\ u_2, & \text{muulloin.} \end{cases}$$

Pysäytyspalkkion rakenne määrää samalla pahimman tapauksen mitan. Kullakin alueella pahin tapaus valitsee posteriorin siten, että odotushyöty minimoituu. Valintaan a_0 liittyvällä alueella valitaan yläpääte piste \bar{m}_t , ja valintaan a_1 liittyvällä alueella alapääte piste \underline{m}_t . Tästä seuraa, että pahimman tapauksen mitta P^* on yksikäsitteinen ja sillä on tiheysgeneraattori (seuraa aikaisemmin esitetystä CE-mallista)

$$-\eta_t = (\tilde{\theta}_t^{\bar{u}}/\sigma)\mathbf{1}_{\{Z_t \leq \tilde{z}_t\}} + (\tilde{\theta}_t^u/\sigma)\mathbf{1}_{\{Z_t > \tilde{z}_t\}}, \quad (5.3)$$

missä \tilde{z}_t on kriittinen signaalin realisaatio, jolla posteriori saavuttaa tason $\bar{\pi}$. Myöhemmin osoitetaan, että P^* on pahin tapaus koko luokassa \mathcal{P}_0 .

Todistus tapaukselle (a.ii)

Tarkastellaan ensin klassista optimaalisen pysäytyksen ongelmaa pahimman tapauksen mitan P^* alla:

$$\sup_{\tau} \mathbb{E}_{P^*}[X_{\tau} - c\tau]. \quad (5.4)$$

Tässä P^* on edellä määritelty pahimman tapauksen todennäköisyysmitta, ja pysäytyspalkkio X_t on johdettu kohdassa 5.1.

Koska posterioritodennäköisyys $m_t \in (0, 1)$ kuvaa kaikkea saatavilla olevaa informaatiota parametrin θ arvosta, arvofunktiota voidaan esittää sen funktiona. Lisäksi ongelma ei riipu eksplisiittisesti ajasta, joten oletetaan

$$v(t, z) = V(m_t).$$

Tällöin arvofunktiota toteuttaa jatkoalueella HJB-yhtälön, joka ajasta riippumattomuuden vuoksi supistuu yksiulotteiseksi differentiaaliyhtälöksi

$$-c + \frac{1}{2}\hat{c}^2 y^2 (1-y)^2 V''(y) = 0,$$

missä $y = m_t$ ja \hat{c} on pahimman tapauksen mitasta P^* johtuva vakio.

Tämän yhtälön yleinen ratkaisu on muotoa

$$g(y; C_{2i-1}, C_{2i}) = \hat{c}(2y-1) \log\left(\frac{y}{1-y}\right) + C_{2i-1}y + C_{2i}, \quad i = 1, 2, \quad (5.5)$$

missä vakiot C_1, \dots, C_4 määräytyvät myöhemmin sileysehdosta.

Näiden perusteella oletetaan, että arvofunktiota voidaan kirjoittaa paloittain seuraavassa muodossa:

$$v(t, z) = \begin{cases} (u_{00} - u_{01})(1 - \bar{m}_t(z)) + u_{01}, & z < \tilde{f}(t, r^l), \\ g_1(\bar{m}_t(z); C_1, C_2), & \tilde{f}(t, r^l) \leq z < \tilde{z}_t, \\ g_2(m_t(z); C_3, C_4), & \tilde{z}_t \leq z < \tilde{f}(t, r^R), \\ (u_{11} - u_{10})m_t(z) + u_{10}, & \tilde{f}(t, r^R) \leq z, \end{cases} \quad (5.6)$$

missä $\tilde{f}(t, r^l)$ ja $\tilde{f}(t, r^R)$ ovat pysäytysrajat, jotka määritellään osana ratkaisua.

Vakiot C_1, \dots, C_4 määräytyvät edellytetystä arvofunktion jatkuvasta differentioituvuudesta pysäytysreunojen yli. Erityisesti pätee

$$C_1 = -\hat{c}\ell(\bar{\pi}), \quad C_2 = (u_{00} - u_{01})(1 - r^l) + u_{01} - \hat{c}(2r^l - 1) \log\left(\frac{r^l}{1-r^l}\right) - \ell(\bar{\pi})r^l,$$

$$C_3 = -\hat{c}\ell(\underline{\pi}), \quad C_4 = (u_{11} - u_{10})r^R + u_{10} - \hat{c}(2r^R - 1) \log\left(\frac{r^R}{1-r^R}\right) - \ell(\underline{\pi})r^R.$$

Huomataan lisäksi, että kriittinen vakiohyöty u_2^{**} , joka määriteltiin kohdassa (4.5), toteuttaa

$$u_2^{**} = g_1(\bar{\pi}; C_1, C_2) = g_2(\bar{\pi}; C_3, C_4) = v(t, \tilde{z}_t),$$

eli u_2^{**} vastaa arvofunktion arvoa kriittisessä pisteessä \tilde{z}_t .

Lemma 3. *Arvofunktio v on klassisen optimaalisen pysäytyksen ongelman (5.4) arvofunktio. Toisin sanoen, kaikilla $t \geq 0$ pätee*

$$v(t, z) = \sup_{\tau \geq t} \mathbb{E}_{P^*} [X_\tau - c(\tau - t) \mid Z_t = z].$$

Lisäksi arvofunktio v toteuttaa HJB-yhtälön

$$\max \left\{ X(t, z) - v(t, z), -c + v_t(t, z) + \frac{1}{2} \sigma^2 v_{zz}(t, z) + f(t, z) v_z(t, z) \right\} = 0, \quad (5.7)$$

missä ajautumatermi $f(t, z)$ määräytyy pahimman tapauksen mitan P^ mukaisesti ja on muotoa*

$$f(t, z) = \left(\theta_1 - \frac{\theta_1 - \theta_0}{1 + \frac{\bar{m}_0}{1 - \bar{m}_0} \varphi(t, z)} \right) \mathbf{1}_{\{z < \bar{z}_t\}} + \left(\theta_1 - \frac{\theta_1 - \theta_0}{1 + \frac{m_0}{1 - m_0} \varphi(t, z)} \right) \mathbf{1}_{\{z \geq \bar{z}_t\}}. \quad (5.8)$$

Jatkoalueella, eli kun z kuuluu pysäytysrajojen väliin $z \in (\tilde{f}(t, r^l), \tilde{f}(t, r^R))$, pysäytystä ei ole vielä tapahtunut ja HJB-yhtälö supistuu muotoon

$$-c + v_t(t, z) + \frac{1}{2} \sigma^2 v_{zz}(t, z) + f(t, z) v_z(t, z) = 0. \quad (5.9)$$

Todistus. Todistetaan, että v on ongelman (5.4) arvofunktio ja v toteuttaa variaatioepäyhtälön (HJB) ja jatkoalueella siitä seuraavan osittaisdifferentiaaliyhtälön.

Olkoon P^* pahimman tapauksen mitta ja oletetaan, että Z on sen alla diffuusio-prosessi

$$dZ_s = f(s, Z_s) ds + \sigma dW_s^{P^*},$$

missä f on (5.8):n mukainen paloittainen ajautuma ja $\sigma > 0$ vakio. Oletetaan lisäksi, että kandidaattifunktio v on jatkuva kaikkialla ja jatkoalueella kerran jatkuvasti derivoituva ajan t suhteen sekä kahdesti jatkuvasti derivoituva muuttujan z suhteen, ja että se täyttää sileys-ehdot (kuten luvussa 5 johdetaan).

Määritellään klassisen optimaalisen pysäytyksen arvofunktio

$$V(t, z) = \sup_{\tau \geq t} \mathbb{E}_{P^*} [X(\tau, Z_\tau) - c(\tau - t) \mid Z_t = z]. \quad (A)$$

Koska Z on Markov-prosessi ja palkkio on additiivinen ajassa, pätee jokaiselle pienelle $h > 0$ (Peskir-Shiryaevin DPP-optimaalisen pysäytyksen tuloksista)[20] rekursio

$$V(t, z) = \max \{ X(t, z), \mathbb{E}_{P^*} [V(t + h, Z_{t+h}) - ch \mid Z_t = z] \}. \quad (B)$$

Intuitio: joko pysäytetään heti ja saadaan $X(t, z)$ tai jatketaan ainakin aikaan $t + h$ ja maksetaan kustannus ch , minkä jälkeen toimitaan optimaalisesti.

Oletetaan nyt, että kandidaatti v toteuttaa ehdot

$$v(t, z) \geq X(t, z) \quad \text{kaikilla } (t, z), \quad (C1)$$

ja että jatkoalueella $\mathcal{C} = \{(t, z) : v(t, z) > X(t, z)\}$ pätee osittaisdifferentiaaliyhtälö

$$-c + v_t(t, z) + \frac{1}{2} \sigma^2 v_{zz}(t, z) + f(t, z) v_z(t, z) = 0. \quad (C2)$$

Osoitetaan, että edellä johdetuista ehdoista seuraa variaatioepäyhtälö (5.7), eli

$$\max \{ X(t, z) - v(t, z), -c + v_t + \frac{1}{2} \sigma^2 v_{zz} + f v_z \} = 0.$$

Huomataan, että

- jos $(t, z) \in \mathcal{C}$, niin $X(t, z) - v(t, z) < 0$ ja ehto (C2) antaa toisen termin arvoksi nollan, jolloin maksimi on nolla
- jos (t, z) kuuluu pysäytysalueelle $\mathcal{S} = \{v = X\}$, niin $X(t, z) - v(t, z) = 0$ ja maksimi on jälleen nolla riippumatta differentiaaliyhtälön termistä.

Tämä on yhtälö (5.7).

Määritellään prosessi

$$M_s = v(s, Z_s) - c(s - t), \quad s \geq t.$$

Näytetään, että $M_{s \wedge \tau}$ on P^* -supermartingaali kaikille pysäytysajoille $\tau \geq t$, ja että sopivalla τ^* se on martingaali. Tämä antaa $v = V$.

(a) *Supermartingaali.* Olkoon $\tau \geq t$ mielivaltainen pysäytysaika ja tarkastellaan ennen pysäytystä prosessia $s \mapsto v(s, Z_s)$ jatkoalueella. Itön kaavalla saadaan

$$dv(s, Z_s) = \left(v_t + \frac{1}{2} \sigma^2 v_{zz} + f v_z \right) (s, Z_s) ds + \sigma v_z(s, Z_s) dW_s^{P^*}.$$

Tästä seuraa

$$dM_s = \left(v_t + \frac{1}{2} \sigma^2 v_{zz} + f v_z - c \right) (s, Z_s) ds + \sigma v_z(s, Z_s) dW_s^{P^*}.$$

Variaatioepäyhtälöstä (5.7) seuraa, että jatkoalueella $v_t + \frac{1}{2} \sigma^2 v_{zz} + f v_z - c = 0$, ja pysäytysalueella emme tarvitse differentioituvuutta, koska pysäytämme. Siten $M_{s \wedge \tau}$ on lokaali martingaali jatkoalueella ja rajoittuneisuusehdoilla supermartingaali. Tästä saadaan valinnaisen pysäytyksen teoreemalla

$$\mathbb{E}_{P^*}[M_\tau \mid Z_t = z] \leq M_t = v(t, z). \quad (\text{D})$$

(b) *Yläraja arvolle.* Koska $v(\tau, Z_\tau) \geq X(\tau, Z_\tau)$ kaikilla τ (ehto (C1)), yhtälöstä (D) seuraa kaikille pysäytysajoille $\tau \geq t$:

$$\mathbb{E}_{P^*}[X(\tau, Z_\tau) - c(\tau - t) \mid Z_t = z] \leq \mathbb{E}_{P^*}[v(\tau, Z_\tau) - c(\tau - t) \mid Z_t = z] \leq v(t, z).$$

Ottamalla supremumin $\tau \geq t$ saadaan

$$V(t, z) \leq v(t, z). \quad (\text{E})$$

(c) *Optimaalinen pysäytysaika ja alaraja.* Määritellään ensimmäinen osuma pysäytysalueelle

$$\tau^* = \inf \{s \geq t : v(s, Z_s) = X(s, Z_s)\}.$$

Jatkoalueella ajautumaosa on nolla, joten $M_{s \wedge \tau^*}$ on martingaali. Valinnaisen pysäytyksen teoreemalla saadaan

$$v(t, z) = \mathbb{E}_{P^*}[v(\tau^*, Z_{\tau^*}) - c(\tau^* - t) \mid Z_t = z].$$

Määritelmän vuoksi $v(\tau^*, Z_{\tau^*}) = X(\tau^*, Z_{\tau^*})$, joten

$$v(t, z) = \mathbb{E}_{P^*}[X(\tau^*, Z_{\tau^*}) - c(\tau^* - t) \mid Z_t = z] \leq V(t, z). \quad (\text{F})$$

Yhdistämällä (E) ja (F) saadaan $v(t, z) = V(t, z)$ ja τ^* on optimaalinen.

Kun z on pysäytysrajojen välissä, eli jatkoalueella $z \in (\tilde{f}(t, r^l), \tilde{f}(t, r^R))$, pätee $v > X$ ja (5.7) supistuu toiseksi termiksi, jolloin saadaan ODY (5.9):

$$-c + v_t(t, z) + \frac{1}{2}\sigma^2 v_{zz}(t, z) + f(t, z)v_z(t, z) = 0.$$

Edellä on osoitettu, että kandidaattifunktio v on ongelman (5.4) arvofunktiio ja että se toteuttaa HJB-yhtälön (5.8) sekä jatkoalueen ODY:n (5.9). \square

Nyt Lemma 3 ollaan saatu todistettua, joten voidaan jatkaa lauseen 5 todistusta hyödyntäen tätä tulosta.

Seuraavaksi osoitetaan, että funktio v on myös ei-klassisen optimaalisen pysäytysongelman (3.1) arvofunktiio. Pelkkä HJB-yhtälön ratkaiseminen ei vielä riitä tämän osoittamiseen. Rajoitutaan tarkastelemaan ajanhetkeä $t = 0$ ja näytetään, että

$$v(0, z) = \sup_{\tau \geq 0} \inf_{P \in \mathcal{P}_0} \mathbb{E}_P[X(Z_\tau) - c\tau].$$

Lemman 3 nojalla tiedetään, että

$$v(0, z) = \sup_{\tau \geq 0} \mathbb{E}_{P^*}[X(Z_\tau) - c\tau] \geq \sup_{\tau \geq 0} \inf_{P \in \mathcal{P}_0} \mathbb{E}_P[X(Z_\tau) - c\tau],$$

joten jäljellä on vastakkaisen epäyhtälön osoittaminen.

Määritellään pysäytysaika

$$\tau^* = \inf\{t \geq 0 : Z_t \leq \tilde{f}(t, r^l) \text{ tai } Z_t \geq \tilde{f}(t, r^R)\},$$

eli ensimmäinen hetki, jolloin prosessi Z_t osuu pysäytysalueelle.

Kun $t \leq \tau^*$, ollaan jatkoalueella, ja Itön kaavaa käyttäen yhdessä HJB-yhtälön (5.7) ja jatkoalueen yhtälön (5.9) kanssa saadaan

$$\begin{aligned} dv(t, Z_t) &= \left[v_t(t, Z_t) + \frac{1}{2}\sigma^2 v_{zz}(t, Z_t) \right] dt + v_z(t, Z_t) dZ_t \\ &= [c - f(t, Z_t)v_z(t, Z_t)] dt + v_z(t, Z_t) dZ_t. \end{aligned} \quad (5.10)$$

Olkoon nyt $P^\eta \in \mathcal{P}_0$ mielivaltainen todennäköisyysmitta ja η_t sitä vastaava tiheysgeneraattoriprosessi. Tällöin Z_t voidaan kirjoittaa muodossa

$$dZ_t = \tilde{f}(t, Z_t, \eta_t) dt + \sigma dW_t^\eta,$$

missä W^η on Brownin liike mitan P^η alla ja

$$\tilde{f}(t, Z_t, \eta_t) = \theta_1 - \frac{\theta_1 - \theta_0}{1 + \frac{\eta_t}{1 - \eta_t} \varphi(t, Z_t)}.$$

Sijoittamalla tämä edelliseen saadaan

$$dv(t, Z_t) = \left[c + (\tilde{f}(t, Z_t, \eta_t) - f(t, Z_t))v_z(t, Z_t) \right] dt + \sigma v_z(t, Z_t) dW_t^\eta.$$

Huomataan, että

$$(\tilde{f}(t, Z_t, \eta_t) - f(t, Z_t))v_z(t, Z_t) \geq 0.$$

Todellakin, jos $Z_t < \tilde{z}_t$, niin $v_z(t, Z_t) \leq 0$ ja $\tilde{f}(t, Z_t, \eta_t) - f(t, Z_t) \leq 0$, koska $\theta_1 - \frac{\theta_1 - \theta_0}{1 + \frac{m_0}{1 - m_0} \varphi(t, z)}$ on kasvava funktio posteriorista. Vastaava argumentti pätee tapauksessa $Z_t > \tilde{z}_t$.

Ottamalla odotusarvo mitan P^η alla ja integroimalla ajanhetkeen τ^* saadaan

$$v(0, z) \leq \mathbb{E}_{P^\eta}[v(\tau^*, Z_{\tau^*}) - c\tau^*] = \mathbb{E}_{P^\eta}[X_{\tau^*} - c\tau^*],$$

Koska pysäytysaika τ^* on määritelty ensimmäisenä hetkenä, jolloin pysäytys on optimaalinen, arvofunktion arvo yhtyy pysäytyspalkkioon hetkellä τ^* .

Stokastinen integraalitermi katoaa odotusarvossa, koska

$$\mathbb{E}_{P^\eta} \left[\int_0^{\tau^*} \sigma v_z(t, Z_t) dW_t^\eta \right] = 0,$$

mikä on taattu ehdosta

$$\sup_{P \in \mathcal{P}_0} \mathbb{E}_P[\tau^*] < \infty. \quad (5.11)$$

Tämä ehto seuraa tässä asetelmassa pysäytyspalkkion X_t rajoittuneisuudesta.

Koska P^η oli mielivaltainen mitta joukossa \mathcal{P}_0 , saadaan

$$v(0, z) \leq \inf_{P \in \mathcal{P}_0} \mathbb{E}_P[X_{\tau^*} - c\tau^*] \leq \sup_{\tau \geq 0} \inf_{P \in \mathcal{P}_0} \mathbb{E}_P[X(Z_\tau) - c\tau].$$

Yhdessä aiemmin saadun vastakkaisen epäyhtälön kanssa tämä osoittaa, että

$$v(0, z) = \sup_{\tau \geq 0} \inf_{P \in \mathcal{P}_0} \mathbb{E}_P[X(Z_\tau) - c\tau],$$

eli v on robustin optimaalisen pysäytysongelman (3.1) arvofunktio ja τ^* on optimaalinen pysäytysaika.

Todistus tapaukselle (a.i)

Todistus on rakenteeltaan samanlainen kuin tapauksessa (a.ii). Ainoa ero on, että arvofunktio v saa hieman erilaisen paloittaisen muodon. Tarkemmin, arvofunktio voidaan kirjoittaa muodossa

$$v(t, z) = \begin{cases} (u_{00} - u_{01})(1 - \bar{m}_t(z)) + u_{01}, & z < \tilde{f}(t, r_1^l), \\ g_3(\bar{m}_t(z); C_5, C_6), & \tilde{f}(t, r_1^l) \leq z < \tilde{f}(t, r_1^l), \\ u_2, & \tilde{f}(t, r_1^l) \leq z < \tilde{f}(t, r_1^R), \\ g_4(m_t(z); C_7, C_8), & \tilde{f}(t, r_1^R) \leq z < \tilde{f}(t, r_2^R), \\ (u_{11} - u_{10})m_t(z) + u_{10}, & \tilde{f}(t, r_2^R) \leq z, \end{cases} \quad (5.12)$$

missä funktiot g_3 ja g_4 ovat muodoltaan samat kuin funktiot g_1 ja g_2 kohdassa (5.5), mutta niissä esiintyvät vakiot poikkeavat. Täsmällisesti vakioiden C_1, \dots, C_4 tilalle tulevat vakiot C_5, \dots, C_8 , jotka ovat

$$\begin{aligned} C_5 &= -\hat{c}\ell(r_1^l), & C_7 &= -\hat{c}\ell(r_1^R), \\ C_6 &= u_2 - \hat{c}(2r_1^l - 1) \log\left(\frac{r_1^l}{1-r_1^l}\right) - \ell(r_1^l)r_1^l, \\ C_8 &= u_2 - \hat{c}(2r_1^R - 1) \log\left(\frac{r_1^R}{1-r_1^R}\right) - \ell(r_1^R)r_1^R. \end{aligned}$$

Muilta osin todistus etenee täsmälleen kuten tapauksessa (a.ii): edellä määritelty kandidaattifunktio toteuttaa HJB-yhtälön jatkoalueella sekä pysäytysalueiden jatkuvuus- ja sileysvaatimukset, minkä perusteella se on ongelman arvofunktiio.

Todistus tapaukselle (b)

Tapauksessa (b) valintaa a_2 ei ole milloinkaan optimaalista valita, joten se voidaan poistaa sallittujen valintojen joukosta. Todistus etenee tämän jälkeen kuten tapauksessa (a.ii), mutta arvofunktiio saa seuraavan muodon:

$$v(t, z) = \begin{cases} (u_{00} - u_{01})(1 - \bar{m}_t(z)) + u_{01}, & z < \tilde{f}(t, r_1^l), \\ g_5(m_t(z); C_9, C_{10}), & \tilde{f}(t, r_1^l) \leq z < \tilde{z}_t, \\ g_6(m_t(z); C_{11}, C_{12}), & \tilde{z}_t \leq z < \tilde{f}(t, r_1^R), \\ (u_{11} - u_{10})m_t(z) + u_{10}, & \tilde{f}(t, r_1^R) \leq z. \end{cases}$$

Tässä funktiot g_5 ja g_6 ovat muodoltaan samat kuin funktiot g_1 ja g_2 kohdassa (5.5), mutta vakioiden C_1, \dots, C_4 tilalle tulevat vakiot C_9, \dots, C_{12} , jotka määräytyvät seuraavasti:

$$\begin{aligned} C_9 &= -\hat{c}\ell(\bar{r}^R) + u_{11} - u_{10}, & C_{11} &= -\hat{c}\ell(\bar{r}^l) + u_{01} - u_{00}, \\ C_{10} &= u_{10} - \hat{c}[1 - \ell(\bar{r}^R)], & C_{12} &= u_{00} - \hat{c}[1 - \ell(\bar{r}^l)]. \end{aligned}$$

Näillä vakioilla määritelty funktio v täyttää HJB-yhtälön jatkoalueella sekä pysäytysrajoilla vaaditut jatkuvuus- ja sileysominaisuudet. Näin ollen v on tapauksen (b) mukaisen optimaalisen pysäytysongelman arvofunktiio. \square

Tämän jälkeen lähdetään todistamaan muita tutkielman tärkeitä tuloksia.

5.2 Lemma 2:n todistus

Yhtälöiden (4.3) ja (4.4) ratkaisujen yksikäsitteisyys

Derivoidaan l :

$$l'(r) = 2\left(\frac{1}{r} + \frac{1}{1-r}\right) + \frac{1}{r^2} + \frac{1}{(1-r)^2} > 0, \quad r \in (0, 1).$$

Siis l on aidosti kasvava ja jatkuva välillä $(0, 1)$. Lisäksi

$$\lim_{r \downarrow 0} l(r) = -\infty, \quad \lim_{r \uparrow 1} l(r) = +\infty.$$

Tästä seuraa, että $l : (0, 1) \rightarrow \mathbb{R}$ on bijektio. Siksi kummallakin yhtälöllä

$$l(r^R) - l(\pi) = \frac{u_{11} - u_{10}}{\hat{c}}, \quad l(r^l) - l(\pi) = -\frac{u_{00} - u_{01}}{\hat{c}}$$

on täsmälleen yksi ratkaisu $r^R, r^l \in (0, 1)$. Tämä todistaa (4.3):n olemassaolon ja yksikäsitteisyyden.

Kiinnitetään $\hat{r}^l \in (0, \pi)$. Etsitään funktiota $V^l(\cdot; \hat{r}^l)$, joka ratkaisee

$$\partial_{yy} V^l(y; \hat{r}^l) = \hat{c} l'(y), \quad y \in (0, 1),$$

sekä täyttää ehdot

$$V^l(\hat{r}^l; \hat{r}^l) = (u_{00} - u_{01})\hat{r}^l + u_{00}, \quad \partial_y V^l(\hat{r}^l; \hat{r}^l) = u_{01} - u_{00}.$$

Integroimalla kerran saadaan

$$\partial_y V^l(y; \hat{r}^l) = \hat{c} l(y) + C_1,$$

missä C_1 on vakio.

Integroimalla uudelleen saadaan

$$V^l(y; \hat{r}^l) = \hat{c} \int_{y_0}^y l(s) ds + C_1 y + C_2,$$

missä C_2 on vakio.

Vakio C_1 määräytyy derivaattaehdosta:

$$C_1 = (u_{01} - u_{00}) - \hat{c} l(\hat{r}^l).$$

Tämän jälkeen arvoehto määrää yksikäsitteisesti vakion C_2 . Siis jokaiselle $\hat{r}^l \in (0, \pi)$ ratkaisu $V^l(\cdot; \hat{r}^l)$ on olemassa ja yksikäsitteinen. Lisäksi

$$\partial_{yy} V^l(y; \hat{r}^l) = \hat{c} \hat{l}''(y) = \hat{c} l'(y) > 0,$$

eli $V^l(\cdot; \hat{r}^l)$ on aidosti konvekksi.

Määritellään seuraavaksi funktio $V^R(\cdot; \hat{r}^l)$ välille $[\pi, 1)$ muodossa

$$V^R(y; \hat{r}^l) = \hat{c} \hat{l}(y) + C_3 y + C_4,$$

missä vakiot C_3, C_4 valitaan niin, että V^R on C^1 -jatkuva pisteessä π :

$$V^R(\pi; \hat{r}^l) = V^l(\pi; \hat{r}^l), \quad \partial_y V^R(\pi; \hat{r}^l) = \partial_y V^l(\pi; \hat{r}^l).$$

Nämä ehdot määräävät C_3, C_4 yksikäsitteisesti, ja samoin

$$\partial_{yy} V^R(y; \hat{r}^l) = \hat{c} l'(y) > 0,$$

eli $V^R(\cdot; \hat{r}^l)$ on aidosti konvekksi.

Määritellään nyt suora

$$L_1(y) = (u_{11} - u_{10})y + u_{10}.$$

Konveksisuuden nojalla yhtälöpari

$$V^R(\hat{r}^R; \hat{r}^l) = L_1(\hat{r}^R), \quad \partial_y V^R(\hat{r}^R; \hat{r}^l) = u_{11} - u_{10}$$

tarkoittaa, että L_1 on funktion $V^R(\cdot; \hat{r}^l)$ tangenti pisteessä $\hat{r}^R \in (\pi, 1)$. Konveksisuudesta seuraa, että jos tällainen \hat{r}^R on olemassa, se on yksikäsitteinen.

Koska vakiot C_1, C_2, C_3, C_4 riippuvat jatkuvasti parametrilla \hat{r}^l , riippuu myös $V^R(\cdot; \hat{r}^l)$ jatkuvasti tästä parametrilla. Tästä seuraa väliarvolauseargumentin avulla, että on olemassa $\tilde{r}^l \in (0, \pi)$, jolle tangenttiehto toteutuu jollain $\tilde{r}^R \in (\pi, 1)$.

Näiden perusteella saamme yksikäsitteisyyden. Rakennelma on deterministinen: parametri \hat{r}^l määrää yksikäsitteisesti funktion V^l , siitä seuraa yksikäsitteisesti V^R , ja tangenttiehto määrää \hat{r}^R yksikäsitteisesti konveksisuuden vuoksi. Lisäksi tangenttiehdon toteuttava \hat{r}^l voi olla korkeintaan yksi jatkuvuuden ja konveksisuuden yhdistelmän vuoksi. Näin ratkaisupari $(\tilde{r}^R, \tilde{r}^l)$ on yksikäsitteinen ja toteuttaa yhtälöryhmän (4.4).

Osoitetaan seuraavaksi Lemman jäljellä olevat väitteet.

(i) Järjestyshdot (4.6). Ratkaisut r_i^l, r_i^R määräytyvät yhtälöistä muotoa

$$l(r) - l(\pi) = C_i,$$

missä vakio C_i riippuu tarkasteltavasta indeksistä $i \in \{1, 2\}$ ja vastaavasta palkkiorakenteesta. Koska funktio l on aidosti kasvava välillä $(0, 1)$, on ratkaisupiste r aidosti kasvava funktiona vakiosta C_i .

Koska indeksin 2 tapauksessa palkkiorakenne on epäedullisempi kuin indeksin 1 tapauksessa, niitä vastaavat vakiot toteuttavat

$$C_2 < C_1.$$

Tästä seuraa suoraan

$$r_2^l < r_1^l \quad \text{ja} \quad r_1^R < r_2^R.$$

(ii) Arvon u_2^* ja kynnyksen \tilde{y} määrittely (4.1)-(4.2). Määritelmän ja kohdan (4.3) perusteella saadaan, että

$$u_2^* > u_{01}.$$

Täten voidaan määritellä

$$\tilde{y} = \frac{u_{00} - u_2^*}{u_{00} - u_{01}} \in (0, 1).$$

Määritellään tätä vastaava paloittain määritelty palkkiofunktio

$$V(y) = \begin{cases} -(u_{00} - u_{01})y + u_{00}, & y \in (0, \tilde{y}), \\ u_2^*, & y \in (\tilde{y}, 1). \end{cases}$$

Tämä funktio on jatkuva ja muodostaa vertailupalkkion, jonka avulla voidaan tarkastella optimaalisen pysäytyksen ongelmaa indeksin 2 tapauksessa.

Rakenteellisesti kyse on samasta vapaaseen reunaan johtavasta ongelmasta kuin kohdassa (4.4), ja edellä esitetty olemassaolo- ja yksikäsitteisyysargumentti pätee sellaisenaan. Tästä seuraa, että myös yhtälölle (4.2) on olemassa yksikäsitteinen ratkaisupiste pari (r_2^l, r_2^R) .

Vastaava argumentti indeksin 1 tapauksessa antaa ratkaisun yhtälölle (4.1).

(4.7) Palkkiorakenteen symmetrian ja määritelmien perusteella pätee

$$\underline{\pi} + \bar{\pi} = 1.$$

Lisäksi funktiolle l pätee identiteetti

$$l(1-r) = -l(r), \quad r \in (0, 1),$$

mikä seuraa suoraan sen eksplisiittisestä muodosta.

Yhtälöistä (4.3) saadaan

$$l(r^R) - l(\pi) = C \quad \text{ja} \quad l(r^l) - l(\pi) = -C$$

jollakin vakiolla C . Yhdistämällä nämä ja käyttäen edellä mainittua symmetriaa saadaan

$$l(r^R) = -l(r^l),$$

mistä seuraa

$$r^R = 1 - r^l.$$

Täten ratkaisut toteuttavat symmetriaehdon

$$r^l + r^R = 1.$$

(4.8) Osoitetaan ensin identiteetti

$$\frac{1}{2} l(r) = \tilde{l}(r) - \frac{1}{2r(1-r)} + 1, \quad r \in (0, 1).$$

Määritelmistä

$$l(r) = 2 \log\left(\frac{r}{1-r}\right) - \frac{1}{r} + \frac{1}{1-r}, \quad \tilde{l}(r) = \log\left(\frac{r}{1-r}\right) + \frac{r}{1-r}$$

saadaan

$$\frac{1}{2} l(r) = \log\left(\frac{r}{1-r}\right) - \frac{1}{2r} + \frac{1}{2(1-r)}.$$

Toisaalta

$$\begin{aligned} \tilde{l}(r) - \frac{1}{2r(1-r)} + 1 &= \log\left(\frac{r}{1-r}\right) + \frac{r}{1-r} - \frac{1}{2r(1-r)} + 1 \\ &= \log\left(\frac{r}{1-r}\right) + \frac{1}{1-r} - \frac{1}{2r(1-r)} \\ &= \log\left(\frac{r}{1-r}\right) + \frac{2r-1}{2r(1-r)} \\ &= \log\left(\frac{r}{1-r}\right) - \frac{1}{2r} + \frac{1}{2(1-r)}, \end{aligned}$$

joten identiteetti pätee.

Tämän avulla voidaan kirjoittaa (4.1) uudelleen seuraavaanlaisesti. Koska lausekkeessa esiintyy erotus, integraalivakio kumoutuu. Saadaan

$$\frac{1}{2}(l(r_2^R) - l(r_1^R)) = \tilde{l}(r_2^R) - \tilde{l}(r_1^R) - \frac{1}{2r_2^R(1-r_2^R)} + \frac{1}{2r_1^R(1-r_1^R)}.$$

Sijoittamalla tämä (4.1):een saadaan ekvivalentti muoto

$$\tilde{l}(r_2^R) - \tilde{l}(r_1^R) = -\frac{1}{2r_2^R(1-r_2^R)} + \frac{1}{2r_1^R(1-r_1^R)} + \frac{u_{11} - u_{10}}{\hat{c}}, \quad \tilde{l}(r_2^R) - \tilde{l}(r_1^R) = \frac{u_2 - u_{10}}{\hat{c}},$$

mikä on väite (4.8).

Oletetaan seuraavaksi $u_2 = u_2^*$. Palkkiosymmetrian nojalla voidaan tarkistaa, että tällöin $r_2^R = r^R$ ja $r_1^R = \pi$ toteuttavat (4.1):n, ja koska ratkaisu on yksikäsitteinen, ne ovat ainoat ratkaisut.

Näytetään nyt, että ratkaisu r_1^R riippuu aidosti kasvavasti parametrusta u_2 . Merkitään ratkaisua muodossa $(r_1^R(u_2), r_2^R(u_2))$. Huomataan, että

$$l'(r) = \frac{1}{r^2(1-r)^2}, \quad \tilde{l}'(r) = \frac{1}{r(1-r)^2}, \quad r \in (0, 1).$$

Derivoidaan (4.1):n ensimmäinen yhtälö r_1^R :n suhteen ketjusäännöllä:

$$l'(r_2^R) \frac{dr_2^R}{dr_1^R} - l'(r_1^R) = 0.$$

Derivoidaan (4.1):n toinen yhtälö u_2 :n suhteen:

$$\tilde{l}'(r_2^R) \frac{dr_2^R}{du_2} - \tilde{l}'(r_1^R) \frac{dr_1^R}{du_2} = \frac{1}{\hat{c}}.$$

Koska $\frac{dr_2^R}{du_2} = \frac{dr_2^R}{dr_1^R} \frac{dr_1^R}{du_2}$, voidaan yhdistää edelliset ja ratkaista $\frac{dr_1^R}{du_2}$:

$$\left(\tilde{l}'(r_2^R) \frac{dr_2^R}{dr_1^R} - \tilde{l}'(r_1^R) \right) \frac{dr_1^R}{du_2} = \frac{1}{\hat{c}}.$$

Sijoittamalla $\frac{dr_2^R}{dr_1^R} = \frac{l'(r_1^R)}{l'(r_2^R)}$ sekä yllä olevat derivaatat saadaan

$$\frac{dr_1^R}{du_2} = \frac{(r_1^R)^2(1-r_1^R)^2}{\hat{c}(r_2^R - r_1^R)}.$$

Koska $\hat{c} > 0$ ja (4.1):stä seuraa $r_2^R > r_1^R$, pätee $\frac{dr_1^R}{du_2} > 0$. Siis $u_2 \mapsto r_1^R(u_2)$ on aidosti kasvava.

Edellä todetun ja tapauksen $u_2 = u_2^*$ nojalla seuraa ekvivalenssi

$$r_1^R \geq \pi \iff u_2 \geq u_2^*.$$

Vastaavasti (symmetrian avulla tai samanlaisella tarkastelulla vasemmalle rajalle) saadaan

$$r_1^l \leq \bar{\pi} \iff u_2 \geq u_2^*.$$

Tämä osoittaa kohdan (4.8).

(4.9) Koska funktio $f(t, \cdot)$ on aidosti kasvava, sillä on käänteisfunktio. Määritelmien nojalla pätee tällöin

$$m_t(f(t, r)) = r = \bar{m}_t(f(t, r)), \quad r \in (0, 1).$$

Lisäksi kynnysten määritelmistä $\bar{z}_t = f(t, \bar{\pi})$ ja $\underline{z}_t = f(t, \underline{\pi})$ sekä funktion $f(t, \cdot)$ aidosta kasvavuudesta seuraavat ekvivalenssit

$$\bar{f}(t, r_1) \leq \bar{z}_t \iff r_1 \leq \bar{\pi}, \quad f(t, r_2) \geq \underline{z}_t \iff r_2 \geq \underline{\pi}.$$

Tämä osoittaa väitteen.

Näin ollaan osoitettu jokainen Lemma 2:n kohta ja todistus on valmis. □

5.3 Sovelluksien todistukset

Lauseen 2 todistus

Kiinnitetään Ellsberg-tapauksessa parametrit siten, että

$$\hat{c} = \frac{c\sigma^2}{2\alpha^2}, \quad \bar{z}_t = 0, \quad \underline{\pi} = \frac{1-\varepsilon}{2}, \quad \bar{\pi} = \frac{1+\varepsilon}{2}.$$

Tällöin yhtälöt (4.1) ja (4.2) sievenevät muotoon

$$r_2^R + r_1^R = 1, \quad l(r_2^R) = \frac{2\alpha^3}{c\sigma^2}, \quad r_2^l + r_1^l = 1, \quad l(r_1^l) = \frac{2\alpha^3}{c\sigma^2},$$

missä on käytetty lisäksi sitä, että $u_2 = \frac{1}{2}(u_{00} + u_{10})$, ja funktiot \bar{f} ja f saavat eksplisiittiset muodot

$$\bar{f}(t, r) = \frac{\sigma^2}{2\alpha} \log\left(\frac{1-\varepsilon}{1+\varepsilon} \frac{r}{1-r}\right), \quad f(t, r) = \frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon} \frac{r}{1-r}\right).$$

1. Oletetaan ensin, että $r_1^l < \frac{1+\varepsilon}{2} = \bar{\pi}$. Tällöin

$$\bar{f}(t, r_1^l) \leq 0 \leq f(t, r_1^R),$$

ja koska $Z_0 = 0$, saadaan Lauseen 5(a.i) nojalla, että Z_0 kuuluu pysäytysalueeseen. Siis optimaalinen pysäytyshetki on $\tau^* = 0$. Tämä todistaa kohdan 1. valinnalla $\hat{r} = r_1^l$.

2. Yhtälö (4.3) sievenee muotoon

$$r^R + r^l = 1, \quad l(r^R) + l\left(\frac{1+\varepsilon}{2}\right) = \frac{4\alpha^3}{c\sigma^2},$$

ja lisäksi

$$\bar{z} = \bar{f}(t, r^R) = -f(t, r^l) = \frac{\sigma^2}{2\alpha} \left[\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) + \log\left(\frac{r^R}{1-r^R}\right) \right].$$

Lauseen 5(a.ii) nojalla optimaalinen pysäytyshetki on

$$\tau^* = \inf\{t \geq 0 : |Z_t| \geq \bar{z}\}.$$

Määritellään vielä

$$\bar{\bar{z}} = \frac{\sigma^2}{2\alpha} \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) < \bar{z}.$$

Yhtälöistä (3.9) ja (3.3) seuraa, että millä tahansa ajanhetkellä t (ei välttämättä optimaalisella pysäytysajalla) päätös "valitse ambiguuteettiurna" on parempi kuin "valitse riskinen urna" täsmälleen silloin, kun $|Z_t| \geq \bar{\bar{z}}$. Tästä seuraa erityisesti, että pysäytyshetkellä $\tau^* > 0$ pätee

$$|Z_{\tau^*}| = \bar{z} \geq \bar{\bar{z}},$$

joten ambiguuteettiurnan valitseminen on optimaalista pysäytyksessä.

Näytetään lopuksi, että \bar{z} kasvaa ε :n funktiona. Koska $\bar{z} = \bar{f}(t, r^R)$ ja $r^R = r^R(\varepsilon)$ määräytyy edellisistä yhtälöistä, riittää tarkastella derivaatan etumerkkiä $\frac{d\bar{z}}{d\varepsilon}$. Derivoidaan identiteetti

$$l(r^R) + l\left(\frac{1+\varepsilon}{2}\right) = \frac{4\alpha^3}{c\sigma^2}$$

ε :n suhteen ja käytetään kaavaa

$$l'(r) = \frac{1}{r^2(1-r)^2} > 0, \quad r \in (0, 1).$$

Saadaan

$$l'(r^R) \frac{dr^R}{d\varepsilon} + \frac{1}{2} l'\left(\frac{1+\varepsilon}{2}\right) = 0, \quad \text{eli} \quad \frac{dr^R}{d\varepsilon} = -\frac{1}{2} \frac{l'\left(\frac{1+\varepsilon}{2}\right)}{l'(r^R)} < 0.$$

Toisaalta

$$\bar{z}(\varepsilon) = \frac{\sigma^2}{2\alpha} \left[\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) + \log\left(\frac{r^R}{1-r^R}\right) \right],$$

joten

$$\frac{d\bar{z}}{d\varepsilon} = \frac{\sigma^2}{2\alpha} \left[\frac{2}{1-\varepsilon^2} + \left(\frac{1}{r^R} + \frac{1}{1-r^R} \right) \frac{dr^R}{d\varepsilon} \right].$$

Koska $r^R > \bar{\pi} = \frac{1+\varepsilon}{2}$ (tästä syystä $r^R(1-r^R) < \bar{\pi}(1-\bar{\pi}) = \frac{1-\varepsilon^2}{4}$), saadaan

$$\left(\frac{1}{r^R} + \frac{1}{1-r^R} \right) \left| \frac{dr^R}{d\varepsilon} \right| = \frac{1}{r^R(1-r^R)} \cdot \frac{1}{2} \frac{l'\left(\frac{1+\varepsilon}{2}\right)}{l'(r^R)} < \frac{1}{\bar{\pi}(1-\bar{\pi})} \cdot \frac{1}{2} = \frac{2}{1-\varepsilon^2}.$$

Tästä seuraa, että edellä esiintyvä lauseke

$$\frac{2}{1-\varepsilon^2} + \left(\frac{1}{r^R} + \frac{1}{1-r^R} \right) \frac{dr^R}{d\varepsilon}$$

on positiivinen ja siis $\frac{d\bar{z}}{d\varepsilon} > 0$. Tämä päättää kohdan 2. ja koko lauseen todistuksen. \square

5.4 Lauseen 4 todistus

Lauseen 5(b) nojalla riittää osoittaa väite (3.22) olettaen, että $a = b$. Palkkiorakenteen symmetrian perusteella yhtälö (4.4) sievenee muotoon (4.3). Lemman 2 nojalla seuraa tällöin, että $r^l = 1 - r^R$ ja että r^R ratkaisee yhtälön

$$l(r^R) = l(\pi) + \frac{b}{\hat{c}}.$$

Bayesilaisessa tapauksessa pätee $\pi = \frac{a}{a+b}$, ja yhtälöstä (3.20) saadaan, että $\hat{r}_B^l = 1 - \hat{r}_B^R$ ja että \hat{r}_B^R ratkaisee yhtälön

$$l(\hat{r}_B^R) = \frac{a+b}{2\hat{c}} = \frac{b}{\hat{c}}.$$

Tästä seuraa, että $\hat{r}^R < \hat{r}_B^R$. \square

5.5 Nash-tasapainotulkinta Ellsbergin mallissa

Tuodaan viimeisenä kohtana tässä tutkielmassa esille Nash-tasapainotulkinta Ellsbergin mallille. Kyseessä on tässä tutkielmassa esitetty tulkinta, joka perustuu Epstein ja Ji:n malliin mutta ei esiinny heidän analyysissään eksplisiittisesti. Tarkastellaan nollasummapeleä, jossa päätöksentekijä valitsee pysäytysajan $\tau \in \Gamma$ ja “luonto” valitsee mitan $Q \in \mathcal{P}_0$. Näiden strategioiden generoima arvo on

$$J(\tau, Q) = \mathbb{E}^Q[X_\tau - c\tau].$$

Pari (τ^*, Q^*) on satulapiste (eli Nash-tasapaino[17] nollasummapelissä), jos

$$J(\tau, Q^*) \leq J(\tau^*, Q^*) \leq J(\tau^*, Q) \quad \forall \tau \in \Gamma, \forall Q \in \mathcal{P}_0.$$

Koska pysäytysaikojen joukko Γ ei yleensä ole konvekksi, laajennetaan strategiajoukkoa satunnaistettuihin pysäytysstrategioihin. Olkoon $U \sim \text{Unif}[0, 1]$ satunnaismuuttuja, joka on riippumaton prosessin (Z_t) generoimasta filtraatiosta. Tämä ulkoinen satunnaistus mahdollistaa strategioiden satunnaistamisen ilman että signaaliprosessin dynamiikka muuttuu.

Jos $\tau_1, \tau_2 \in \Gamma$ ja $\lambda \in [0, 1]$, määritellään satunnaistettu pysäytysstrategia

$$\tau^\lambda = \tau_1 \mathbf{1}_{\{U \leq \lambda\}} + \tau_2 \mathbf{1}_{\{U > \lambda\}}.$$

Koska tapahtuma $\{U \leq \lambda\}$ on riippumaton prosessin (Z_t) generoimasta filtraatiosta ja τ_1, τ_2 ovat \mathcal{G}_t -pysäytysaikoja, on myös τ^λ pysäytysaika. Merkitään Γ^r :llä kaikkien tällaisten ulkoisen satunnaistuksen avulla määriteltyjen pysäytysstrategioiden joukkoa. Tällöin $\Gamma \subset \Gamma^r$.

Määritellään tuotto satunnaistetulle strategialle kaavalla

$$J(\tau, Q) = \mathbb{E}^Q[X_\tau - c\tau].$$

Tällöin

$$X_{\tau^\lambda} - c\tau^\lambda = (X_{\tau_1} - c\tau_1) \mathbf{1}_{\{U \leq \lambda\}} + (X_{\tau_2} - c\tau_2) \mathbf{1}_{\{U > \lambda\}}.$$

Koska U on riippumaton prosessin (Z_t) generoimasta filtraatiosta, saadaan

$$J(\tau^\lambda, Q) = \lambda J(\tau_1, Q) + (1 - \lambda) J(\tau_2, Q).$$

Siis Γ^r on konvekksi satunnaistamisen suhteen ja kuvaus

$$\tau \longmapsto J(\tau, Q)$$

on affiini joukossa Γ^r jokaisella $Q \in \mathcal{P}_0$

Lause 6 (Nash-tasapaino Ellsbergin mallissa). *Oletetaan Ellsbergin spesifikaatio luvusta 3.2. Oletetaan lisäksi, että*

1. \mathcal{P}_0 voidaan parametrisoida välin $m \in [m_0, \bar{m}_0]$ avulla muodossa

$$Q = Q^m, \quad \mathcal{P}_0 = \{Q^m : m \in [m_0, \bar{m}_0]\},$$

ja väli $[m_0, \bar{m}_0]$ on kompakti ja konvekksi.

2. jokaiselle $\tau \in \Gamma^r$ satunnaismuuttuja X_τ on \mathcal{G}_τ -mitallinen ja

$$\sup_{Q \in \mathcal{P}_0} \mathbb{E}^Q |X_\tau| < \infty.$$

3. jokaisella $\tau \in \Gamma^r$ kuvaus

$$m \mapsto J(\tau, Q^m)$$

on jatkuva ja affiini.

4. satunnaistettujen pysäytysstrategioiden joukko Γ^r on varustettu sellaisella topologialla, että Γ^r on kompakti ja konvekksi sekä kuvaus

$$(\tau, m) \mapsto J(\tau, Q^m)$$

on affiini kummankin muuttujan suhteen. Lisäksi kuvaus on jatkuva muuttujassa m , ja muuttujassa τ se määräytyy odotusarvona

$$J(\tau, Q) = \mathbb{E}^Q[X_\tau - c\tau],$$

minkä vuoksi se ei voi kasvaa rajatta strategian pienillä muutoksilla.

Tällöin on olemassa satulapiste

$$(\hat{\tau}, \hat{Q}) \in \Gamma^r \times \mathcal{P}_0.$$

Lisäksi \hat{Q} voidaan valita välin päätepisteestä, eli

$$\hat{Q} \in \{Q^{m_0}, Q^{\bar{m}_0}\}.$$

Jos lisäksi luvussa 3.2 johdettu osuma-aika

$$\tau^* = \inf\{t \geq 0 : |Z_t| \geq \bar{z}\} \quad (\text{oppimisen ollessa optimaalista, muutoin } \tau^* = 0)$$

saavuttaa saman maximin-arvon kuin satunnaistettu peli, eli

$$\inf_{Q \in \mathcal{P}_0} J(\tau^*, Q) = \sup_{\tau \in \Gamma^r} \inf_{Q \in \mathcal{P}_0} J(\tau, Q),$$

niin pari (τ^*, \hat{Q}) on satulapiste myös alkuperäisessä pelissä $\Gamma \times \mathcal{P}_0$.

Todistus. Merkitään

$$v = \sup_{\tau \in \Gamma^r} \inf_{Q \in \mathcal{P}_0} J(\tau, Q).$$

Koska $\mathcal{P}_0 = \{Q^m : m \in [m_0, \bar{m}_0]\}$, voidaan kirjoittaa

$$v = \sup_{\tau \in \Gamma^r} \inf_{m \in [m_0, \bar{m}_0]} J(\tau, Q^m).$$

Jokaisella $\tau \in \Gamma^r$ kuvaus $m \mapsto J(\tau, Q^m)$ on oletuksen mukaan affiini välillä $[m_0, \bar{m}_0]$. Siksi sen infimum saavutetaan välin päätepisteessä:

$$\inf_{m \in [m_0, \bar{m}_0]} J(\tau, Q^m) = \min\{J(\tau, Q^{m_0}), J(\tau, Q^{\bar{m}_0})\}.$$

Oletusten mukaan Γ^r on kompakti ja konvekksi, väli $[m_0, \bar{m}_0]$ on kompakti ja konvekksi, ja funktio

$$(\tau, m) \mapsto J(\tau, Q^m)$$

on affiini kummankin muuttujan suhteen. Lisäksi kuvaus on jatkuva muuttujassa m , ja muuttujassa τ sen arvo määräytyy odotusarvon $\mathbb{E}^Q[X_\tau - c\tau]$ kautta.

Näin Sionin minimax-lauseen[22] ehdot täyttyvät, ja saadaan

$$\sup_{\tau \in \Gamma^r} \inf_{m \in [m_0, \bar{m}_0]} J(\tau, Q^m) = \inf_{m \in [m_0, \bar{m}_0]} \sup_{\tau \in \Gamma^r} J(\tau, Q^m).$$

Tästä seuraa satulapisteen olemassaolo: on olemassa pari

$$(\hat{\tau}, \hat{m}) \in \Gamma^r \times [m_0, \bar{m}_0]$$

siten, että merkitsemällä $\hat{Q} = Q^{\hat{m}}$ saadaan

$$J(\tau, \hat{Q}) \leq J(\hat{\tau}, \hat{Q}) \leq J(\hat{\tau}, Q) \quad \forall \tau \in \Gamma^r, \forall Q \in \mathcal{P}_0.$$

Koska kiinteällä $\hat{\tau}$:lla minimi muuttujassa m saavutetaan päätepisteessä, voidaan lisäksi valita

$$\hat{Q} \in \{Q^{m_0}, Q^{\bar{m}_0}\}.$$

Oletetaan lisäksi, että luvussa 3.2 johdettu osuma-aika τ^* saavuttaa saman maximin-arvon kuin satunnaistettu peli:

$$\inf_{Q \in \mathcal{P}_0} J(\tau^*, Q) = v.$$

Tällöin

$$J(\tau^*, Q) \geq v \quad \forall Q \in \mathcal{P}_0.$$

Soveltamalla tätä mittaan \hat{Q} saadaan

$$J(\tau^*, \hat{Q}) \geq v.$$

Toisaalta, koska $\tau^* \in \Gamma \subset \Gamma^r$ ja $(\hat{\tau}, \hat{Q})$ on satulapiste joukossa $\Gamma^r \times \mathcal{P}_0$, pätee

$$J(\tau^*, \hat{Q}) \leq J(\hat{\tau}, \hat{Q}) = v.$$

Siis

$$J(\tau^*, \hat{Q}) = v.$$

Nyt kaikilla $\tau \in \Gamma$ pätee

$$J(\tau, \hat{Q}) \leq v = J(\tau^*, \hat{Q}),$$

ja kaikilla $Q \in \mathcal{P}_0$ pätee

$$J(\tau^*, Q) \geq v = J(\tau^*, \hat{Q}).$$

Siis

$$J(\tau, \hat{Q}) \leq J(\tau^*, \hat{Q}) \leq J(\tau^*, Q) \quad \forall \tau \in \Gamma, \forall Q \in \mathcal{P}_0.$$

Näin (τ^*, \hat{Q}) on satulapiste myös alkuperäisessä pelissä $\Gamma \times \mathcal{P}_0$. □

Huomio. Yllä oleva todistus jakautuu kahteen osaan. Ensiksi satulapisteen olemassaolo osoitetaan satunnaistettujen pysäytysstrategioiden joukossa Γ^r minimax-argumentilla. Toiseksi satulapiste palautetaan puhtaaseen pysäytysaikaan Ellsbergin mallin erityisrakenteen avulla: pahimman tapauksen mitta voidaan valita välin päätepisteestä, ja luvussa 3.2 johdettu osuma-aika saavuttaa saman maximin-arvon kuin satunnaistettu peli.

Vastaava palautus puhtaisiin strategioihin yleisemmässä Chen–Epstein-kehyksessä olisi huomattavasti teknisempi, eikä sitä tämän tutkielman puitteissa tarkastella.

Viitteet

- [1] Abate A, Redig F, Tkachev I (2013) On the effect of perturbation of conditional probabilities in total variation. *Statistics and Probability Letters* 88:1–8.
- [2] Arrow KJ, Blackwell D, Girshick MA (1949) Bayes and minimax solutions of sequential decision problems. *Econometrica* 17(3–4):213–244.
- [3] Ash RB (1972) *Real Analysis and Probability*. Academic Press, New York.
- [4] Borodin A, Salminen P (2015) *Handbook of Brownian Motion Facts and Formulae*. 2nd ed. Birkhäuser, Basel.
- [5] Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* 75(4):643–669.
- [6] El Karoui N, Kapoudjian C, Pardoux E, Peng S, Quenez MC (1997) Reflected solutions of backward SDEs and related obstacle problems for PDEs. *Annals of Probability* 25(2):702–737.
- [7] Epstein LG, Chen Z (2002) Ambiguity, risk and asset returns in continuous time. *Econometrica* 70(4):1403–1443.
- [8] Epstein LG, Schneider M (2003) Recursive multiple-priors. *Journal of Economic Theory* 113(1):1–31.
- [9] Epstein LG, Schneider M (2007) Learning under ambiguity. *Review of Economic Studies* 74(4):1275–1303.
- [10] Epstein LG, Schneider M (2008) Ambiguity, information quality and asset pricing. *Journal of Finance* 63(1):197–228.
- [11] Epstein LG, Schneider M (2010) Ambiguity and asset markets. *Annual Review of Financial Economics* 2:315–346.
- [12] Epstein LG, Ji S (2019) Optimal learning under robustness and time-consistency. *Operations Research* 70(3):1317–1329.
- [13] Gilboa I, Schmeidler D (1989) Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18(2):141–153.

- [14] Gilboa I (2009) *Theory of Decision under Uncertainty*. Cambridge University Press, New York.
- [15] Gilboa I (2015) Rationality and the Bayesian paradigm. *Journal of Economic Methodology* 22(3):312–334.
- [16] Knight FH (1921) *Risk, uncertainty and profit*. Houghton Mifflin, Boston.
- [17] Nash J (1951) Non-cooperative games. *Annals of Mathematics* 54(2)
- [18] Liptser R, Shiryaev A (2001) *Statistics of Random Processes I: General Theory*. Springer, Berlin.
- [19] Marinacci M (2002) Learning from ambiguous urns. *Statistical Papers* 43(1):143–151.
- [20] Peskir G, Shiryaev A (2006) *Optimal Stopping and Free-Boundary Problems*. Springer, Berlin.
- [21] Shiryaev A (2008) *Optimal Stopping Rules*. 2nd ed. Springer, Berlin.
- [22] Sion M (1958) On general minimax theorems. *Pacific Journal of Mathematics* 8(1):171–176.
- [23] Øksendal B (2005) *Stochastic Differential Equations*. 5th ed. Springer, Berlin.
- [24] Zhang J (2017) *Backward Stochastic Differential Equations*. Springer, New York.