



Aalto-yliopisto
Perustieteiden
korkeakoulu

Ylikerroinstrategiat ja Poisson- mallit vedonlyönnissä (aihe-esittely)

Jussi Kolehmainen

23.01.2012

Ohjaaja: *Jussi Kangaspunta*

Valvoja: *Ahti Salo*

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Kertoimet ja todennäköisyydet

- ✦ Vedonlyönnin kohteena ottelun voittaja
 - 1 = kotivoitto, X = tasapeli, 2 = vierasvoitto
- ✦ Vedonlyöntitoimisto myy merkeille kertoimia q_1 , q_X ja q_2

Wigan Athletic - Manchester City

8.00

4.25

1.40

- Määritys: asiantuntijat, mallit, pelikäyttäytyminen jne.
- Kertoimista on vähennetty toimiston komissio
- ✦ Palautusprosentti on osuus, joka palautuu odotusarvoisesti pelaajille voittoina
 - Pieni komissio Suuri palautusprosentti

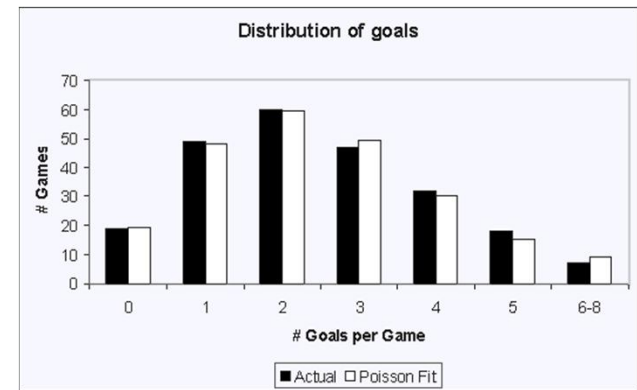
Ylikertoimet

- ✦ Toimiston kertoimet eivät välttämättä kuvaa oikeita todennäköisyyksiä
 - Virheet joukkueiden voimasuhteiden määrittämisessä
 - Asiakkaiden pelikäyttäytymisen vaikutus
- ✦ Pelaaja voi määrittää omat todennäköisyysarviot p_1 , p_X ja p_2
- ✦ Ylikerroin: $q > 1/p$
 - Kerroin liian hyvä omaan arvioon nähden
 - Odotusarvoisesti yli 100% palautus, mikäli oma arvio on oikea
- ✦ Ongelma: Miten laskea tarkat todennäköisyysarviot?

Poisson-jakauma maalimäärille

- ✧ Moroney, 1956: Joukkueiden maalimääriä voi mallintaa riippumattomilla Poisson-jakaumilla

$$X_1 \sim \text{Poisson}(\lambda_1) \quad X_2 \sim \text{Poisson}(\lambda_2)$$



- Intensiteetit maaliodotusarvoja edellisten pelien perusteella
- ” Summaamalla kotivoitto-, tasapeli- ja vierasvoittotulosten todennäköisyydet saadaan 1X2-jakauma

Muunnelmia 1

- ✧ Maher, 1982: Joukkueille hyökkäys- ja puolustusparametrit edellisistä peleistä
 - Maali-intensiteetit lasketaan näistä parametreista
 - Mukana myös parametri λ kotikenttäedulle+
 - Parametrit lasketaan Poisson-regressiolla
- ✧ Ongelma: Joukueiden maalimäärät eivät ole riippumattomia (esim. Lee, 1997)
 - Havaitaan erityisesti vähämaalisisissa peleissä
- ✧ Dixon & Coles, 1997: Korjaustermi 0-0, 0-1, 1-0 ja 1-1 -tuloksille

Muunnelmia 2

- ✧ Karlis & Ntzoufras, 2003: Kaksimuuttujainen Poisson-jakauma $BP(\lambda_1, \lambda_2, \lambda_3)$
 - Kovarianssi λ_3 kuvaa maalimäärien riippuvuutta
 - Reunajakaumat tavallisia Poisson-jakaumia
 - Maalimäärien erotus riippumaton kovarianssista => 1X2

- ✧ Parametrien estimointi vaikeampaa
 1. Momenttimenetelmä ja SU-menetelmä (Kocherlakota, 1992)
 2. Bayesilainen estimointi (Tsionas, 1999)

Muunnelmia 3

- ✧ Karlis & Ntzoufras, 2001: Painotusermi $pD(x, \cdot)$ tasapeleille
 - $D(x, \cdot)$ Poisson-, geometrinen tai Bernoulli-jakauma
 - Huomioidaan vain 5 maalia/joukkue/peli, millä eliminoidaan uusien maalien riippuvuutta aikaisemmista (Poisson-oletus)
- ✧ Rue & Salvesen, 1997: Joukkueiden hyökkäys- ja puolustusparametrit muuttuvat ajassa
 - Satunnaisuus Brownin liikkeestä
 - Päivitys MCMC-algoritmillä

Empiirinen osio: data



- ✧ Testidataa osoitteesta Football-Data.co.uk CSV-muodossa
 - Valitaan Englannin Valioliiga ja Mestaruussarja
 - Kaikki tulokset kaudesta 93-94 alkaen
 - Saatavilla suurimpien brittiläisten toimistojen kertoimet vuodesta 2000 alkaen
- ✧ Ottelutulokset ja kertoimet parsitaan CSV-tiedostoista
 - Tallennetaan MySQL-tietokantaan helpompaa käsittelyä varten
- ✧ Mallit toteutetaan Java-kielellä
 - É Nopea kehitystyö, tehokkuus, helpot rajapinnat

Empiirinen osio: strategia

- ✦ Strategiana on löytää Poisson-mallien avulla ylikertoimia ja panostaa vain niihin
 - Vain singlejä eli ei tehdä yhdistelmävetoja
 - Saman kohteen mahdollisesta kahdesta ylikertoimesta panostetaan sitä, jolla on suurempi odotusarvo (kerroin*todennäköisyysarvio)
- ✦ Kelly, 1956: Optimaalinen panostusosuus kassasta:

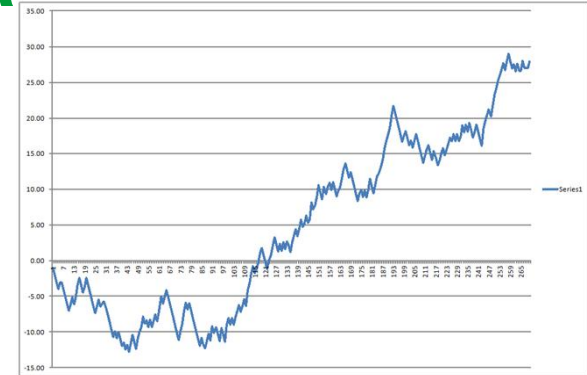
$$B = \frac{pq - 1}{q - 1}$$

missä p on todennäköisyysarvio ja q on kerroin

- Pelikassan maksimaalinen kasvunopeus

Empiirinen osio: tavoitteita

- ⤴ Kuvaajat pelikassan kehityksestä
 - Voitollisuus pitkällä aikavälillä?
 - Mikä on pitkä aikaväli?
- ⤴ Poikkeavuudet mallin eri muunnelmien väleillä
 - Ylikertoimien määrä?
 - Tiettyjen tulosten (esim. tasapelien) ennustettavuus?
- ⤴ Yhteensopivuustestit maaliaineistolle ja jakaumalle
- ⤴ Panostusstrategia
 - Kelly vs tasapanos



Kirjallisuutta

- ⋄ Kelly J. L. Jr., 1956: A new interpretation of information rate
- ⋄ Dixon M. & Coles S., 1995: Modelling association football scores and inefficiencies in the football betting market
- ⋄ Maher M. J., 1997: Modelling association football scores
- ⋄ Lee A. J., 1997: Modelling scores in the Premier League: Is Manchester United really the best?
- ⋄ Rue H. & Salvesen O., 2000: Prediction and retrospective analysis of soccer matches in a league
- ⋄ Marttinen N., 2001: Creating a profitable betting strategy for football by using statistical modelling
- ⋄ Karlis D. & Ntzoufras I., 2003: Analysis of sports data by using bivariate Poisson models

Aikataulu

- ✦ Kevät 2012
 - Malleihin tutustuminen
 - Mallien toteutus
 - Ensimmäiset testit
- ✦ Kesä 2012
 - Testejä aineistolle
 - Työn kirjoitus
- ✦ Syksy 2012
 - Työn viimeistely
 - Valmiin työn esittely