



Stockholms
universitet

3:e perioden
- en statistisk analys av utfallet i jämna
ishockeymatcher

Niklas Warren

Kandidatuppsats 2016:29
Matematisk statistik
December 2016

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

3:e perioden

- en statistisk analys av utfallet i jämna ishockeymatcher

Niklas Warren*

December 2016

Sammanfattning

I denna uppsats analyseras utfallet av jämna ishockeymatcher från SHL. Medelst multinomial logistisk regression anpassas modeller vilka syftar till att förklara sannolikheterna för de olika utfallen. Vidare undersöks modellerna i prediktionssammanhang innan en primär modell väljs ut. Prediktionsförmågan visar sig undermålig och bland annat predikterar alla modeller oavgjort resultat i alldeles för hög grad. Genom att istället prediktera utifrån kriterier som begränsar antalet oavgjorda resultat förbättras prediktionen något. Lite överraskande visar analysen att ordningen i vilken lagen gör målen i matchen kan spela en betydande roll, men inte på det sätt man kanske förväntar sig. Ett lag som hämtar upp ett större underläge står sällan som segrare efter tre spelade perioder.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: niklaswarren@live.se. Handledare: Jan-Olov Persson.

Abstract

In this paper the outcome of tight ice hockey games from SHL are analyzed. By means of multinomial logistic regression, models which seek to explain the probabilities of the different outcomes are fitted. Further, prediction results for these models on new observations are examined before a primary model is selected. The predictive capability proves inadequate, including all models predicting way too high proportion of tied games. If prediction instead is based on criteria that limit the number of tied games the results are slightly improved. Somewhat surprising, it turns out that the order in which the teams score the goals in the game can play a significant role, but not in the way one might expect. A team that comes back from a large deficit are rarely victorious after three periods of play.

Förord

Detta är en kandidatuppsats i matematisk statistik hösten 2016. Jag vill passa på att tacka min handledare Jan-Olov Persson för vägledning under arbetets gång. Ett tack går även till Svenska Ishockeyförbundet som tillåtit användandet och bearbetandet av data.

Innehåll

| | | |
|----------|--|-----------|
| 1 | Introduktion | 5 |
| 2 | Teori | 6 |
| 2.1 | Logistisk och multinomial logistisk regression | 6 |
| 2.2 | Purposeful selection | 7 |
| 2.3 | Test som används i analysen | 8 |
| 2.3.1 | Likelihood Ratio-test | 8 |
| 2.3.2 | Wald-test | 9 |
| 2.3.3 | Cramérs V | 9 |
| 2.3.4 | Fishers exakta test | 9 |
| 2.3.5 | AIC | 9 |
| 2.3.6 | Hosmer-Lemeshows test | 10 |
| 2.3.7 | ROC | 10 |
| 2.3.8 | Undersökning av enskilda observationer | 10 |
| 3 | Beskrivning av data | 12 |
| 3.1 | Allmänt | 12 |
| 3.2 | Variabler | 12 |
| 4 | Analys av data | 18 |
| 4.1 | Statistisk modellering | 19 |
| 4.1.1 | Steg 1 | 19 |
| 4.1.2 | Steg 2 | 19 |
| 4.1.3 | Steg 3 | 20 |
| 4.1.4 | Steg 4 | 21 |
| 4.1.5 | Steg 5 | 21 |
| 4.1.6 | Steg 6 | 24 |
| 4.1.7 | Steg 7 | 27 |
| 4.2 | Prediktion | 32 |
| 5 | Resultat | 35 |
| 6 | Diskussion | 37 |
| A | Appendix | 39 |
| A.1 | Samband mellan variabler | 39 |
| A.2 | Tabeller över enskilda kategorivariabler | 40 |
| A.3 | Statistisk modellering | 42 |
| A.3.1 | Steg 1 | 42 |
| A.3.2 | Steg 2-6 | 43 |
| A.3.3 | Steg 7 - Hosmer-Lemeshows test | 44 |
| A.3.4 | Steg 7 - Klassificeringstabeller | 48 |
| A.4 | Prediktion | 49 |

| | | |
|-------|---|----|
| A.4.1 | Resultat då prediktion sker på utfallet som ges högst sannolikhet | 49 |
| A.4.2 | Resultat då prediktion sker enligt andra kriterier . . . | 51 |

1 Introduktion

Ishockey, likväl som idrott i stort, är något som intresserar många människor. Varför det är så finns det säkert lika många svar på som det finns idrottsintresserade. Vi kommer heller inte undersöka den saken närmare, utan nöjer oss med antagandet att den spänning som följer av idrottens oförutsägbara natur förmodligen är en viktig ingrediens. I denna uppsats gör vi ett försök till att ta död på lite av denna spänning genom att titta närmare på vad som kan avgöra utfallet i jämna ishockeymatcher; inte som ett sabotage riktat mot ishockeyintresserade utan som ett sätt att bidra med ökad förståelse inom en fascinerande idrott.

Det finns självfallet fler än ett sätt att definiera vad en jämn ishockeymatch är, men för genomförandet av denna undersökning kommer vi att titta på matcher där resultatet efter två perioder är oavgjort och se vilket resultat som följer efter tre perioder. Matcherna som är aktuella kommer från SHL (tidigare Elitserien i ishockey) och omfattar tio säsonger, från säsongen 2006/2007 till säsongen 2015/2016. Under alla säsonger spelas en grundserie med 55 omgångar mellan 12 lag, bortsett från säsongen 2015/2016 där grundserien består av 52 omgångar och 14 lag. Upplägget skiljer sig åt mellan en del av säsongerna, men generellt gäller för ishockeymatcher att full tid är resultatet efter tre perioder, där varje period pågår i 20 minuters effektiv speltid. Vid oavgjort resultat efter tre spelade perioder har sedan olika förfaranden tillämpats; en kort förlängning på 5 minuter har följt, ibland med färre spelare på isen, och från och med säsongen 2010/2011 har ett efterföljande straffavgörande tillkommit för att utse en vinnare om inget lag har avgjort under förlängningen. Att denna undersökning görs utifrån resultatet vid full tid är ett sätt att kringgå denna inkonsekvens av hur matcher avgjorts över dessa tio säsonger.

Målet med uppsatsen är att undersöka faktorer som kan vara av betydelse för utgången i jämna ishockeymatcher. Detta innefattar mer specifikt att medelst en statistisk modell dels försöka förklara vad som påverkar sannolikheterna för de olika utfallen och dels se hur denna modell lyckas prediktera utfall. Matcher från de nio äldsta säsongerna kommer användas till att ta fram en förklarande modell och matcher från den senaste säsongen kommer användas för validering av modellen.

2 Teori

Materialet som presenteras i denna del har inte för avsikt att ge en heltäckande teoretisk bakgrund utan syftar endast till att kortfattat beskriva för analysen fundamentala områden. Teorin är hämtad från [4] om inte annat anges.

2.1 Logistisk och multinomial logistisk regression

När en responsvariabel Y är kategorisk och kan anta två olika värden, $Y = 0$ respektive $Y = 1$, kan logistisk regression användas för att modellera den betingade sannolikheten för utfallen, givet en uppsättning värden på ett antal förklarande variabler. Detta kan formuleras som $\pi(\mathbf{x}) = P(Y = 1|\mathbf{x}) = (e^{\beta'\mathbf{x}})/(1 + e^{\beta'\mathbf{x}})$, där $\mathbf{x}' = (1, x_1, \dots, x_p)$ är en vektor bestående av en konstantterm och p förklaringsvariabler och $\beta' = (\beta_0, \beta_1, \dots, \beta_p)$ är en vektor bestående av $p + 1$ parametrar.

Oddsset mellan de båda utfallen, givet en viss uppsättning värden på de förklarande variablerna, definieras som $\pi(\mathbf{x})/(1 - \pi(\mathbf{x}))$ och kan anta alla värden större än 0. Ett odds på 1 innebär att inget utfall är mer sannolikt än det andra. Ett ofta använt mått är odds-kvoten (OR), vilken ger en uppskattning i termer av odds på hur mycket mer eller mindre troligt ett utfall är då värdet på en förklaringsvariabel ändras. Med $\mathbf{x}'_{\text{alt}} = (1, x_1, \dots, x_j + 1, \dots, x_p)$ och $\mathbf{x}' = (1, x_1, \dots, x_j, \dots, x_p)$ kan detta formuleras som

$$\mathbf{OR} = \frac{\pi(\mathbf{x}_{\text{alt}})/(1 - \pi(\mathbf{x}_{\text{alt}}))}{\pi(\mathbf{x})/(1 - \pi(\mathbf{x}))}.$$

En logit-transformation innebär en logaritmering av oddsset och är central för logistisk regression. Det gäller nämligen att logit-funktionen $g(\mathbf{x}) = \log(\pi(\mathbf{x})/(1 - \pi(\mathbf{x}))) = \beta'\mathbf{x}$ är linjär i parametrarna och kan anta alla reella värden samtidigt som de betingade sannolikheterna fortfarande antar värden mellan 0 och 1. Det är denna funktion som definierar den logistiska regressionsmodellen.

I detta arbete är den berörda responsvariabeln Y kategorisk och kan anta tre olika värden utan inbördes ordning, $Y = 0, Y = 1$ och $Y = 2$. En generalisering av logistisk regression som kallas multinomial logistisk regression kan då användas. Med $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$ för $j = 0, 1, 2$ och $\beta'_k = (\beta_{k0}, \beta_{k1}, \dots, \beta_{kp})$ för $k = 1, 2$ ges de betingade sannolikheterna för de olika utfallen, givet en uppsättning värden på förklaringsvariablerna, av

$$\pi_0(\mathbf{x}) = \frac{1}{1 + e^{\beta'_1\mathbf{x}} + e^{\beta'_2\mathbf{x}}},$$
$$\pi_1(\mathbf{x}) = \frac{e^{\beta'_1\mathbf{x}}}{1 + e^{\beta'_1\mathbf{x}} + e^{\beta'_2\mathbf{x}}},$$

$$\pi_2(\mathbf{x}) = \frac{e^{\beta_2' \mathbf{x}}}{1 + e^{\beta_1' \mathbf{x}} + e^{\beta_2' \mathbf{x}}}.$$

Genom att välja en av kategorierna, förslagsvis $Y = 0$, som referenskategori vilken övriga kategorier jämförs mot kan två logit-funktioner skapas enligt

$$g_1(\mathbf{x}) = \log\left(\frac{\pi_1(\mathbf{x})}{\pi_0(\mathbf{x})}\right) = \beta_1' \mathbf{x},$$

$$g_2(\mathbf{x}) = \log\left(\frac{\pi_2(\mathbf{x})}{\pi_0(\mathbf{x})}\right) = \beta_2' \mathbf{x}.$$

Antag att n oberoende försök genomförs, där varje enskilt försök kan hamna i exakt en av tre kategorier, 0, 1 respektive 2, och där sannolikheten för att hamna i respektive kategori, π_0, π_1 respektive π_2 , är given. Sannolikheten för ett visst utfall av de n försöken ges då av multinomialfördelningen. Med n_0 försök i kategori 0, n_1 i kategori 1 och n_2 i kategori 2, där $\sum_{i=0}^2 n_i = n$, ges sannolikhetsfunktionen av

$$\frac{n!}{n_0! n_1! n_2!} \pi_0^{n_0} \pi_1^{n_1} \pi_2^{n_2} \propto \pi_0^{n_0} \pi_1^{n_1} \pi_2^{n_2}.$$

För logit-funktionerna skattas $\beta = (\beta_1, \beta_2)$ med maximum likelihood-metoden. Givet n oberoende observationer $(y_i, x_{1i}, \dots, x_{pi})$ för $i = 1, \dots, n$ så ges likelihood-funktionen av

$$L(\beta) = \prod_{i=1}^n \pi_0(\mathbf{x}_i)^{y_{0i}} \pi_1(\mathbf{x}_i)^{y_{1i}} \pi_2(\mathbf{x}_i)^{y_{2i}},$$

där $y_{ji} = 1$ då $y_i = j$ och $y_{ji} = 0$ då $y_i \neq j$ för $j = 0, 1, 2$. Efter partiell derivering av $\log(L(\beta))$ med avseende på varje enskild parameter fås likelihood-ekvationerna genom att sätt derivatorna lika med noll. Parameterskattningarna erhålls sedan genom att med iterativa metoder lösa ekvationerna för $\hat{\beta}$.

2.2 Purposeful selection

I situationer där många potentiella förklaringsvariabler kan ingå i en multipel regressionsmodell behövs metoder för att hantera dessa. En sådan är sju-stepsproceduren purposeful selection vilken beskrivs nedan.

I steg 1 anpassas en enkel regressionsmodell för varje enskild förklaringsvariabel. Hosmer et al. (2013) föreslår här att varje variabel vars p-värde är under 0.25 vid ett signifikanstest identifieras för att inte missa variabler av betydelse.

I steg 2 anpassas en multipel modell bestående av de förklaringsvariabler som identifierades i steg 1. Utifrån denna modell undersöks sedan betydelsen av varje variabel, och de som inte bidrar på mer traditionella signifikansnivåer

utesluts ur modellen.

I steg 3 så jämförs parameterskattningarna mellan den nya reducerade modellen och den ursprungliga multipla modellen. Syftet är att identifiera om exkluderandet av en variabel har haft betydande påverkan på andra variablers parameterskattningar, och därför bör inkluderas i modellen igen. Hosmer et al. (2013) föreslår här att parameterskattningar som förändrats mer än 20% ska uppmärksammas.

I steg 4 återkommer de förklaringsvariabler som föll bort i steg 1. För var och en av dessa testas nu signifikansen av ett inkluderande i den reducerade multipla modellen från steg 3.

I steg 5 undersöks variablerna i den multipla modellen närmare. För varje kontinuerlig variabel i modellen kontrolleras antagandet att logit-funktionen växer eller avtar linjärt som en funktion av variabeln. För detta ändamål finns flera metoder varav en kallas metoden med designvariabler. Denna går ut på att den kontinuerliga variabeln görs om till en kategorivariabel med fyra nivåer. Datamaterialets gruppering baseras sedan utifrån den kontinuerliga variabelns kvartiler. En ny anpassning av den multipla modellen görs nu med den kontinuerliga variabeln ersatt av kategorivariabeln varpå linjäriteten kontrolleras i ett spridningsdiagram av parameterskattningarna mot mittpunkterna för kvartilnivåerna.

I steg 6 undersöks om det finns intressanta samspel mellan variabler i den multipla modellen. Ett eventuellt inkluderande av en samspelsterm i modellen sker efter att såväl statistiska som praktiska överväganden gjorts.

Innan en modell fastställs sker i steg 7 en kontroll av dess anpassning och lämplighet. Olika test för detta finns beskrivna i avsnitt 2.3.

2.3 Test som används i analysen

2.3.1 Likelihood Ratio-test

I ett Likelihood Ratio-test så jämförs en modell mot en alternativ modell, där den första modellen är ett specialfall av den senare på så sätt att den innehåller färre parametrar. Testet baseras på likelihood-kvoten, $\Lambda = L_0/L_1$, där L_0 är värdet på den maximerade likelihood-funktionen för specialfallet och L_1 värdet på den maximerade likelihood-funktionen för alternativet. Det gäller att test-statistikan $-2 \log(\Lambda)$ är asymptotiskt χ^2 -fördelad under antagandet att specialfallet håller, med frihetsgrader motsvarande differensen i parameterummens dimension mellan de båda modellerna. Detta test hjälper oss således att bedöma om en enklare modell ska förkastas till förmån

för en med ett utökat antal parametrar [1].

2.3.2 Wald-test

Maximum likelihood-skattningar av parametrar i regressionsmodeller följer en asymptotisk normalfördelning. Med test-statistikan $\hat{\beta}_j / \widehat{SE}(\hat{\beta}_j)$, där $\hat{\beta}_j$ är skattningen för parameter j och $\widehat{SE}(\hat{\beta}_j)$ dess skattade medelfel, kan vi utföra Wald-test på enskilda parametrar. Under hypotesen att parametern av intresse är noll gäller att statistikan approximativt följer en standardnormalfördelning och vi kan därför med hjälp av erhållet p-värde bedöma om denna hypotes ska förkastas [1].

2.3.3 Cramérs V

Cramérs V är ett mått på associationen mellan två nominella variabler, A, B , i vilka A består av r kategorier $A_i, i = 1, \dots, r$ och B består av k kategorier $B_j, j = 1, \dots, k$. Måttet bygger på Pearsons χ^2 -statistika vilken ges av $\chi^2 = \sum_{i=1}^r \sum_{j=1}^k ((O_{i,j} - E_{i,j})^2 / E_{i,j})$, där $O_{i,j}$ är antalet observationer av typ (A_i, B_j) och $E_{i,j}$ är det förväntade antalet observationer av typ (A_i, B_j) . Om vi låter n beteckna stickprovsstorleken ges Cramérs V av $\sqrt{\chi^2 / (n \cdot \min(r-1, k-1))}$. Värdet som erhålls ligger i det slutna intervall mellan 0 och 1, där 0 innebär ingen association mellan variablerna och 1 innebär fullständig association [2].

2.3.4 Fishers exakta test

Vid analys av kontingenstabeller kan Fishers exakta test vara användbart. Testet fungerar för alla stickprovsstorlekar och således även för små stickprov där antaganden för Pearsons χ^2 -test inte är uppfyllda. Med utgångspunkt i en 2×2 -kontingenstabell där oberoende observationer klassificeras på två olika sätt testas nollhypotesen att klassificeringarna är oberoende mot alternativhypotesen att det finns en signifikant association mellan de två. Under nollhypotesen är tabellens cellfrekvenser hypergeometriskt fördelade så för en given uppsättning frekvenser kan p-värdet för ett minst lika extremt utfall beräknas exakt genom fördelningens sannolikhetsfunktion. I viss statistisk programvara kan testet generaliseras för kontingenstabeller med godtyckligt antal rader och kolumner [1].

2.3.5 AIC

Värdet av likelihood-funktionen vid maximum likelihood-skattningen beskriver en modells anpassning. Om detta värde kombineras med modellens komplexitet, dimensionen av parametervektorn, så kan olika modeller jämföras enligt bestämda kriterier. Akaike's informationskriterium (AIC) jämför den relativa kvaliteten på olika modeller för ett givet dataset enligt

AIC = $-2l(\hat{\beta}_{ML}) + 2p$, där $l(\hat{\beta}_{ML})$ är den maximerade log-likelihooden och p är antalet parametrar i modellen. Den modell som har lägst AIC-värde förespråkas enligt detta kriterium, vilket innebär att modeller med många parametrar bestraffas [3].

2.3.6 Hosmer-Lemeshows test

Hosmer-Lemeshows test är ett Goodness of fit-test för modeller med binär responsvariabel och kan även användas genom att dela upp en multinomial logistisk regressionsmodell i flera binära modeller vilka testas separat. Data grupperas här i ett bestämt antal grupper g (ofta är $g = 10$) baserat på antingen percentiler eller fixerade värden av de skattade sannolikheterna för utfallen. Statistikan, \hat{C} , bygger på Pearsons χ^2 -statistika och ges av $\hat{C} = \sum_{i=1}^g ((O_{1i} - E_{1i})^2 / E_{1i}) + ((O_{0i} - E_{0i})^2 / E_{0i})$ där O_{1i} respektive O_{0i} är observerat antal av utfallen 1 respektive 0 inom grupp i , och E_{1i} respektive E_{0i} är förväntat antal av utfallen 1 respektive 0 inom grupp i . Under antagandet att modellen är sann följer \hat{C} en asymptotisk $\chi^2(g - 2)$ -fördelning. För mer om detta test, se [4].

2.3.7 ROC

Om responsvariabeln är binär kan en modells prediktionsförmåga uppskattas med hjälp av en ROC-kurva (Receiver Operating Characteristic Curve). Detta kan även utnyttjas genom att dela upp en multinomial modell med trinär responsvariabel i två binära modeller vilka testas separat. Om vi för observation i låter $\hat{\pi}(\mathbf{x}_i)$ beteckna den skattade sannolikheten för utfallet 1 kan vi jämföra $\hat{\pi}(\mathbf{x}_i)$ mot en given skärningspunkt $s \in (0, 1)$. Vi låter sedan det predikerade värdet för observation i vara 1 om $\hat{\pi}(\mathbf{x}_i) > s$ och 0 annars. För varje möjlig skärningspunkt $s \in (0, 1)$ kan vi sedan beräkna sensitivitet, definierat som sannolikheten att prediktionsvärdet är 1 givet att utfallet är 1, och specificitet, definierat som sannolikheten att prediktionsvärdet är 0 givet att utfallet är 0. En ROC-kurva erhålls sedan om sensitivitet plottas mot $1 - \text{specificitet}$, och ett AUC-värde (Area Under Curve) ges av arean under den uppritade kurvan. Generellt innebär en större area också en bättre prediktionsförmåga, men det ska samtidigt påpekas att AUC-värdet inte tar hänsyn till antalet parametrar i modellen. Hosmer et al. (2013) ger följande generella tumregler för att bedöma AUC-värdet: 0.5 så kan vi lika gärna övergå till slantsingling; mellan 0.5 och 0.7 bedöms som dåligt och inte mycket bättre än slantsingling; mellan 0.7 och 0.8 är acceptabelt; mellan 0.8 och 0.9 är riktigt bra; över 0.9 är enastående.

2.3.8 Undersökning av enskilda observationer

En modells anpassning kan beskrivas genom en summeringsstatistika som med ett enda tal summerar överensstämmelsen mellan observerade och an-

passade värden. Inom logistisk regression är Pearson-residualer och devians-residualer två vanliga mått på skillnaden mellan observerade och anpassade värden, se [4] för detaljer. Genom att för respektive mått summera kvadraten av varje residual så erhålls Pearson χ^2 -statistikan, X^2 , respektive deviansen, D . Innan det kan slås fast att anpassningen är god behöver emellertid studeras om anpassningen stöds över hela datasetet. För att upptäcka observationer med dålig anpassning föreslår Hosmer et al. (2013) bland annat måtten ΔX_j^2 och ΔD_j , vilka anger förändringen i respektive statistika om observation j (eller alla observationer med samma uppsättning värden från modellens förklaringsvariabler som observation j) utesluts. Stora värden på dessa mått hjälper till att identifiera observationer där anpassningen tycks sämre och Hosmer et al. (2013) rekommenderar att de båda måtten plottas mot observationernas skattade sannolikheter för de olika utfallen. Vad som anses vara ett stort värde är en bedömning som måste baseras på det aktuella datasetet. Punkter där mätvärdet överstiger $\chi_{0.95}^2(1) = 3.84$ kan då ge vägledning, men Hosmer et al. (2013) förespråkar i första hand att punkter i övre högra och vänstra hörnet av diagrammet som tydligt avviker från övriga punkter analyseras. Huruvida en observation med dålig anpassning också har stort inflytande kan därefter undersökas genom att studera hur parameterskattningarna förändras om observationen utesluts.

Ett mått som direkt ser till en observations inflytande på parameterskattningarna är $\Delta \hat{\beta}_j$, vilket är analogt med måttet Cook's D för linjär regression. Detta fås som skillnaden mellan maximum likelihood-skattningarna med alla observationer inkluderade och där observation j (eller alla observationer med samma uppsättning värden från modellens förklaringsvariabler som observation j) utesluts, efter standardisering med kovariansmatrisen för $\hat{\beta}$. Även detta mått kan med fördel plottas mot observationernas skattade sannolikheter för de olika utfallen. Hosmer et al. (2013) påpekar att av erfarenhet måste mätvärdet i regel vara större än 1 för att ha effekt på parameterskattningar men att det alltid finns undantag och att alla värden som avviker från övriga kan uppmärksammas.

För beslut gällande om inflytelsrika observationer ska tas bort ur datamaterialet påpekar Hosmer et al. (2013) att rimlighetsbedömningar först måste göras och att inte per automatik exkludera allt som anses olägligt. Ovan beskrivna undersökningar av enskilda observationer syftar till att identifiera sådana där ett uteslutande kan vara aktuellt men att observationen också ska innehålla uppenbara felaktigheter för att uteslutas.

3 Beskrivning av data

3.1 Allmänt

Data har ursprungligen hämtats från Svenska Ishockeyförbundets statistikdatabas, se [5]. Det har dels varit alla matcher som spelats under respektive säsong från 2006/2007 till 2015/2016, med tillhörande resultat samt en del annan matchinformation och dels statistik som endast funnits tillgänglig i sammanställd form efter varje färdigspelad säsong. Dessa data har sedan bearbetats i Microsoft Excel för att passa den statistiska analysen mer ändamålsenligt och möjliggöra skapandet av flertalet förklaringsvariabler. Totalt har 3334 matcher spelats under de tio säsongerna, vilka används vid skapandet av förklaringsvariablerna. Ställningen har varit oavgjord efter två perioder i 771 matcher och dessa utgör observationerna för den kommande analysen. Vidare gäller att de 684 observationerna från de nio första säsongerna kommer ingå i modellframtagandet och de 87 observationerna från den senaste säsongen kommer användas för att testa modellens prediktionsförmåga. Observationerna behandlas som oberoende även om samma lag kan förekomma i flertalet observationer från samma säsong.

Innan vi går in på vilka variabler som kommer användas i analysen tittar vi närmare på vad som kan vara av intresse att undersöka beträffande utgången i jämna ishockeymatcher. Utifrån vad som är tillgängligt samt praktiskt möjligt att ta fram i våra data ses följande information som intressant att inhämta: lagens aktuella tabellplacering; lagens förmåga att prestera på hemma- respektive bortaplan; lagens benägenhet att göra samt släppa in mål; publikantalets påverkan; i vilken ordning lagen gjort eventuella mål i matchen; lagens prestationer mot varandra tidigare under säsongen; lagens prestationer de senaste matcherna; lagens prestationer under tidigare jämna matcher; lagens prestationer under tidigare tredje perioder; lagens prestationer i situationer med utvisade spelare; huruvida lagen förfogar över särskilda matchvinnare. Denna information försöker vi om möjligt täcka in genom variablerna i nästa avsnitt. Vissa variabler skiljer sig endast åt marginellt, vilket är för att undersöka om den ena framställningen representerar ett för analysen lämpligare mått på önskad information än den andra.

3.2 Variabler

I detta avsnitt presenteras först en sammanställning i tabellform av de variabler som ingår i analysen (Tabell 1) och därefter följer förklaringar av samma variabler. För alla variabler där värdet bygger på lagens tidigare prestationer gäller att detta endast avser tidigare prestationer under aktuell säsong. Vidare gäller att värden i de två sista kolumnerna i Tabell 1 endast ser till de nio första säsongerna.

Tabell 1: Variabler

| Variabel | Typ | Värden/Nivåer | Medianer/Frekvenser |
|-----------------------------|-----------------------|--|-------------------------------------|
| <i>resultat</i> | Kategori (respons) | 0 = oavgjort (ref) 1 = hemmaseger 2 = bortaseger | 0: 296 st 1: 225 st 2: 163 st |
| <i>omgång</i> | Kvantitativ | 1, ..., 55 | Median: 28 |
| <i>tabell</i> | Kvantitativ | -11, ..., 11 | Median: -1 |
| <i>tabell_kategori</i> | Kategori | 0 = {-11, ..., -4} 1 = {-3, ..., 3} (ref) 2 = {4, ..., 11} | 0: 195 st 1: 317 st 2: 172 st |
| <i>hemmaborta</i> | Kvantitativ | -11, ..., 11 | Median: 0 |
| <i>hemmaborta_kategori</i> | Kategori | 0 = {-11, ..., -4} 1 = {-3, ..., 3} (ref) 2 = {4, ..., 11} | 0: 170 st 1: 352 st 2: 162 st |
| <i>hemmaborta_kategori2</i> | Kategori | 0 = {-11, ..., -4} (ref) 1 = {-3, ..., 11} | 0: 522 st 1: 162 st |
| <i>gjorda_mål</i> | Kvantitativ | [-6, 4] | Median: 0 |
| <i>insläppta_mål</i> | Kvantitativ | [-4, 3] | Median: 0 |
| <i>publik</i> | Kvantitativ | 2424, ..., 18884 | Median: 5494 |
| <i>fullsatt</i> | Kategori | 0 = under 90% (95%) (ref) 1 = över 90% (95%) | 0: 460 (552) st 1: 224 (132) st |
| <i>första_mål</i> | Kategori | 0 = inget lag (ref) 1 = hemmalaget 2 = bortalaget | 0: 104 st 1: 300 st 2: 294 st |
| <i>sista_mål</i> | Kategori | 0 = inget lag (ref) 1 = hemmalaget 2 = bortalaget | 0: 104 st 1: 286 st 2: 294 st |
| <i>sista2_mål</i> | Kategori | 0 = inget lag (ref) 1 = hemmalaget 2 = bortalaget | 0: 591 st 1: 44 st 2: 49 st |
| <i>inbördes_totalt</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>inbördes_senaste</i> | Kategori | 0 = inget lag (ref) 1 = hemmalaget 2 = bortalaget | 0: 157 st 1: 213 st 2: 314 st |
| <i>senaste</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>senaste3</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>senaste5</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>jämn_vinst</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>jämn_förlust</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>tredje_vinst</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>tredje_förlust</i> | Kvantitativ | [-1, 1] | Median: 0 |
| <i>pp</i> | Kvantitativ | [-0.143, 0.1472] | Median: 0 |
| <i>pk</i> | Kvantitativ | [-0.1686, 0.1667] | Median: 0 |
| <i>topp10</i> | Kvantitativ | -4, ..., 4 | Median: 0 |
| <i>topp25</i> | Kvantitativ | -6, ..., 6 | Median: 0 |

Här följer nu förklaringar av analysens variabler.

resultat: Analysens responsvariabel, där resultatet efter tre spelade perioder tillhör någon av kategorierna 1, X, 2.

omgång: Varje observation tilldelas det värde som motsvarar den omgång under säsongen vilken matchen tillhör.

tabell: Observationerna tilldelas ett värde som motsvarar skillnaden i aktuell tabellplacering mellan lagen. Värdet är med avseende på hemmalaget, så exempelvis betyder -6 att bortalaget ligger sex placeringar högre upp i tabellen medan värdet 1 betyder att hemmalaget ligger en placering högre upp. Observationer från omgång 1 tilldelas värdet 0.

tabell_kategori: Variabeln är ett alternativt sätt att representera skillnaden i tabellplacering mellan lagen. Denna är tänkt att tolkas som betydande/marginell skillnad i placering och observationerna delas upp i någon av de tre kategorierna: mer än tre tabellplaceringars fördel för hemmalaget; mer än tre tabellplaceringars fördel för bortalaget; lagen skiljer sig inte mer än tre tabellplaceringar.

hemmaborta: Observationerna tilldelas ett värde som motsvarar skillnaden mellan hemmalagets placering i en aktuell tabell som enbart ser till resultat på hemmaplan och bortalagets placering i en motsvarande tabell för bortaplan. Värdet är även här med avseende på hemmalaget och för denna variabel gäller att observationer från omgång 1 och 2 tilldelas värdet 0, då varje lag först inför omgång 3 har spelat en hemmamatch och en bortamatch.

hemmaborta_kategori: Variabeln är uppbyggd på samma sätt som *tabell_kategori*, men nu med avseende på *hemmaborta* istället för på *tabell*.

gjorda_mål: Först beräknas det aktuella genomsnittliga antalet gjorda mål per match för respektive lag, varpå observationerna tilldelas ett värde som motsvarar skillnaden mellan dessa genomsnitt. Värdet är med avseende på hemmalaget, så positiva värden innebär att hemmalaget gör fler mål per match än bortalaget och omvänt för negativa värden.

insläppta_mål: Variabeln är uppbyggd likt *gjorda_mål* men med insläppta mål per match i fokus. Noterbart är att här innebär ett positivt värde att hemmalaget släpper in fler mål per match än bortalaget, vilket kan ses som en fördel för bortalaget.

publik: Observationerna tilldelas ett värde som motsvarar antalet åskådare som tittat på matchen.

fullsatt: Denna variabel ser till andelen åskådare i förhållande till arenans kapacitet. Observationerna delas upp i två kategorier baserat på om andelen är mer eller mindre än 90%, där procenttalet valts godtyckligt för att motsvara en någorlunda fullsatt arena.

första_mål: Observationerna delas upp i tre kategorier utifrån om hemmalaget eller bortalaget har gjort matchens första mål, förutsatt att målet kommit före andra periodens slut, eller om de två första perioderna varit mållösa.

sista_mål: Observationerna delas upp i tre kategorier utifrån om hemmalaget eller bortalaget har gjort det sista målet före andra periodens slut, eller om de två första perioderna varit mållösa.

sista2_mål: Observationerna delas upp i tre kategorier utifrån om de två sista målen före andra periodens slut har gjorts av enbart hemmalaget, enbart bortalaget, eller om inget av föregående har inträffat.

inbördes_totalt: Variabeln ser till hur tidigare möten under samma säsong har slutat mellan berörda lag. Först beräknas andelen poäng som varje lag har tagit av det totala antalet poäng som kunnat erhållas, varpå observationerna tilldelas ett värde som motsvarar skillnaden i dessa andelar. Värdet är med avseende på hemmalaget, så positiva värden innebär att hemmalaget har tagit fler poäng än bortalaget sett till deras tidigare möten under säsongen. Om lagen ännu inte mötts så tilldelas observationen värdet 0.

inbördes_senaste: Denna variabel ser endast till det senaste mötet mellan lagen. Observationerna är uppdelade i tre kategorier utifrån om hemmalaget, bortalaget eller inget av lagen tagit flest poäng i senaste mötet. Notera att vinst efter full tid ger tre poäng, vinst efter förlängning eller straffar två poäng, förlust efter förlängning eller straffar samt oavgjord förlängning före säsongen 2010/2011 en poäng och förlust efter full tid noll poäng. Därav följer att kategorin där inget av lagen tagit flest poäng förändras från och med 2010/2011 genom att matcher inte längre kan sluta oavgjort.

senaste: Först beräknas andelen poäng som varje lag har tagit av de totalt tre poäng som kunnat erhållas den senast spelade matchen, varpå observationerna tilldelas ett värde som motsvarar skillnaden i dessa andelar. Värdet är med avseende på hemmalaget, så positiva värden innebär att hemmalaget presterade poängmässigt bättre än bortalaget den senast spelade matchen. Observationer från omgång 1 tilldelas värdet 0.

senaste3: Precis som för *senaste* är detta en variabel som ser till berörda lags form, men denna gång beaktas de tre senast spelade matcherna. Obser-

vationernas värden beräknas analogt men här gäller att värden från omgång 2 endast tar hänsyn till senaste matchen och värden från omgång 3 endast tar hänsyn till senaste två matcherna.

senaste5: Som sista formvariabel är denna tänkt att undersöka om form sträcker sig så långt tillbaka som fem matcher och vilken påverkan på utfallet det i så fall har. Observationerna beräknas analogt med de två variablerna ovan men med justeringar i början av varje säsong fram till och med omgång 5.

jämn_vinst: Variabeln berör tidigare jämna matcher (oavgjort efter två spelade perioder) under säsongen och visar om något av lagen har större tendens att vinna dessa (under full tid). Observationernas värden utgörs av skillnaden mellan lagens andel vunna jämna matcher och där positiva värden innebär att hemmalaget har högre vinstandel. Om inte båda lagen har ingått i en jämn match tidigare under säsongen så tilldelas observationen värdet 0.

jämn_förlust: Denna variabel berör också tidigare jämna matcher under säsongen men visar om något av lagen har större tendens att förlora dessa. Observationernas värden följer samma princip som i *jämn_vinst* och här innebär positiva värden att hemmalaget har högre förlustandel, vilket kan ses som en fördel för bortalaget.

tredje_vinst: Variabeln berör prestationer i tredje perioden och visar om något av lagen har större tendens att vinna dessa. Observationernas värden utgörs av skillnaden mellan lagens andel vunna tredjeperioder och där positiva värden innebär att hemmalaget har högre vinstandel. Observationer från omgång 1 tilldelas värdet 0.

tredje_förlust: Denna variabel visar, i motsats till föregående, om något av lagen har större tendens att förlora tredjeperioder. Observationernas värden följer samma princip som i *tredje_vinst* och här innebär positiva värden att hemmalaget har högre förlustandel, vilket kan ses som en fördel för bortalaget.

För följande fyra variabler har det, till skillnad från variablerna ovan, inte varit möjligt att erhålla information som är aktuell sett till varje observation och omgång. Här gäller istället att värdena är en sammanställning för hela säsongen efter att den har spelats färdigt. Varje enskilt lag under en given säsong har således tilldelats värden på dessa variabler som inte förändras över säsongens olika omgångar. För varje observation gäller att dessa värden då också är delvis baserade på sådant som inträffat efter att aktuell match har spelats. Det här är givetvis problematiskt, inte minst ur ett prediktionsperspektiv, och hur det hanteras återkommer vi till senare i rapporten.

De fyra variablerna anses av författaren emellertid besitta potentiellt stor förklaringsförmåga, varför de också får ingå i variabeluppsättningen. Tolkningen som måste göras är att de värden som gäller efter avslutad säsong också är representativa som genomsnittliga värden sett till hela säsongen oavsett omgång. Det ska avslutningsvis tilläggas att aktuell information för dessa variabler kommer finnas tillgänglig om man ämnar prediktera utfallet för framtida jämna ishockeymatcher medan matcherna fortfarande pågår.

pp: PP är kortform för powerplay, vilket är ett internationellt vedertaget ord för att ett lag har fler spelare på isen än motståndarlaget till följd av utvisning(-ar). Denna variabel ser till skillnaden mellan lagens PP-effektivitet, dvs andelen PP-tillfällen som resulterar i mål för laget med fler spelare på isen. Positiva observationsvärden innebär att hemmalaget är effektivare i PP än bortalaget.

pk: PK är kortform för penalty killing, och i motsats till *pp* så är det istället laget med färre spelare på isen som betraktas. Denna variabel ser till skillnaden mellan lagens PK-effektivitet, dvs andelen PK-tillfällen som inte resulterar i mål för laget med fler spelare på isen. Positiva observationsvärden innebär att hemmalaget är effektivare på att 'döda utvisningar' än bortalaget.

topp10: Vid varje mål tilldelas målskytten, samt eventuellt en eller två passningsläggare en poäng. Om erhållna poäng summeras och jämförs mellan alla spelare får man vad som kallas för poängligan. Denna variabel ser till skillnaden mellan berörda lag i antalet spelare som återfinns bland poängligans tio första platser. Positiva observationsvärden innebär att hemmalaget har fler Topp 10-spelare än bortalaget.

topp25: Denna variabel är uppbyggd likt föregående men där jämförelsen görs mellan antalet Topp 25-spelare i berörda lag.

4 Analys av data

I denna del presenteras dels arbetet med att ta fram förklarande modeller genom multinomial logistisk regression och dels modellernas förmåga att prediktera utfallet av matcher från säsongen 2015/2016. Analysen utförs med hjälp av programspråket R. Som utgångspunkt följer modelleringen strategin *purposeful selection*, men även AIC-värde och klassificeringsförmåga ges viss betydelse. Med klassificeringsförmåga avses här andelen rätt klassificerade resultat på samma data som modelleringen utförs på, givet att klassificeringen sker på det utfall som ges högst sannolikhet av modellen.

Innan vi går vidare så tittar vi på hur utfallen för de 771 observationerna har fördelat sig. I Tabell 2 åskådliggörs antal och andelar av de olika utfallen, såväl totalt som över respektive säsong, samt även hur observationerna är fördelade över de olika säsongerna. Vi ser att det vanligaste resultatet är oavgjort, följt av hemmaseger och slutligen bortaseger. Skillnader mellan säsongerna ligger inom 10 procentenheter för bortaseger medan det för oavgjort skiljer 18 procentenheter som mest och för hemmaseger 22 procentenheter. I övrigt kan nämnas att säsongen 2013/2014 innehåller fler bortasegrar än hemmasegrar samt att säsongen 2015/2016, vilken ska predikteras, har högre andel hemmasegrar och lägre andel bortasegrar jämfört med de totala andelarna. Fler tabeller över enskilda variabler återfinns i Appendix A.2, Tabell 14-21.

Tabell 2: Observationernas utfall.

| Säsong | 1 | X | 2 | Totalt |
|---------------|-----------|-----------|-----------|--------|
| 2006/2007 | 26 (40%) | 27 (42%) | 12 (18%) | 65 |
| 2007/2008 | 25 (36%) | 25 (36%) | 19 (28%) | 69 |
| 2008/2009 | 33 (43%) | 25 (33%) | 18 (24%) | 76 |
| 2009/2010 | 23 (29%) | 38 (49%) | 17 (22%) | 78 |
| 2010/2011 | 27 (35%) | 35 (45%) | 16 (21%) | 78 |
| 2011/2012 | 24 (28%) | 42 (49%) | 19 (22%) | 85 |
| 2012/2013 | 21 (28%) | 34 (45%) | 20 (27%) | 75 |
| 2013/2014 | 16 (21%) | 39 (51%) | 21 (28%) | 76 |
| 2014/2015 | 30 (37%) | 31 (38%) | 21 (26%) | 82 |
| 2015/2016 | 36 (41%) | 35 (40%) | 16 (18%) | 87 |
| Totalt | 261 (34%) | 331 (43%) | 179 (23%) | 771 |

För att få en överblick över eventuella samband och beroenden mellan förklarande variabler så tittar vi även på korrelation mellan kvantitativa variabler och Cramérs V mellan kategorivariabler. Dessutom undersöks eventuella samband mellan kvantitativa och kategoriska variabler med hjälp av

lådagram. I Appendix A.1 återfinns delar av resultatet men där lådagramen har utelämnats då antalet möjliga kombinationer är mycket stort. Vi noterar att samband mellan olika variabler förekommer, men väljer att låta variabeluppsättningen vara intakt då många variabler troligen kommer försvinna i steg 1 av purposeful selection. Vidare bevakar vi osäkerheten i parameterskattningarna för variabler i multipla modeller genom att hålla ett öga på skattningarnas medelfel.

4.1 Statistisk modellering

Nedan följer en genomgång av de sju stegen för den statistiska modelleringen. Referenskategori för responsvariabeln är oavgjort resultat och i logit-funktion 1 jämförs hemmasegrar mot referenskategori medan bortasegrar jämförs mot referenskategori i logit-funktion 2.

4.1.1 Steg 1

I steg 1 gäller att för varje förklarande variabel anpassa en enkel multinomial logistisk regressionsmodell. Om p-värdet vid ett Likelihood Ratio-test (LR-test), för att testa hypotesen att variabeln inte har någon inverkan, är under 0.25 så kommer variabeln ingå i en första multipel regressionsmodell i steg 2. Övriga variabler försvinner däremot ur modelleringsarbetet tills vidare.

Följande variabler klarar signifikanskravet (ordnade efter lägst p-värde): *pk*, *topp25*, *sista2_mål*, *insläppta_mål*, *hemmaborta_kategori*, *sista_mål*, *fullsatt*, *hemmaborta*, *topp10*, *pp*. Då både *hemmaborta_kategori* och *hemmaborta* är signifikanta på 25%-nivån plockas den senare bort tills vidare. Detta ger en inledande multipel regressionsmodell med nio förklarande variabler till steg 2. För en fullständig tabell över p-värden från LR-testen, se Appendix A.3.1.

Det visar sig att två av nivåerna i variabeln *hemmaborta_kategori* inte skiljer sig signifikant i någon av logit-funktionerna. Vi väljer därför att slå ihop dessa nivåer och benämner den nya variabeln *hemmaborta_kategori2* (se Tabell 1 i föregående kapitel samt Appendix A.3.1 för detaljer kring variabeln).

4.1.2 Steg 2

En initial multipel regressionsmodell består nu av variablerna *pk*, *topp25*, *sista2_mål*, *hemmaborta_kategori2*, *insläppta_mål*, *sista_mål*, *fullsatt*, *topp10* och *pp*, och denna benämner vi MODELL0. LR-test visar att *pp*, *topp10* och *sista_mål* kan exkluderas från modellen. Däremot erhålls p-värden runt 0.07 vid LR-test för exkluderandet av såväl *fullsatt* som *insläppta_mål* och *pk*. Modeller utan dessa variabler får dessutom ökade AIC-värden och försämrade klassificeringsförmåga, varför vi låter alla tre finnas kvar. Fortsatta LR-test för ännu ej nämnda variabler ger signifikanta resultat och vi behåller därför

en modell innehållande pk , $topp25$, $sista2_mål$, $hemmaborta_kategori2$, $insläppta_mål$ och $fullsatt$, vilken vi benämner MODELL1.

Vi tar även fram modeller där alla ingående variabler måste vara signifikanta på 5%-nivån. Beroende på i vilken ordning variablernas signifikans testas erhålls två olika modeller. Den första, vilken fortsättningsvis benämns MODELL2, består av variablerna pk , $topp25$ och $sista2_mål$. Den andra modellen består av variablerna $topp25$, $sista2_mål$, $hemmaborta_kategori2$ samt $insläppta_mål$, och benämns framöver MODELL3.

I Appendix A.3.2, Tabell 23, återfinns AIC-värden och klassificeringsförmåga för modellerna som behålls under analysen. Vidare presenteras i slutet av avsnitt 4.1.7 en sammanställning av ingående variabler för alla modeller som kommer undersökas i prediktionssammanhang. Allmänt gäller att parameterskattningar endast redovisas i rapportens resultatdel för den modell vilken slutligen väljs som analysens primära modell.

4.1.3 Steg 3

I detta steg kommer förändringar i variablernas parameterskattningar mellan de olika modellerna att studeras närmare. På så sätt fås en indikation på om exkluderandet av någon variabel i steg 2 haft betydande inverkan på någon kvarvarande variabels skattningar.

Då parameterskattningarna i MODELL1 jämförs med samma variabelers parameterskattningar i MODELL0 visar det sig att fem skattningar har förändrats mer än 25%. Fyra av dessa är inte signifikant skilda från noll i berörd logit-funktion, och små förändringar i absoluta tal leder då till relativt stora förändringar i procent. Den femte skattningen är av större intresse och berör en av nivåerna i $sista2_mål$ i logit-funktion 2. En närmare titt visar att skattningen förändras med knappt 36% då variabeln $sista_mål$ exkluderas. Förklaringen ligger troligen i att variablerna överlappar en del, på så sätt att gör ett lag de två sista målen före andra periodens slut så gör de följaktligen även det sista målet före andra periodens slut. Vidare är värdet på Cramérs V för variablerna 0.31. Då det inte rör sig om någon avskräckande stor förändring av parameterskattningen, och då förändringen även tycks ha en rimlig förklaring, görs heller ingen ändring av MODELL1.

Vid jämförelse av parameterskattningarna mellan MODELL2 och MODELL1 så förändras två av skattningarna mer än 25%. Ingen av dessa är emellertid signifikant skild från noll i berörd logit-funktion och skattningarna undersöks därför inte närmare. Det är även två skattningar som förändras mer än 25% mellan MODELL3 och MODELL1 varav den ena kräver en närmare undersökning. Den berörda skattningen gäller för $insläppta_mål$ i

logit-funktion 1 och denna förändras nästan 77% då variabeln *pk* exkluderas. Det finns en svag korrelation mellan de båda variablerna; om det ena laget i genomsnitt släpper in fler mål per match så tenderar det även att släppa in fler mål än motståndaren i spelformen PK. Dessutom finns ett svagt samband med en annan ingående variabel, *hemmaborta_kategori2*, för både *pk* och *insläppta_mål* vilket kan leda till ökad osäkerhet i skattningen. Vidare har parameterskattningen för *pk* ett klart större medelfel än övriga parameterskattningars medelfel vilket också kan påverka skattningen av den svagt korrelerade variabeln *insläppta_mål*. Trots parameterskattningens förändring låter vi *pk* vara fortsatt exkluderad från MODELL3, då vi utöver denna modell också har en modell innehållande båda variablerna och en modell endast innehållande *pk*.

4.1.4 Steg 4

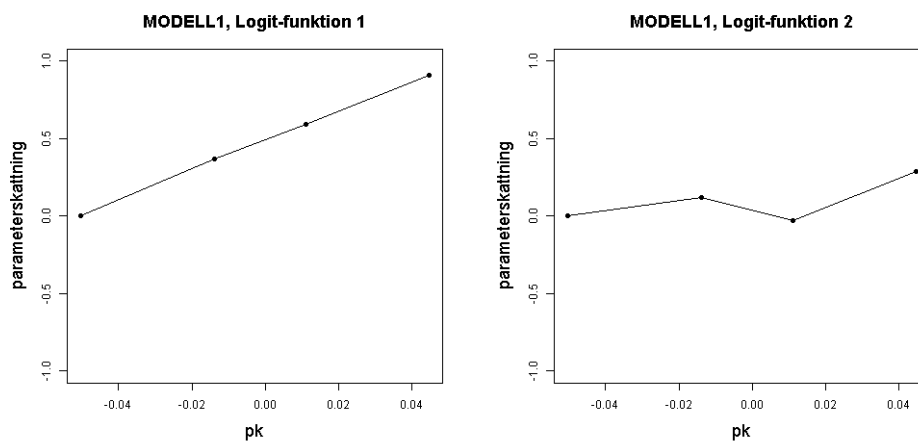
Steg 3 resulterade inte i någon förändring av modellerna, men däremot i en utökad kännedom om olika variablers påverkan på varandra. Nu följer steg 4, där det undersöks om någon av de exkluderade variablerna i steg 1 är signifikant i någon av de tre modellerna.

Det visar sig att ingen variabel klarar signifikansnivån 5%. Närmast ett signifikant resultat, i alla tre modellerna, är variabeln *jämn_vinst*, följt av *inbördes_totalt* och *inbördes_senaste*. Det lägsta p-värdet som erhålls för *jämn_vinst* är 0.16 i MODELL1. Noterbart är emellertid att om *inbördes_totalt* inkluderas i MODELL1 så uppnås det hittills bästa klassificeringsvärdet, se Appendix A.3.2, Tabell 23. Av den anledningen behåller vi en sådan modell under namnet MODELL4 och återupptar intresset för denna när det är dags för prediktion av säsongen 2015/2016.

4.1.5 Steg 5

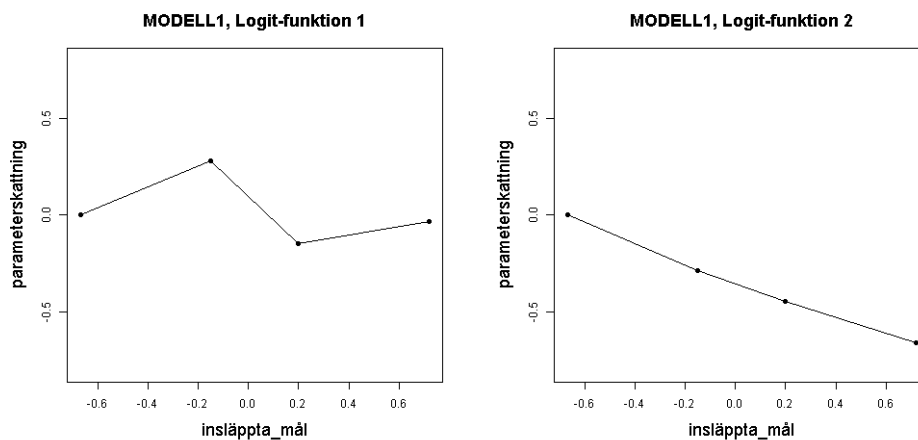
Syftet med steg 5 är att för varje kvarvarande kvantitativ förklaringsvariabel undersöka antagandet om att logit-funktionerna växer eller avtar linjärt som en funktion av förklaringsvariabeln. De kvantitativa variabler som finns representerade i modellerna är *pk*, *topp25* och *insläppta_mål*. Innan analysen av dessa sker så tittar vi på hur det påverkar variabeln *fullsatt* om dess nivåindelning justeras. Initialt valdes 90% som en godtycklig gräns för att avgöra om arenan under en match är nästan fullsatt eller inte. En ändring till 85% gör att variabeln inte längre är signifikant i någon modell, medan en ändring till 95% ger lägre p-värden vid LR-test i alla tre modeller. I Appendix A.3.2, Tabell 24, återfinns resultat från dessa test. Nivåindelningens gräns kommer i fortsättningen vara 95% och dessutom inkluderas också *fullsatt* i MODELL3. En kontroll visar att övriga variabler i den modellen fortfarande är signifikanta.

Vi går vidare med Steg 5 och först undersöks linjäriteten i logit-funktionerna för variabeln pk , vilken ingår i MODELL1 och MODELL2. Genom att omvandla denna till en kategorivariabel med fyra nivåer där den lägsta kvartilen utgör referensgrupp så kan parameterskattningar och medianvärden inom respektive nivå jämföras. Resultatet för de båda logit-funktionerna i MODELL1 kan ses i Figur 1. Diagrammen som erhålls för MODELL2 är mycket snarlika och presenteras inte. Linjäriteten i logit-funktion 1 är tydlig medan den är mer tveksam i logit-funktion 2. Däremot är variabeln endast signifikant i logit-funktion 1, och således är det linjäriteten i denna funktion som är av primärt intresse.



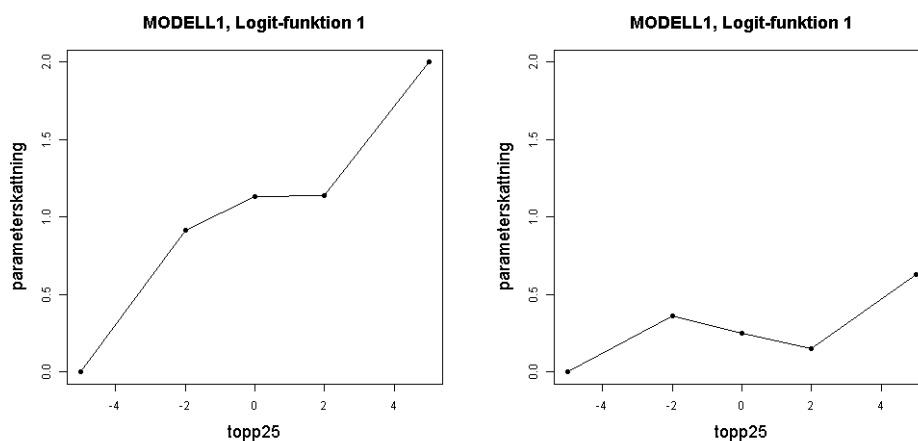
Figur 1: Diagram för parameterskattningar och medianvärden inom kvartilnivåer av pk .

Härnäst undersöks *insläppta_mål*, vilken ingår i MODELL1 och MODELL3. Även denna omvandlas till en kategorivariabel med fyra nivåer och förfarandet är analogt med hur pk undersöktes. I Figur 2 presenteras resultatet för de båda logit-funktionerna i MODELL1. Diagrammen med MODELL3 är närmast identiska och utelämnas därför. Denna gång är linjäriteten i logit-funktion 2 uppenbar medan den är högst tvivelaktig i logit-funktion 1. För *insläppta_mål* gäller emellertid det omvända mot pk , att variabeln är signifikant i logit-funktion 2 men inte i logit-funktion 1.



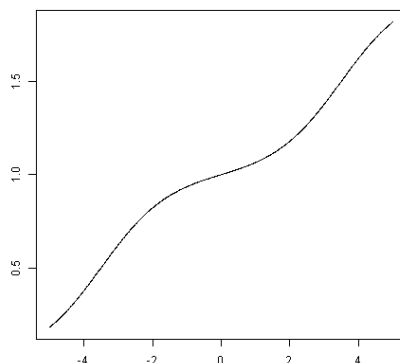
Figur 2: Diagram för parameterskattningar och medianvärden inom kvartilnivåer av *insläppta_mål*.

Avslutningsvis undersöks *topp25*, vilken ingår i alla tre modeller. Variabeln antar 13 olika värden av vilka över 50% av observationerna har något av värdena -1, 0 eller 1. I detta fall blir det mer lämpligt att frångå kvartiluppdelning och istället använda fem nivåer indelade enligt följande: -6 till -4, -3 till -1, 0, 1 till 3 samt 4 till 6. Bortsett från detta så är tillvägagångssättet detsamma som för de två tidigare variablerna. Diagrammen mellan de tre olika modellerna skiljer sig inte nämnvärt och i Figur 3 presenteras resultatet för logit-funktionerna i MODELL1. I logit-funktion 1, där parameterskattningen är signifikant, finns vissa linjäritetstendenser men resultatet tyder mer på att funktionen växer enligt en S-kurva.



Figur 3: Diagram för parameterskattningar och medianvärden inom nivåer av *topp25*.

På grund av kurvans utseende i logit-funktion 1 testar vi att ersätta *topp25* med en sigmoid-funktion i modellerna. Enligt Figur 3 så verkar kurvan plana ut på mitten och vara brantare i början samt på slutet. Vi testar därför med en funktion på formen $\frac{1}{1+e^{-3.5-topp25}} + \frac{1}{1+e^{3.5-topp25}}$, vilken kan ses i Figur 4.



Figur 4: Sigmoid-funktion.

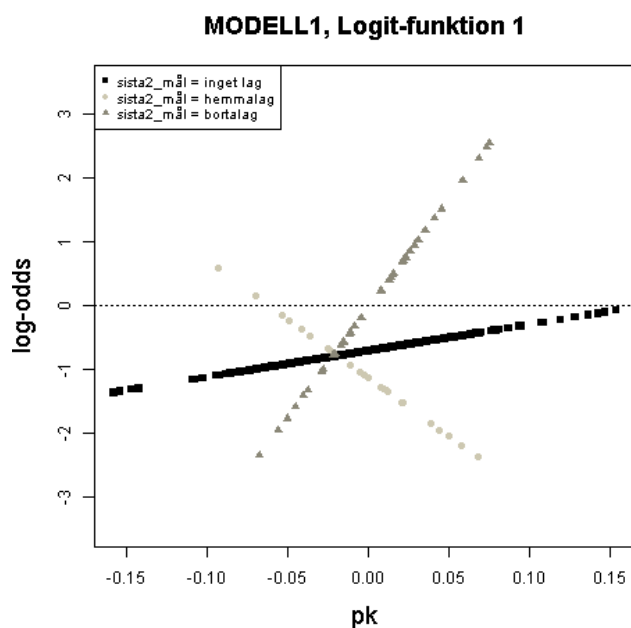
Att göra denna ändring i modellerna ger aningen lägre p-värden vid LR-test samt vid Wald-test av parameterskattningen för logit-funktion 1. Nackdelarna med att införa en ändring av den här typen är dels att modellen riskerar att bli överanpassad och för komplicerad i förhållande till antalet observationer vi har att arbeta med och dels att tolkningen av variabelns påverkan på utfallet av jämna ishockeymatcher möjligen försvåras. Det finns emellertid en rimlig tolkning av en sigmoid-funktion av detta slag för variabeln. Vid stora skillnader i antalet Topp 25-spelare mellan lagen så ökar sannolikheten för hemmaseger/oavgjort medan om endast en eller ingen spelare skiljer lagen så är variabeln av liten betydelse. En till aspekt som talar för en ändring är att det hittills lägsta AIC-värdet erhålls då MODELL1 ändras, se Appendix A.3.2, Tabell 23. Därför kommer en sådan förändring av MODELL1 behållas under namnet MODELL5.

4.1.6 Steg 6

Steg 6 berör eventuella samspelseffekter i modellerna och LR-test utförs för alla tänkbara tvåvägssamspel mellan variablerna inom varje modell. För inkluderande av en samspelsterm i en modell gäller att den ska vara statistiskt signifikant men också kunna tolkas på ett rimligt sätt. LR-testen resulterar i signifikanta samspel på 5%-nivån mellan variablerna *pk* och *sista2_mål* samt mellan *insläppta_mål* och *sista2_mål* i de modeller där dessa variabler förekommer. I Appendix A.3.2, Tabell 25 återfinns en sammanställning av alla

samspel som är signifikanta på 10%-nivån.

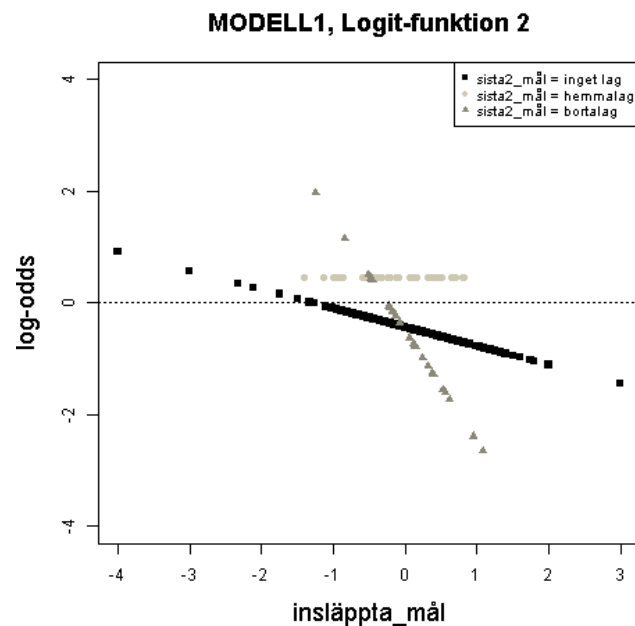
Vad gäller samspelet mellan pk och $sista2_mål$ så är det svårt att göra en annan tolkning än att det tycks långsökt. Samspelet är emellertid signifikant för båda nivåerna av $sista2_mål$ i logit-funktion 1 och i Figur 5 används MODELL1 för att åskådliggöra detta i berörd funktion.



Figur 5: Diagram av skattad logit som funktion av pk med avseende på de tre nivåerna i $sista2_mål$.

Av diagrammet framgår bland annat att bra presterande i spelformen PK minskar sannolikheten för vinst om man också gjort de två sista målen, vilket inte upplevs som rimligt. Det ska påpekas att endast 25 av de 521 observationer som faller under logit-funktion 1 har haft ett hemmalag som gjort de två sista målen före andra periodens slut, vilket medför en osäkerhet kring giltigheten i denna samspelseffekt. Dessutom är parameterskattningarnas medelfel för samspelet flera gånger större än medelfelen för modellens övriga skattningar. Med sentensen att enklare ofta är bättre beslutas därför att denna samspelseffekt inte kommer ingå i fortsatt analys.

Vi går vidare och undersöker samspelet mellan $insläppta_mål$ och $sista2_mål$ och även här är det svårt att tolka det som annat än långsökt. Återigen används MODELL1 för att åskådliggöra samspelet och i Figur 6 ses log-oddset i logit-funktion 2 av $insläppta_mål$ för de tre nivåerna i $sista2_mål$.

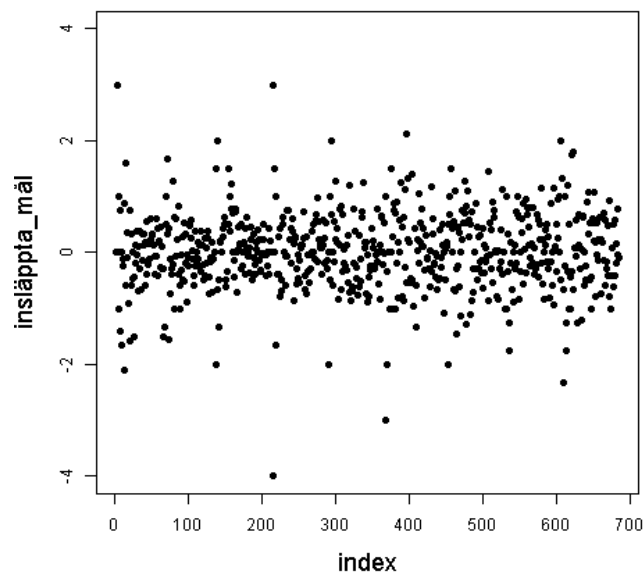


Figur 6: Diagram av skattad logit som funktion av *insläppta_mål* med avseende på de tre nivåerna i *sista2_mål*.

Diagrammet tycks inte ge vägledning till en rimlig tolkning av samspelet och vi väljer därför att inte inkludera detta i någon modell. Däremot framkommer något intressant gällande variabeln *insläppta_mål* när diagrammet studeras. De svarta fyrkanterna i Figur 6 visar att om bortalaget i genomsnitt släpper in färre mål än hemmalaget så minskar log-oddset för bortaseger jämfört med oavgjort. Detta talar mot vad som känns intuitivt riktigt, men samtidigt kan det tänkas att ett bortalag som släpper in få mål ofta får ett oavgjort resultat efter tre perioder genom att tredje perioden då oftare skulle bli mållös.

Vi undersöker *insläppta_mål* i ett spridningsdiagram (Figur 7) för att se om allt står rätt till. I diagrammet ses att ett fåtal observationer kan tänkas ha väldigt stort inflytande på variabeln. Det visar sig att de fyra mest extrema observationsvärdena alla är matcher som tillhör omgång 2 och har tilldelats värden på variabeln som skulle vara i det närmaste otänkbara längre in på säsongen. Dessutom gäller att de två mest extrema negativa värdena på *insläppta_mål* har slutat med bortaseger och de två mest extrema positiva värdena har slutat oavgjort efter full tid. Om dessa observationer plockas bort ger nya LR-test att variabeln inte längre är signifikant i någon av modellerna som den ingår i. Däremot kvarstår tendensen i logit-funktion 2; om bortalaget släpper in färre mål i genomsnitt så ökar sannolikheten för ett

oavgjort resultat efter full tid.



Figur 7: Spridningsdiagram för *insläppta_mål* över de 684 observationerna.

Med de nya resultat som erhålls för *insläppta_mål* blir variabelns medverkan i modellerna mer tveksam. Vi väljer därför att dela upp MODELL1 i en modell där *insläppta_mål* i sin ursprungliga form ingår och en där den är bortplockad helt. Vidare exkluderas variabeln från MODELL3, vilken endast består av variabler som klarar signifikansnivån 5%. Det här medför att det enda som skiljer MODELL3 från MODELL1 efter exkluderandet av *insläppta_mål* är variabeln *pk*, vilken är signifikant i MODELL1. MODELL3 bedöms därför vara överflödigt och vi tar bort den från fortsatt analys. Istället låter vi MODELL1 med *insläppta_mål* exkluderad ärvä benämningen MODELL3.

4.1.7 Steg 7

I detta slutsteg av purposeful selection kommer modellernas anpassning studeras genom Goodness of fit-tester. Vi inleder med att titta på övergripande mått av anpassning och avslutar med att studera individuella komponenter, vilket förvisso tjuvstartade redan i steg 6 med de inflytelserika observationsvärdena för variabeln *insläppta_mål*. Testerna kommer genomföras efter att de multinomiala modellerna delats upp i binära modeller: en där vi jämför hemmasegrar mot oavgjorda resultat och en där vi jämför bortasegrar mot oavgjorda resultat. Variablernas parameterskattningar förändras inte

nämnvärt till följd av denna uppdelning. Vidare gäller att antalet observationer, av de totalt 684, som har en unik uppsättning värden på förklaringsvariablerna är 682, 640, 657 respektive 682 för MODELL1, MODELL2, MODELL3 respektive MODELL5. För de icke unika uppsättningarna gäller att samma mönster endast förekommer två gånger.

Först utförs Hosmer-Lemeshows test med uppdelning i tio grupper baserat på percentiler av de skattade sannolikheterna för utfallen. I Tabell 3 presenteras p-värden för de åtta olika testen och då inget av dessa ligger nära noll finns heller inget som indikerar att anpassningen är dålig. Vidare kan ses i Appendix A.3.3, Tabell 26-33, att observerade och förväntade antal observationer för de olika grupperna överensstämmer väl i de flesta fall.

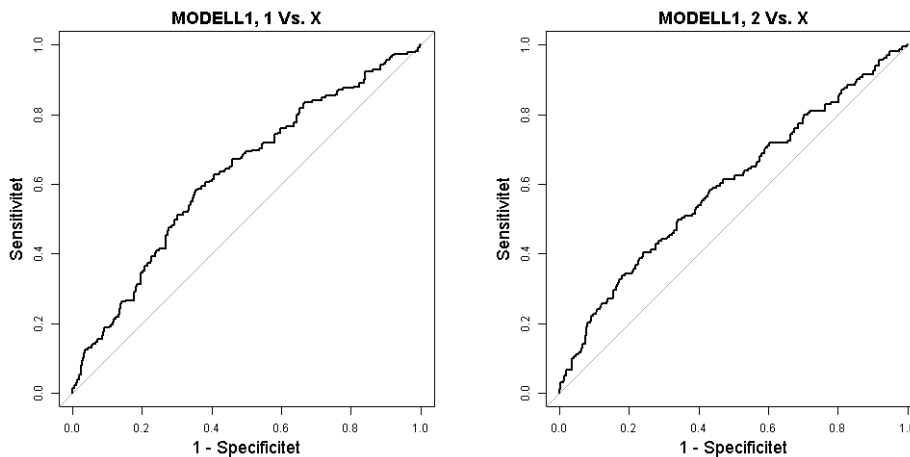
Tabell 3: Hosmer-Lemeshow-testens p-värden för de olika modellernas binära uppdelningar.

| Modell | 1 Vs. X | 2 Vs. X |
|---------------|----------------|----------------|
| MODELL1 | 0.82 | 0.55 |
| MODELL2 | 0.32 | 0.56 |
| MODELL3 | 0.68 | 0.92 |
| MODELL5 | 0.52 | 0.50 |

Vidare undersöks ROC och i synnerhet arean under dessa ROC-kurvor för de olika modellernas binära uppdelningar. I Tabell 4 återfinns de olika AUC-värdena inklusive 95% konfidensintervall och i Figur 8 ses de båda ROC-kurvorna för de binära uppdelningarna av MODELL1. Vi ser att värdena är för ändamålet låga och inte mycket bättre än att singla slant enligt de generella tumregler som presenterades i avsnitt 2.3.7. Om vi istället plockar bort alla oavgjorda resultat och undersöker uppdelningen i hemma- och bortasegrar för MODELL1 så ökar AUC-värdet något till 0.65. Tittar vi istället på uppdelningen i oavgjort och inte oavgjort resultat för MODELL1 så blir AUC-värdet så lågt som 0.60.

Tabell 4: AUC-värden för de olika modellernas binära uppdelningar.

| Modell | 1 Vs. X | 2 Vs. X |
|---------------|-------------------|-------------------|
| MODELL1 | 0.63 (0.58, 0.67) | 0.61 (0.55, 0.66) |
| MODELL2 | 0.61 (0.56, 0.66) | 0.57 (0.51, 0.62) |
| MODELL3 | 0.62 (0.58, 0.67) | 0.59 (0.54, 0.65) |
| MODELL5 | 0.63 (0.58, 0.68) | 0.61 (0.55, 0.66) |



Figur 8: ROC-kurvor för de binära uppdelningarna av MODEL1.

Eftersom en ROC-kurva endast tar hänsyn till två av de tre utfallen så tittar vi även på hur de anpassade modellerna lyckas klassificera alla tre utfallen. I Appendix A.3.4, Tabell 34-37, presenteras klassificeringstabeller för de fyra modellerna och i Tabell 5 sammanfattas resultatet mellan modellerna. Procenttalen för rätt klassificering av de olika resultaten är beräknade som antalet rätt klassificerade av resultatet dividerat med antalet observerade utfall av resultatet. Vi ser att MODEL1 och MODEL5 klassificerar något bättre totalt sett, men att alla modellerna är dåliga på att klassificera resultaten 1 och 2 rätt. Om de olika procenttalen istället beräknas som antalet rätt klassificerade av resultatet dividerat med totala antalet klassificerade av resultatet ligger värdena för rätt klassificering av 1, X och 2 mellan 44-55% i samtliga fall.

Tabell 5: Jämförelse mellan modellernas klassificeringsförmåga.

| Modell | Rätt totalt | Rätt 1 | Rätt X | Rätt 2 | Andelar 1, X, 2 |
|--------|-------------|--------|--------|--------|-----------------------|
| MODEL1 | 47.08% | 33.33% | 76.01% | 13.50% | 23.54%, 70.32%, 6.14% |
| MODEL2 | 44.88% | 25.78% | 78.38% | 10.43% | 18.27%, 75.88%, 5.85% |
| MODEL3 | 45.76% | 32.00% | 76.01% | 9.82% | 23.25%, 71.64%, 5.12% |
| MODEL5 | 46.64% | 32.44% | 75.34% | 14.11% | 23.83%, 70.03%, 6.14% |

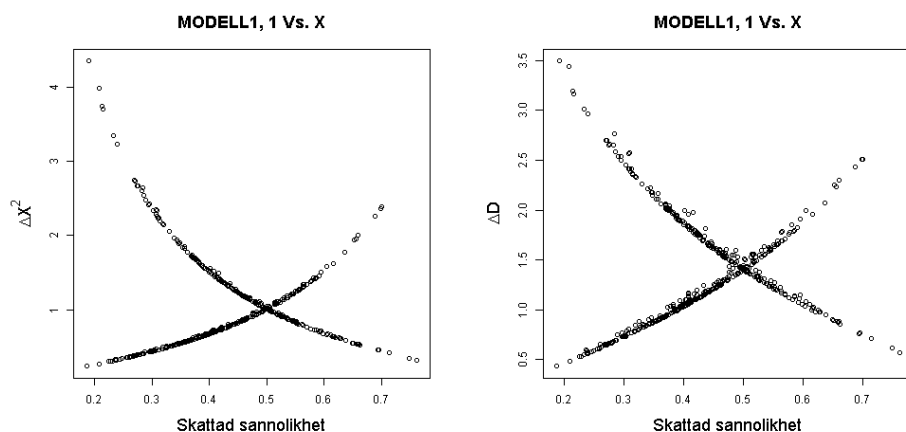
Genom att testa hur en modell endast bestående av ett intercept klassificerar erhålls 43.27% rätt klassificerade totalt sett, så modellerna är endast marginellt bättre än slumpen. Det mest anmärkningsvärda i Tabell 5 är annars hur modellerna klassificerar för stor andel resultat som X, då de riktiga andelarna för resultaten 1, X, 2 är 32.89%, 43.27% respektive 23.83%. I ett försök att korrigera detta försöker vi att begränsa antalet klassificerade X till ungefär rätt andel, samtidigt som vi främjar klassificeringen av 2 aning-

en. Vi vet att antalet oavgjorda utfall är 296 st, och finner att beträffande MODELL1 så har 296 st observationer en skattad sannolikhet för oavgjort resultat som överstiger 0.445. Detta sannolikhetsvärde ger även ett liknande resultat för övriga modeller. En ny klassificering sker därför enligt följande kriterier: om sannolikheten för oavgjort resultat är större än 0.445 så klassificera som X; annars, om sannolikheten för hemmaseger är mer än 0.05 större än sannolikheten för bortaseger så klassificera som 1; annars, klassificera som 2. En sammanfattning av klassificeringen enligt de nya kriterierna finns i Tabell 6. Vi ser att, bortsett från MODELL2, så överensstämmer klassificeringsandelarna nu ganska väl med de riktiga andelarna för de olika resultaten.

Tabell 6: Ny jämförelse mellan modellernas klassificeringsförmåga.

| Modell | Rätt totalt | Rätt 1 | Rätt X | Rätt 2 | Andelar 1, X, 2 |
|---------|-------------|--------|--------|--------|------------------------|
| MODELL1 | 48.10% | 53.78% | 52.70% | 31.90% | 37.84%, 43.27%, 18.86% |
| MODELL2 | 44.15% | 59.56% | 48.65% | 14.72% | 47.51%, 42.54%, 9.94% |
| MODELL3 | 46.20% | 55.11% | 47.97% | 30.67% | 40.35%, 40.64%, 19.01% |
| MODELL5 | 48.10% | 53.78% | 52.36% | 32.52% | 38.01%, 42.69%, 19.30% |

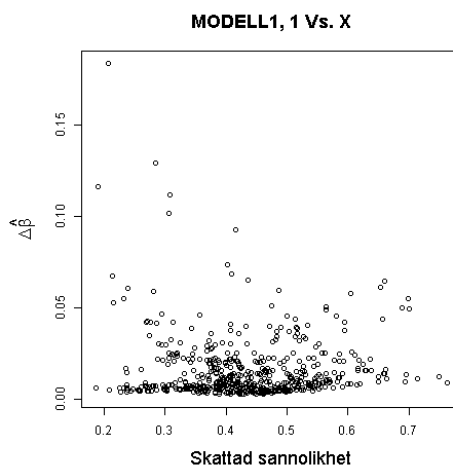
Vi övergår nu till att titta på om det finns enskilda observationer som påverkar anpassningen markant eller som har betydande påverkan på parameterskattningarna. Först undersöks ΔX_j^2 respektive ΔD_j , vilka beskrivs i avsnitt 2.3.8, för den binära uppdelningen i hemmasegrar och oavgjorda resultat. Var och ett av dessa mått plottas mot den skattade sannolikheten för de observerade utfallen. Diagrammen för de fyra olika modellerna skiljer sig inte nämnvärt och i Figur 9 visas resultatet för MODELL1.



Figur 9: Diagram för ΔX^2 respektive ΔD och skattade sannolikheter för hemmasegrar och oavgjorda resultat.

Vi ser att de båda diagrammen är snarlika och uppvisar kurvor av kvadratisk karaktär, där kurvan som startar i övre vänstra hörnet motsvarar hemmasegrar och kurvan som startar i nedre vänstra hörnet motsvarar oavgjorda resultat. I enlighet med rekommendationerna som beskrevs i avsnitt 2.3.8 väljs de sex observationer som avviker något i övre vänstra hörnet ut för vidare analys. Det är samma observationer som sticker ut i båda diagrammen och genom att exkludera dessa, dels en i taget och dels alla samtidigt, så kan parameterskattningarnas förändringar undersökas. Endast då alla sex observationer exkluderas påverkas vissa parameterskattningar nämnvärt. Förklaringen tycks ligga i stora negativa värden för variablerna *pk* och *topp25*, vilket ger stöd för andra resultat än hemmaseger medan det faktiska resultatet i matcherna varit just hemmaseger. Detta är emellertid ingenting orimligt utan snarare en följd av ishockeymatchers oförutsägbara natur. Då dessa observationer inte innehåller några uppenbara felaktigheter finns heller inga skäl till att utesluta dem från våra data.

Vi går vidare med att, för samma binära uppdelning i hemmasegrar och oavgjorda resultat, undersöka $\Delta\hat{\beta}_j$ från avsnitt 2.3.8. Detta mått identifierar om observation *j* har stort inflytande på en eller flera parameterskattningar. Diagrammen för de olika modellerna är snarlika och det är samma observation som har störst $\Delta\hat{\beta}$ -värde i samtliga modeller. I Figur 10 visas resultatet för MODELL1 där $\Delta\hat{\beta}$ plottas mot den skattade sannolikheten för de observerade utfallen.



Figur 10: Diagram för $\Delta\hat{\beta}$ och skattade sannolikheter för hemmasegrar och oavgjorda resultat.

Den observation som sticker ut mest har ett $\Delta\hat{\beta}$ -värde på drygt 0.18, vilket är betydligt mindre än riktmärket 1 som anges i avsnitt 2.3.8. Observatio-

nen är emellertid inte inkluderad bland de föregående sex observationerna och då den avviker en del från övriga observationer undersöks den närmare. Något anmärkningsvärt som skulle tyda på felaktigheter kan däremot inte ses bland observationens uppsättning av variabelvärden och således finns det fortsatt ingen anledning till förändringar i data.

På motsvarande sätt undersöks ΔX^2 , ΔD och $\Delta \hat{\beta}$ för den binära uppdelningen i bortasegrar och oavgjorda resultat. Inget som skulle ge anledning till ett uteslutande av observationer upptäckts och därför utelämnas detaljerna från denna undersökning.

I och med detta så avslutas den statistiska modelleringen och de fem modeller som presenteras i Tabell 7 kommer härnäst undersökas i prediktions-sammanhang:

Tabell 7: Modeller från purposeful selection.

| | MODELL1 | MODELL2 | MODELL3 | MODELL4 | MODELL5 |
|-----------------------------|---------|---------|---------|---------|---------|
| <i>pk</i> | X | X | X | X | X |
| <i>topp25</i> | X | X | X | X | X* |
| <i>sista2_mål</i> | X | X | X | X | X |
| <i>hemmaborta_kategori2</i> | X | | X | X | X |
| <i>fullsatt</i> | X | | X | X | X |
| <i>insläppta_mål</i> | X | | | X | X |
| <i>inbördes_totalt</i> | | | | X | |

*I MODELL5 är *topp25* ersatt av funktionen $\frac{1}{1+e^{-3.5-topp25}} + \frac{1}{1+e^{3.5-topp25}}$.

4.2 Prediktion

Som tidigare har belysts så kommer prediktion att ske på de 87 observationerna från säsongen 2015/2016, vilka inte ingått i modelleringen. Prediktionsresultatet för respektive modell återfinns i Appendix A.4.1, Tabell 38-42. I Tabell 8 sammanfattas resultatet mellan modellerna, där procentalen för rätt prediktion av de olika resultaten är beräknade som antalet rätt predikterade av resultatet dividerat med antalet observerade utfall av resultatet. Hur det förhåller sig mellan modellerna om de olika procentalen istället beräknas som antalet rätt predikterade av resultatet dividerat med totala antalet predikterade av resultatet kan ses i Tabell 9.

Tabell 8: Jämförelse mellan modellernas prediktionsförmåga.

| Modell | Rätt totalt | Rätt 1 | Rätt X | Rätt 2 | Andelar 1, X, 2 |
|---------|-------------|--------|--------|--------|-----------------------|
| MODELL1 | 36.78% | 19.44% | 68.57% | 6.25% | 22.99%, 67.82%, 9.20% |
| MODELL2 | 34.48% | 8.33% | 74.29% | 6.25% | 14.94%, 77.01%, 8.05% |
| MODELL3 | 40.23% | 22.22% | 74.29% | 6.25% | 22.99%, 71.26%, 5.75% |
| MODELL4 | 37.93% | 19.44% | 71.43% | 6.25% | 21.84%, 68.97%, 9.20% |
| MODELL5 | 36.78% | 19.44% | 68.57% | 6.25% | 22.99%, 67.82%, 9.20% |

Tabell 9: Jämförelse mellan modellernas prediktionsförmåga vid alternativ beräkning.

| Modell | Rätt totalt | Rätt 1 | Rätt X | Rätt 2 | Andelar 1, X, 2 |
|---------|-------------|--------|--------|--------|-----------------------|
| MODELL1 | 36.78% | 35.00% | 40.68% | 12.50% | 22.99%, 67.82%, 9.20% |
| MODELL2 | 34.48% | 23.08% | 38.81% | 14.29% | 14.94%, 77.01%, 8.05% |
| MODELL3 | 40.23% | 40.00% | 41.94% | 20.00% | 22.99%, 71.26%, 5.75% |
| MODELL4 | 37.93% | 36.84% | 41.67% | 12.50% | 21.84%, 68.97%, 9.20% |
| MODELL5 | 36.78% | 35.00% | 40.68% | 12.50% | 22.99%, 67.82%, 9.20% |

Vi ser att prediktionsförmågan följer modellernas förmåga att klassificera från avsnitt 4.1.7 och X predikteras i alldeles för hög andel. MODELL3 lyckas prediktera något bättre än övriga modeller totalt sett, men endast på samma nivå som en modell bestående av enbart ett intercept. Vidare ger p-värden från Fishers exakta test för de olika modellerna inget stöd för att förkasta hypotesen om oberoende. Vi utför därför en ny prediktion och ändrar kriterierna analogt med hur de ändrades vid klassificeringen i avsnitt 4.1.7. Detta innebär att den nya prediktionen sker enligt följande: om sannolikheten för oavgjort resultat är större än 0.445 så prediktera som X; annars, om sannolikheten för hemmaseger är mer än 0.05 större än sannolikheten för bortaseger så prediktera som 1; annars, prediktera som 2.

Det nya prediktionsresultatet för respektive modell återfinns i Appendix A.4.2, Tabell 43-47. I Tabell 10 sammanfattas resultatet mellan modellerna, där procentalen för rätt prediktion av de olika resultaten är beräknade som antalet rätt predikterade av resultatet dividerat med antalet observerade utfall av resultatet. Hur det förhåller sig mellan modellerna om de olika procentalen istället beräknas som antalet rätt predikterade av resultatet dividerat med totala antalet predikterade av resultatet kan ses i Tabell 11.

Tabell 10: Ny jämförelse mellan modellernas prediktionsförmåga.

| Modell | Rätt totalt | Rätt 1 | Rätt X | Rätt 2 | Andelar 1, X, 2 |
|---------|-------------|--------|--------|--------|------------------------|
| MODELL1 | 37.93% | 30.56% | 51.43% | 25.00% | 33.33%, 50.57%, 16.09% |
| MODELL2 | 32.18% | 44.44% | 31.43% | 6.25% | 50.57%, 35.63%, 13.79% |
| MODELL3 | 42.53% | 36.11% | 51.43% | 37.50% | 34.48%, 48.28%, 17.24% |
| MODELL4 | 35.63% | 27.78% | 51.43% | 18.75% | 33.33%, 50.57%, 16.09% |
| MODELL5 | 37.93% | 30.56% | 48.57% | 31.25% | 33.33%, 49.43%, 17.24% |

Tabell 11: Ny jämförelse mellan modellernas prediktionsförmåga vid alternativ beräkning.

| Modell | Rätt totalt | Rätt 1 | Rätt X | Rätt 2 | Andelar 1, X, 2 |
|---------|-------------|--------|--------|--------|------------------------|
| MODELL1 | 37.93% | 37.93% | 40.91% | 28.57% | 33.33%, 50.57%, 16.09% |
| MODELL2 | 32.18% | 36.36% | 35.48% | 8.33% | 50.57%, 35.63%, 13.79% |
| MODELL3 | 42.53% | 43.33% | 42.86% | 40.00% | 34.48%, 48.28%, 17.24% |
| MODELL4 | 35.63% | 34.48% | 40.91% | 21.43% | 33.33%, 50.57%, 16.09% |
| MODELL5 | 37.93% | 37.93% | 39.53% | 33.33% | 33.33%, 49.43%, 17.24% |

Återigen ser vi att MODELL3 predikterar bättre än övriga modeller, och resultatet förbättras också något från föregående prediktion. Även om vi inte kan förkasta hypotesen om oberoende vid Fishers exakta test för denna modell är p-värdena låga då 2 jämförs mot 1 respektive X. Beräkning av Cramérs V ger värdet 0.19 då alla utfall beaktas, värdet 0.33 då 2 och X beaktas, värdet 0.32 då 2 och 1 beaktas samt värdet 0 då 1 och X beaktas. Det finns alltså tecken på association mellan predikterade och observerade utfall när bortasegrar jämförs med något av övriga resultat men inte när jämförelsen görs mellan hemmasegrar och oavgjorda resultat.

Avslutningsvis väljer vi en modell som analysens primära modell till att förklara och prediktera utfallet av jämna ishockeymatcher. MODELL3 är den modell som predikterar rätt flest gånger vid båda prediktionstillfällena. Den innehåller dessutom inte den tveksamma variabeln *insläppta_mål* och därför väljs denna som primär modell. MODELL2 predikterar generellt sämre än övriga modeller och är, trots principen om Ockhams rakkniv, för enkel då det bara finns en signifikant parameterskattning i logit-funktion 2. Denna, nivån för att hemmalaget gör de sista två målen i *sista2_mål*, innehåller dessutom bara 44 av de 684 observationerna.

5 Resultat

Analysen av data har lett fram till en primär modell med syftet att förklara och prediktera utfallet av jämna ishockeymatcher. Med oavgjort resultat som referenskategori för responsvariabeln och med hemmasegrar som jämförs med referenskategori i logit-funktion 1 samt bortasegrar som jämförs med referenskategori i logit-funktion 2 så sammanfattas denna modell i Tabell 12. Förklaringar av ingående variabler återfinns i avsnitt 3.2.

Tabell 12: Parameterskattningar, skattade medelfel, Wald-statistikor samt tvåsidigt p-värde för variablerna i den primära modellen.

| | Variabel | Koefficient | Medelfel | z | p |
|----------------------|-------------------------------------|-------------|----------|--------|-------|
| Logit 1 (1 Vs. X) | <i>intercept</i> | -0.756 | 0.283 | -2.668 | 0.008 |
| | <i>pk</i> | 5.602 | 2.070 | 2.706 | 0.007 |
| | <i>topp25</i> | 0.095 | 0.042 | 2.260 | 0.024 |
| | <i>sista2_mål, nivå 1</i> | -0.337 | 0.431 | -0.781 | 0.435 |
| | <i>sista2_mål, nivå 2</i> | 0.662 | 0.336 | 1.971 | 0.049 |
| | <i>hemmaborta_kategori2, nivå 1</i> | -0.002 | 0.230 | -0.011 | 0.991 |
| | <i>fullsatt, nivå 1</i> | -0.545 | 0.234 | -2.330 | 0.020 |
| Logit 2 (2 Vs. X) | <i>intercept</i> | -0.547 | 0.288 | -1.897 | 0.058 |
| | <i>pk</i> | 2.330 | 2.242 | 1.039 | 0.299 |
| | <i>topp25</i> | -0.020 | 0.046 | -0.431 | 0.667 |
| | <i>sista2_mål, nivå 1</i> | 0.863 | 0.359 | 2.405 | 0.016 |
| | <i>sista2_mål, nivå 2</i> | -0.090 | 0.442 | -0.204 | 0.838 |
| | <i>hemmaborta_kategori2, nivå 1</i> | -0.501 | 0.231 | -2.170 | 0.030 |
| | <i>fullsatt, nivå 1</i> | -0.315 | 0.254 | -1.239 | 0.215 |

Nedan följer de skattade odds-kvoter från modellen som har ett signifikant resultat.

- En ökning med 0.05 av *pk*, skillnaden mellan hemmalagets och bortalagets förmåga att inte släppa in mål i spelformen PK, ökar det skattade oddset för hemmaseger jämfört med oavgjort 1.32 gånger (95% konfidensintervall ges av (1.08, 1.62)).
- En ökning med 1 av *topp25*, skillnaden i antalet spelare som återfinns i poängligans topp 25 mellan hemmalaget och bortalaget, ökar det skattade oddset för hemmaseger jämfört med oavgjort 1.10 gånger (1.01, 1.19), samt ökar det skattade oddset för hemmaseger jämfört med bortaseger 1.12 gånger (1.02, 1.24). Det ska emellertid nämnas att linjäriteten i *topp25* är tveksam, se avsnitt 4.1.5.
- För variabeln *sista2_mål* gäller följande: om hemmalaget gör de sista två målen före andra periodens slut jämfört med om inget av lagen ensamt står

för dessa båda mål så ökar det skattade oddset för bortaseger jämfört med oavgjort 2.37 gånger (1.17, 4.79) och vidare så ökar det skattade oddset för bortaseger jämfört med hemmaseger 3.32 gånger (1.44, 7.65); om bortalaget gör de sista två målen före andra periodens slut jämfört med om inget av lagen ensamt står för dessa båda mål så ökar det skattade oddset för hemmaseger jämfört med oavgjort 1.94 gånger (1.00, 3.74); om bortalaget gör de sista två målen före andra periodens slut jämfört med om hemmalaget gör det så ökar det skattade oddset för hemmaseger jämfört med bortaseger 7.04 gånger (2.24, 22.15).

- Om *hemmaborta_kategori2* är till fördel för bortalaget (nivå 0) så ökar det skattade oddset för bortaseger jämfört med oavgjort respektive hemmaseger 1.65 gånger (1.05, 2.59) respektive (1.00, 2.70).

- För *fullsatt* gäller att om arenan är fullsatt till 95% eller mer så ökar det skattade oddset för oavgjort i jämförelse med hemmaseger 1.72 gånger (1.09, 2.73).

6 Diskussion

Det visar sig, på gott och ont, att det inte är helt lätt att på förhand uttala sig om utfallet i jämna ishockeymatcher. Därmed bibehålls mer av den spänning som jämna matcher skapar, men samtidigt har det funnits en, möjligen naiv, förhoppning om att resultatet mer träffsäkert ska kunna förutspås. Datamaterialet till modelleringen upplevs onekligen som tillräckligt stort men kanske skulle ytterligare säsonger vara önskvärt till prediktionen för att mer säkert kunna bedöma modellens prediktionsförmåga. Vidare så finns det faktorer som kan påverka matchutgången men som den här statistiska analysen inte lyckas täcka in. Bland dessa kan nämnas skador på nyckelspelare, utvisningar i kritiska lägen av matcher och säkert även andra yttre omständigheter som exempelvis individuella spelare i exceptionellt bra form.

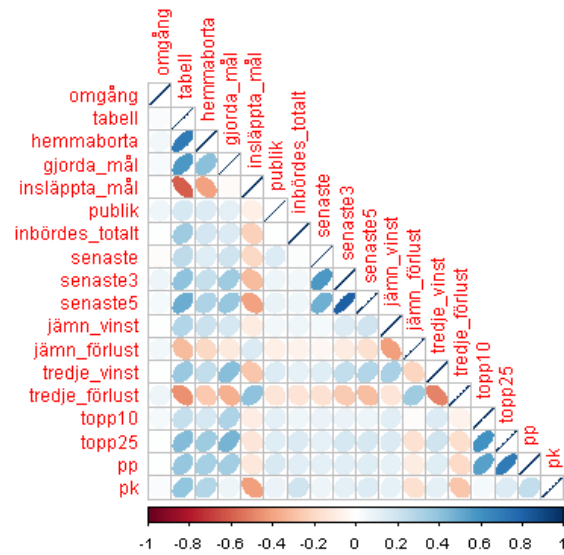
Som analysen visar så är oavgjort det vanligaste resultatet och framför allt den vanligaste klassificeringen/prediktionen som modellerna gör, givet att utfallet med högst skattad sannolikhet väljs. Huruvida detta är ett lämpligt tillvägagångssätt för att klassificera/prediktera utfall är inte självklart. De kriterieändringar som genomfördes innebar i många fall förbättrade resultat. Genom att studera kontingenstabellerna i Appendix A.2 kan ses att det i vissa lägen finns tendenser till större hemma- respektive bortafördelar, men att dessa inte är statistiskt signifikanta och därför inte påverkar den klassificering som görs. Kombinationen av att resultatet ofta är oavgjort efter tre perioder och att ishockeyns oförutsägbara natur medför att troliga hemma- eller bortasegrar inte alls slutar med det, verkar tillsammans bidra till den ofta högt skattade sannolikheten för ett oavgjort resultat.

Vad gäller de fyra variablerna pp , pk , $topp10$ och $topp25$ beskrevs en del problem rörande dessa i avsnitt 3.2. Det hade givetvis ökat trovärdigheten om variablerna varit mer dynamiska och återspeglat aktuella värden för varje enskild observation, i synnerhet vad gäller de två som ingår i den primära modellen. Att pk och $topp25$ faktiskt ingår i den primära modellen visar emellertid på att de innehåller viss förklaringsförmåga för utfallet. Då är det kanske också bättre att detta uppmärksammas än alternativet, att variablerna överhuvudtaget inte hade ingått i analysen. Problemet som uppstår vid prediktion av säsongen 2015/2016, då sannolikheter bestäms utifrån information som är inhämtad efteråt, kan endast kommas runt med tolkningen att värdena också är representativa som genomsnittliga värden sett över hela säsongen. Det är inte orimligt att det också förhåller sig ungefär så, åtminstone från och med ett par omgångar in i varje säsong. Att använda modellen för prediktion av framtida matcher innebär däremot inga problem vad beträffar möjligheten till aktuell information för berörda variabler. Problemet då gäller istället modellens svaga prediktionsförmåga.

Det mest överraskande resultatet av analysen berör variabeln *sista2_mål* och hur pass mycket det påverkar utgången om något lag har lyckats hämta in ett större underläge. Det här kan tolkas som att upphämtningen tagit för mycket på lagets krafter. Mer intuitivt hade kanske varit att tänka sig denna upphämtning som en mental fördel gentemot motståndarna; vad som inom ishockeyn, något märkligt, kommit att benämnas för momentum. Även kontingenstabellen för variabeln *sista_mål* (Appendix A.2, Tabell 19) indikerar, om än inte signifikant, att det verkar ge en fördel att ha släppt in det sista målet före andra periodens slut. Här skulle det vara intressant att också titta mer detaljerat på hur målen utfallit, exempelvis om andra perioden varit mållös eller om något lag har hämtat in tre eller kanske fyra måls underläge. Ett annat intressant resultat är att matcher då arenan är i det närmaste fullsatt inte ger den hemmafördel man kan tänka sig, utan istället talar för ett oavgjort resultat. Detta kan kanske förklaras som att hemmalagen blir mer benägna att inte förlora matchen till förmån för att vinna den.

A Appendix

A.1 Samband mellan variabler



Figur 11: Korrelation (Pearson) mellan kvantitativa variabler.

Tabell 13: Cramérs V mellan kategorivariabler.

| | <i>säsong</i> | <i>tabell_kategori</i> | <i>hemmaborta_kategori</i> | <i>fullsatt</i> | <i>första_mål</i> | <i>sista_mål</i> | <i>sista2_mål</i> | <i>inbördes_senaste</i> |
|----------------------------|---------------|------------------------|----------------------------|-----------------|-------------------|------------------|-------------------|-------------------------|
| <i>säsong</i> | 1 | | | | | | | |
| <i>tabell_kategori</i> | 0.12 | 1 | | | | | | |
| <i>hemmaborta_kategori</i> | 0.12 | 0.48 | 1 | | | | | |
| <i>fullsatt</i> | 0.16 | 0.11 | 0.08 | 1 | | | | |
| <i>första_mål</i> | 0.15 | 0.05 | 0.04 | 0.08 | 1 | | | |
| <i>sista_mål</i> | 0.15 | 0.04 | 0.04 | 0.06 | 0.82 | 1 | | |
| <i>sista2_mål</i> | 0.14 | 0.04 | 0.05 | 0.05 | 0.24 | 0.31 | 1 | |
| <i>inbördes_senaste</i> | 0.15 | 0.17 | 0.10 | 0.05 | 0.03 | 0.03 | 0.05 | 1 |

A.2 Tabeller över enskilda kategorivariabler

Tabell 14: Kontingenstabell för *resultat* och *tabell_kategori*.

| | | <i>resultat</i> | | | Tot |
|------------------------|----------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>tabell_kategori</i> | 0 | 91 | 52 | 52 | 195 |
| | 1 | 131 | 114 | 72 | 317 |
| | 2 | 74 | 59 | 39 | 172 |
| Tot | | 296 | 225 | 163 | 684 |

Tabell 15: Kontingenstabell för *resultat* och *hemmaborta_kategori*.

| | | <i>resultat</i> | | | Tot |
|----------------------------|----------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>hemmaborta_kategori</i> | 0 | 71 | 45 | 54 | 170 |
| | 1 | 155 | 126 | 71 | 352 |
| | 2 | 70 | 54 | 38 | 166 |
| Tot | | 296 | 225 | 163 | 684 |

Tabell 16: Kontingenstabell för *resultat* och *fullsatt* med 90% som gräns.

| | | <i>resultat</i> | | | Tot |
|-----------------|------------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>fullsatt</i> | 0 | 186 | 159 | 115 | 460 |
| | 1 | 110 | 66 | 48 | 224 |
| | Tot | 296 | 225 | 163 | 684 |

Tabell 17: Kontingenstabell för *resultat* och *fullsatt* med 95% som gräns.

| | | <i>resultat</i> | | | Tot |
|-----------------|------------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>fullsatt</i> | 0 | 228 | 189 | 135 | 552 |
| | 1 | 68 | 36 | 28 | 132 |
| | Tot | 296 | 225 | 163 | 684 |

Tabell 18: Kontingenstabell för *resultat* och *första_mål*.

| | | <i>resultat</i> | | | Tot |
|-------------------|----------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>första_mål</i> | 0 | 46 | 34 | 24 | 104 |
| | 1 | 130 | 101 | 69 | 300 |
| | 2 | 120 | 90 | 70 | 280 |
| Tot | | 296 | 225 | 163 | 684 |

Tabell 19: Kontingenstabell för *resultat* och *sista_mål*.

| | | <i>resultat</i> | | | Tot |
|------------------|----------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>sista_mål</i> | 0 | 46 | 34 | 24 | 104 |
| | 1 | 116 | 86 | 84 | 286 |
| | 2 | 134 | 105 | 55 | 294 |
| Tot | | 296 | 225 | 163 | 684 |

Tabell 20: Kontingenstabell för *resultat* och *sista2_mål*.

| | | <i>resultat</i> | | | Tot |
|-------------------|----------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>sista2_mål</i> | 0 | 262 | 193 | 136 | 591 |
| | 1 | 16 | 9 | 19 | 44 |
| | 2 | 18 | 23 | 8 | 49 |
| Tot | | 296 | 225 | 163 | 684 |

Tabell 21: Kontingenstabell för *resultat* och *inbördes_senaste*.

| | | <i>resultat</i> | | | Tot |
|-------------------------|----------|-----------------|----------|----------|------------|
| | | 0 | 1 | 2 | |
| <i>inbördes_senaste</i> | 0 | 63 | 60 | 34 | 157 |
| | 1 | 90 | 67 | 56 | 213 |
| | 2 | 143 | 98 | 73 | 314 |
| Tot | | 296 | 225 | 163 | 684 |

A.3 Statistisk modellering

A.3.1 Steg 1

Tabell 22: Likelihood Ratio-test från steg 1

| Variabel | p-värde |
|-----------------------------|----------|
| <i>pk</i> | 0.006353 |
| <i>topp25</i> | 0.008172 |
| <i>sista2_mål</i> | 0.01032 |
| <i>hemmaborta_kategori2</i> | 0.01289 |
| <i>insläppta_mål</i> | 0.04346 |
| <i>hemmaborta_kategori</i> | 0.05039 |
| <i>sista_mål</i> | 0.05556 |
| <i>fullsatt</i> | 0.1003 |
| <i>hemmaborta</i> | 0.1143 |
| <i>topp10</i> | 0.1845 |
| <i>pp</i> | 0.2065 |
| <i>tabell</i> | 0.2598 |
| <i>tabell_kategori</i> | 0.2739 |
| <i>senaste5</i> | 0.285 |
| <i>tredje_förlust</i> | 0.3037 |
| <i>senaste3</i> | 0.4227 |
| <i>inbördes_senaste</i> | 0.4975 |
| <i>gjorda_mål</i> | 0.5083 |
| <i>jämn_vinst</i> | 0.5092 |
| <i>tredje_vinst</i> | 0.6706 |
| <i>publik</i> | 0.6936 |
| <i>inbördes_totalt</i> | 0.759 |
| <i>jämn_förlust</i> | 0.7616 |
| <i>senaste</i> | 0.8612 |
| <i>omgång</i> | 0.8762 |
| <i>första_mål</i> | 0.9818 |

A.3.2 Steg 2-6

Tabell 23: Modeller från steg 2, 4 och 5 med tillhörande AIC-värden och klassificeringsförmåga.

| Modell | AIC | Klass.förmåga |
|---------|----------------|---------------|
| MODELL0 | 1456.65 | 0.4634 |
| MODELL1 | 1447.71 | 0.4678 |
| MODELL2 | 1451.64 | 0.4488 |
| MODELL3 | 1450.62 | 0.4591 |
| MODELL4 | 1447.85 | 0.4722 |
| MODELL5 | 1446.97 | 0.4664 |

Tabell 24: LR-test för olika indelningar av variabeln *fillsatt*. I MODELL1 avses ett exkluderande av variabeln och i MODELL2 respektive MODELL3 avses ett inkluderande.

| Modell | 85% | 90% | 95% |
|------------------|-------|-------|-------|
| Enkel regression | 0.630 | 0.100 | 0.101 |
| MODELL1 | 0.475 | 0.072 | 0.054 |
| MODELL2 | 0.511 | 0.075 | 0.053 |
| MODELL3 | 0.473 | 0.066 | 0.049 |

Tabell 25: Urval av samspelseffekter från steg 6 med tillhörande värden.

| Modell | Samspel | p-värde | AIC | Klass.förmåga |
|---------|--------------------------------------|---------|----------------|---------------|
| MODELL1 | <i>pk:sista2_mål</i> | 0.00087 | 1436.35 | 0.4708 |
| MODELL1 | <i>pk:fullsatt</i> | 0.071 | 1445.84 | 0.4751 |
| MODELL1 | <i>sista2_mål:inläppta_mål</i> | 0.0093 | 1441.67 | 0.4722 |
| MODELL1 | <i>sista2_mål:fullsatt</i> | 0.053 | 1445.76 | 0.4591 |
| MODELL1 | <i>hemmaborta_kategori2:fullsatt</i> | 0.087 | 1446.23 | 0.4649 |
| MODELL2 | <i>pk:sista2_mål</i> | 0.00075 | 1440.53 | 0.4488 |
| MODELL3 | <i>sista2_mål:inläppta_mål</i> | 0.014 | 1444.03 | 0.4591 |
| MODELL3 | <i>sista2_mål:fullsatt</i> | 0.058 | 1447.43 | 0.4444 |
| MODELL3 | <i>hemmaborta_kategori2:fullsatt</i> | 0.077 | 1447.47 | 0.4430 |
| MODELL5 | <i>pk:sista2_mål</i> | 0.00082 | 1436.06 | 0.4708 |
| MODELL5 | <i>pk:fullsatt</i> | 0.066 | 1445.54 | 0.4751 |
| MODELL5 | <i>sista2_mål:inläppta_mål</i> | 0.0098 | 1441.65 | 0.4708 |
| MODELL5 | <i>sista2_mål:fullsatt</i> | 0.056 | 1445.74 | 0.4693 |
| MODELL5 | <i>hemmaborta_kategori2:fullsatt</i> | 0.086 | 1446.06 | 0.4576 |

A.3.3 Steg 7 - Hosmer-Lemeshows test

Tabell 26: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL1 1 Vs. X.

| Sannolikhet | #1 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.187,0.298] | 16 (14) | 37 (39) |
| (0.298,0.348] | 16 (17) | 36 (35) |
| (0.348,0.378] | 18 (19) | 34 (33) |
| (0.378,0.404] | 18 (20) | 34 (32) |
| (0.404,0.423] | 17 (22) | 35 (30) |
| (0.423,0.456] | 25 (23) | 27 (29) |
| (0.456,0.482] | 27 (24) | 25 (28) |
| (0.482,0.517] | 29 (26) | 23 (26) |
| (0.517,0.564] | 27 (28) | 25 (24) |
| (0.564,0.762] | 32 (32) | 20 (20) |

Tabell 27: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL1 2 Vs. X.

| Sannolikhet | #2 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.11,0.247] | 9 (10) | 37 (36) |
| (0.247,0.277] | 18 (12) | 28 (34) |
| (0.277,0.301] | 12 (13) | 34 (33) |
| (0.301,0.317] | 15 (14) | 31 (32) |
| (0.317,0.336] | 11 (15) | 35 (31) |
| (0.336,0.366] | 13 (16) | 32 (29) |
| (0.366,0.391] | 18 (17) | 28 (29) |
| (0.391,0.432] | 20 (19) | 26 (27) |
| (0.432,0.485] | 22 (21) | 24 (25) |
| (0.485,0.679] | 25 (25) | 21 (21) |

Tabell 28: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL2 1 Vs. X.

| Sannolikhet | #1 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.201,0.334] | 14 (15) | 39 (38) |
| (0.334,0.365] | 16 (19) | 37 (34) |
| (0.365,0.385] | 19 (19) | 32 (32) |
| (0.385,0.406] | 20 (21) | 32 (31) |
| (0.406,0.426] | 22 (22) | 30 (30) |
| (0.426,0.446] | 29 (23) | 23 (29) |
| (0.446,0.473] | 27 (24) | 25 (28) |
| (0.473,0.504] | 25 (26) | 28 (27) |
| (0.504,0.541] | 22 (27) | 29 (24) |
| (0.541,0.686] | 31 (31) | 21 (21) |

Tabell 29: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL2 2 Vs. X.

| Sannolikhet | #2 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.273,0.309] | 18 (14) | 28 (32) |
| (0.309,0.318] | 15 (15) | 32 (32) |
| (0.318,0.329] | 10 (15) | 36 (31) |
| (0.329,0.335] | 13 (15) | 33 (31) |
| (0.335,0.342] | 13 (15) | 32 (30) |
| (0.342,0.349] | 16 (16) | 29 (29) |
| (0.349,0.357] | 17 (16) | 29 (30) |
| (0.357,0.366] | 20 (17) | 26 (29) |
| (0.366,0.394] | 19 (18) | 28 (29) |
| (0.394,0.593] | 22 (23) | 23 (22) |

Tabell 30: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL3 1 Vs. X.

| Sannolikhet | #1 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.203,0.292] | 15 (14) | 38 (39) |
| (0.292,0.346] | 18 (17) | 34 (35) |
| (0.346,0.38] | 16 (19) | 36 (33) |
| (0.38,0.406] | 17 (20) | 35 (32) |
| (0.406,0.428] | 20 (22) | 33 (31) |
| (0.428,0.455] | 24 (23) | 27 (28) |
| (0.455,0.482] | 30 (24) | 22 (28) |
| (0.482,0.515] | 28 (26) | 24 (26) |
| (0.515,0.559] | 26 (28) | 26 (24) |
| (0.559,0.756] | 31 (32) | 21 (20) |

Tabell 31: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL3 2 Vs. X.

| Sannolikhet | #2 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.224,0.256] | 14 (11) | 32 (35) |
| (0.256,0.294] | 13 (13) | 33 (33) |
| (0.294,0.312] | 14 (14) | 32 (32) |
| (0.312,0.323] | 15 (15) | 31 (31) |
| (0.323,0.333] | 11 (15) | 35 (31) |
| (0.333,0.347] | 15 (15) | 30 (30) |
| (0.347,0.378] | 15 (17) | 31 (29) |
| (0.378,0.429] | 20 (19) | 26 (27) |
| (0.429,0.455] | 22 (20) | 24 (26) |
| (0.455,0.662] | 24 (24) | 22 (22) |

Tabell 32: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL5 1 Vs. X.

| Sannolikhet | #1 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.172,0.3] | 14 (14) | 39 (39) |
| (0.3,0.345] | 18 (17) | 34 (35) |
| (0.345,0.378] | 16 (19) | 36 (33) |
| (0.378,0.402] | 20 (20) | 32 (32) |
| (0.402,0.428] | 16 (22) | 36 (30) |
| (0.428,0.453] | 24 (23) | 28 (29) |
| (0.453,0.482] | 29 (24) | 23 (28) |
| (0.482,0.516] | 30 (26) | 22 (26) |
| (0.516,0.562] | 28 (28) | 24 (24) |
| (0.562,0.769] | 30 (33) | 22 (19) |

Tabell 33: Observerade respektive förväntade antal utfall vid Hosmer-Lemeshow-test för MODELL5 2 Vs. X.

| Sannolikhet | #2 obs (förv) | #X obs (förv) |
|---------------|---------------|---------------|
| [0.108,0.249] | 8 (10) | 38 (36) |
| (0.249,0.277] | 19 (12) | 27 (34) |
| (0.277,0.3] | 13 (13) | 33 (33) |
| (0.3,0.317] | 13 (14) | 33 (32) |
| (0.317,0.339] | 12 (15) | 34 (31) |
| (0.339,0.366] | 14 (16) | 31 (29) |
| (0.366,0.391] | 18 (17) | 28 (29) |
| (0.391,0.432] | 20 (19) | 26 (27) |
| (0.432,0.482] | 21 (21) | 25 (25) |
| (0.482,0.687] | 25 (25) | 21 (21) |

A.3.4 Steg 7 - Klassificeringstabeller

Tabell 34: Klassificeringstabell för MODELL1.

| | | Resultat | | | Tot |
|----------------|-----|----------|-----|-----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 75 | 58 | 28 | 161 |
| | X | 143 | 225 | 113 | 481 |
| | 2 | 7 | 13 | 22 | 42 |
| | Tot | 225 | 296 | 163 | 684 |

Tabell 35: Klassificeringstabell för MODELL2.

| | | Resultat | | | Tot |
|----------------|-----|----------|-----|-----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 58 | 49 | 18 | 125 |
| | X | 159 | 232 | 128 | 519 |
| | 2 | 8 | 15 | 17 | 40 |
| | Tot | 225 | 296 | 163 | 684 |

Tabell 36: Klassificeringstabell för MODELL3.

| | | Resultat | | | Tot |
|----------------|-----|----------|-----|-----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 72 | 60 | 27 | 159 |
| | X | 145 | 225 | 120 | 490 |
| | 2 | 8 | 11 | 16 | 35 |
| | Tot | 225 | 296 | 163 | 684 |

Tabell 37: Klassificeringstabell för MODELL5.

| | | Resultat | | | Tot |
|----------------|-----|----------|-----|-----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 73 | 60 | 30 | 163 |
| | X | 146 | 223 | 110 | 479 |
| | 2 | 6 | 13 | 23 | 42 |
| | Tot | 225 | 296 | 163 | 684 |

A.4 Prediktion

A.4.1 Resultat då prediktion sker på utfallet som ges högst sannolikhet

Tabell 38: Prediktionsresultat för MODELL1.

| | | Resultat | | | Tot |
|----------------|-----|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 7 | 8 | 5 | 20 |
| | X | 25 | 24 | 10 | 59 |
| | 2 | 4 | 3 | 1 | 8 |
| | Tot | 36 | 35 | 16 | 87 |

Tabell 39: Prediktionsresultat för MODELL2.

| | | Resultat | | | Tot |
|----------------|-----|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 3 | 6 | 4 | 13 |
| | X | 30 | 26 | 11 | 67 |
| | 2 | 3 | 3 | 1 | 7 |
| | Tot | 36 | 35 | 16 | 87 |

Tabell 40: Prediktionsresultat för MODELL3.

| | | Resultat | | | Tot |
|----------------|-----|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 8 | 7 | 5 | 20 |
| | X | 26 | 26 | 10 | 62 |
| | 2 | 2 | 2 | 1 | 5 |
| | Tot | 36 | 35 | 16 | 87 |

Tabell 41: Prediktionsresultat för MODELL4.

| | | Resultat | | | Tot |
|----------------|-----|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 7 | 7 | 5 | 19 |
| | X | 25 | 25 | 10 | 60 |
| | 2 | 4 | 3 | 1 | 8 |
| | Tot | 36 | 35 | 16 | 87 |

Tabell 42: Prediktionsresultat för MODELL5.

| | | Resultat | | | Tot |
|----------------|-----|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 7 | 8 | 5 | 20 |
| | X | 25 | 24 | 10 | 59 |
| | 2 | 4 | 3 | 1 | 8 |
| | Tot | 36 | 35 | 16 | 87 |

A.4.2 Resultat då prediktion sker enligt andra kriterier

Tabell 43: Nya prediktionsresultat för MODELL1.

| | | Resultat | | | Tot |
|----------------|---|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 11 | 12 | 6 | 29 |
| | X | 20 | 18 | 6 | 44 |
| | 2 | 5 | 5 | 4 | 14 |
| Tot | | 36 | 35 | 16 | 87 |

Tabell 44: Nya prediktionsresultat för MODELL2.

| | | Resultat | | | Tot |
|----------------|---|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 16 | 17 | 11 | 44 |
| | X | 16 | 11 | 4 | 31 |
| | 2 | 4 | 7 | 1 | 12 |
| Tot | | 36 | 35 | 16 | 87 |

Tabell 45: Nya prediktionsresultat för MODELL3.

| | | Resultat | | | Tot |
|----------------|---|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 13 | 12 | 5 | 30 |
| | X | 19 | 18 | 5 | 42 |
| | 2 | 4 | 5 | 6 | 15 |
| Tot | | 36 | 35 | 16 | 87 |

Tabell 46: Nya prediktionsresultat för MODELL4.

| | | Resultat | | | Tot |
|----------------|---|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 10 | 11 | 8 | 29 |
| | X | 21 | 18 | 5 | 44 |
| | 2 | 5 | 6 | 3 | 14 |
| Tot | | 36 | 35 | 16 | 87 |

Tabell 47: Nya prediktionsresultat för MODELL5.

| | | Resultat | | | Tot |
|----------------|---|----------|----|----|-----|
| | | 1 | X | 2 | |
| Klassificering | 1 | 11 | 13 | 5 | 29 |
| | X | 20 | 17 | 6 | 43 |
| | 2 | 5 | 5 | 5 | 15 |
| Tot | | 36 | 35 | 16 | 87 |

Referenser

- [1] AGRESTI, ALAN. 2012. *Categorical Data Analysis*. 3. uppl. Hoboken, N.J.: Wiley Series in Probability and Statistics.

- [2] CRAMÉR, HARALD. 1946. *Mathematical Methods of Statistics*. Princeton : Princeton University Press.

- [3] HELD, LEONHARD OCH SABANÉS BOVÉ, DANIEL. 2014. *Applied Statistical Inference: Likelihood and Bayes*. Berlin, Heidelberg: Springer.

- [4] HOSMER, DAVID W., LEMESHOW, STANLEY OCH STURDIVANT, RODNEY X. 2013. *Applied Logistics Regression*. 3. uppl. Hoboken, N.J.: Wiley Series in Probability and Statistics.

- [5] SVENSKA ISHOCKEYFÖRBUNDET. Official statistics.
<http://stats.swehockey.se>