

Förstagångsåkare i Vasaloppet – sannolikhet att återkomma!



Examinator:
Kenneth Carling

Handledare:
Lars Ronnegörd



Författare:
Khabat Abdulla Amin

Innehållsförteckning

FÖRORD	sid. 3
1. INLEDNING	sid. 4
1.1 <i>Historisk bakgrund</i>	sid. 4
1.2 <i>Problemområde</i>	sid. 5
1.3 <i>Syfte</i>	sid. 6
1.4 <i>Definition.....</i>	sid. 6
2. MATERIAL OCH METODER	sid. 7
2.1 <i>Vasaloppets tävlingstyper</i>	sid. 7
2.2 <i>Materialbeskrivning.....</i>	sid. 8
2.3 <i>Statistiska metoder och modeller</i>	sid. 13
3. RESULTAT	sid. 15
3.1 <i>Resultat från analyser av sannolikhet att återkomma bland förstagångsåkare</i>	sid. 15
3.2 <i>Illustration av resultat från själva Vasaloppet.....</i>	sid. 17
3.3 <i>Jämförelse mellan tävlingstyper.....</i>	sid. 19
4. DISKUSSION	sid. 20
5. SAMMANFATTNING	sid. 22
APPENDIX 1	sid. 23
APPENDIX 2	sid. 24
REFERENSER	sid. 25

Förord

Till att börja med vill jag rikta ett stort tack till min handledare Lars Ronnegård, för tillfället lärare i Högskolan Dalarna. Han har kritiserat, stöttat och uppmuntrat mig under arbetets gång med uppsatsen. Jag vill även tacka övriga medarbetare inom Högskolan Dalarna och speciellt professor Kenneth Carling som gett mig möjligheten och varit min examinator, samt hjälpt mig på vägen mot mitt uppsatta mål; magisterexamen i statistik.

Ett varmt tack till Joakim Quaresmini, projekt ledaren inom UNAR(*utomnordiska akademiska resurser*) vid Högskolan Dalarna som under arbetets slutfas gett mig mycket värdefulla och inspirerande synpunkter om uppsatsen.

Till sist men inte minst, ett stort tack till alla som jobbar inom Vasaloppets evenemang särskilt Mats Budh och Monica Eriksson som har ställt sin tid och sina synpunkter till förfogande.

Datum

Khabat Abdulla Amin

1. INLEDNING

1.1 Historisk bakgrund

Vasaloppet – det är få ord i svenska språket som väcker så många känslor och minnen till liv. Både historiska och sportsliga. Vasaloppet är så mycket mer än en skidtävling. Det handlar också om tradition, historia, kultur och frihet.

Vasaloppet har sitt ursprung från 1500-talet, då Sverige befann sig i en påtvingad union med grannlandet Danmark. Missnöjet växte mot den danska kungen, Kristian II. En av de främsta frihetskämparna Gustav Eriksson Vasa, en ung adelsman från Uppsala, bestämde sig år 1520 för att uppmana dalfolket att ta till vapen. Dalfolket var dock betänksamma och Gustav Vasa gav sig ensam i väg mot Norges gräns från Mora i Dalarna. Några dagar senare nåddes Mora av nyheter om kung Kristian II framfart i riket och två Mora karlar bestämde sig för att försöka hinna ifatt Vasa. Ortens två främsta skidåkare, Lars och Engelbrekt skickades iväg i detta historiens första Vasalopp. De fann Gustav Vasa i Sälenstrakten och övertalade honom att återvända till Mora och att leda dalfolkets kamp mot kung Kristian II. Efter två och ett halvt års krigande är Sverige fritt och Gustav Vasa blir vald till Sveriges konung den 6 juni 1523(Vasaloppet, 2000).

Idén till att ordna skidloppet, som skulle påminna om Gustav Vasas flykt från Mora till Sälen, fick Mora sonen Anders Pers, den 10 feb. 1922, som vid den tiden var chefredaktör för Vestmanlands läns Tidning. Den 19 mars 1922 begav sig 119 skidlöpare (bara män) iväg från Vasastenen i Sälen mot målet beläget vid Gustav Vasas staty i Mora. Detta historiska ögonblick blev början till ett stort nationellt och internationellt evenemang och framförallt en stor folkfest för hela Sverige (Vasaloppet, 2003). 1981 fick de första kvinnorna åka Vasaloppet.

Spårsträckningen har varierat under årens lopp med hänsyn till framkomlighet och markägare. 1922-1924 var spåret 92 km långt. 1936-1985 har spårlängden varierat mellan 87-89 km. Efter 1986 är Vasaloppet 90 km långt och det läggs hela 650 km spår i spårområdet för att alla startande skall ha bästa möjliga framkomlighet (Vasaloppet, 2000) .

1.2 Problemområde

Intresset för skidåkning har sedan slutet på 1980-talet minskat hos det svenska folket. Anledningen till det minskade intresset beror antagligen på de klimatförändringar som under slutet av 1990-talet har resulterat i varmare vintrar, och dels på att många människor, särskilt ungdomar, har visat ett minskat intresset för sport och friluftsliv (Söderlund, 2000).

Utvecklingen för olika idrottsevenemang har varit varierande. Vissa arrangemang har gått framåt, medan andra har stött på stora svårigheter för att överleva på grund av vikande intresse (Bohlin, 1994). Skideevenemanget Vasaloppet har genomgått en anmärkningsvärd bra utveckling. Pga den varma vintern 1990, fick evenemanget ställas in. De allt varmare vintrarna har också påverkat träningsförhållanden. Enligt Henriksson (2003) har deltagarantalet en stor geografisk variation, och det är sannolikt att de milda vintrarna påverkar åkare olika beroende på hemort.

Vasaloppsevenemanget har haft sjunkande antal återkommande deltagare under 1992-2002, vilket har medfört att antalet förstagångsåkare har stigit. Medelåldern bland deltagarna har ökat från 34 år till 41 år under samma period (Budh & Eriksson, 2002). Min roll som statistiker är att undersöka vilka skäl som ligger till grund för problemet.

1.3 Syfte

Syftet med studien är att undersöka sannolikheten att förstagångsåkare i Vasaloppsevenemanget återkommer påföljande år. Vilka faktorer, som registrerats i Vasaloppets databas under åren 1992-2001, påverkar denna sannolikhet? Resultaten kan bidra till nya perspektiv på deltagare och kan ge Vasaloppsevenemang viktiga erfarenheter om hur resultaten kan användas som bas för fortsatt marknadsföring.

1.4 Definition

Två viktiga ord som uppsatsen bygger på definieras här.

Förstagångsåkare: Är en löpare som inte har åkt något slags lopp i Vasaloppsevenemanget tidigare.

Återkommande: Är en löpare som kommer tillbaka till något av loppen i Vasaloppsevenemanget påföljande år.

2. MATERIAL OCH METODER

2.1 Vasaloppets tävlingstyper

I verkligheten finns det inte bara de nedanstående loppen, det finns även andra som också har genomförts de två senaste åren t ex Stafettvasan, Skejtvasan. De finns inte med i denna analys.

Vasaloppet (Sälen-Mora 90 km): Loppet går alltid första söndagen i mars. Årligen deltar ca 15000 löpare. Sedan 1987 körs loppet endast i klassisk teknik och 1997 infördes en officiell damklass (Vasaloppet, 2003). Åldergränsen är 19 år.

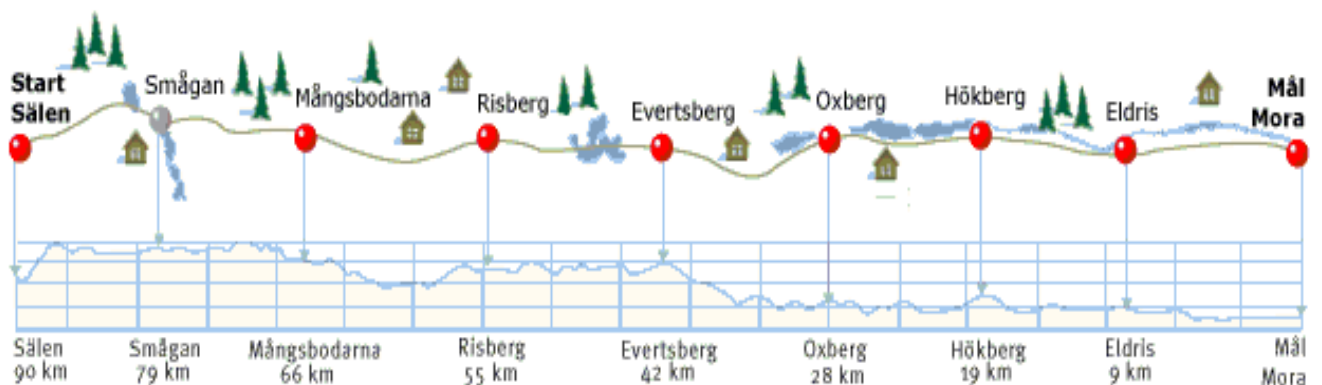
Öppet Spår (Sälen-Mora 90 km): År 1979 genomfördes för första gången detta lopp, där deltagarna åker i vasaloppsspåren, och är lika långt som Vasaloppet. Ca: 12000 skidlöpare deltar i Öppet Spår varje år. Öppet Spår har individuell start med tidtagning och arrangeras under två dagar, söndag och måndag veckan före Vasaloppet (Vasaloppet, 2003). Åldersgränsen är 19 år.

Tjej Vasan (Oxberg-Mora 30 km): Tjejevasan infördes 1988 första gången och är världens största skidtävling för damer, 2000 anmälde sig första gången för loppet. Tjejevasan går på lördagen, en vecka före vasaloppet. Starten sker gruppvis med ca 800 startande åt gången (Vasaloppet, 2003). Åldersgränsen är 17 år.

HalvVasan (Evertsberg-Mora 45 km): Även kallad för 45: an, startades 1997 och var endast för män. Men 1999 fick damerna åka HalvVasan. Loppet är till för alla som tycker att 90 km är för långt och gärna vill uppleva stämningen av Vasaloppet (Vasaloppet, 2003). Åldersgränsen är 17 år.

KortVasan (Oxberg-Mora 30 km): Här kan hela familjen åka tillsammans de 30 km mellan Oxberg och Mora. Även här sker starten gruppvis med c 800 startande i varje grupp (Vasaloppet, 2003). Åldersgränsen är 13 år.

Figur 1: Vasaloppets skidspår.



2.2 Materialbeskrivning:

År 1989 påbörjades insamlingen av information om alla deltagare i Vasaloppet med hjälp av dataprogrammet Access. Accessdatabasen består av tre olika tabeller, Personer, Historik, och Antal lopp (Databas 2002). Första tabellen (personer) innehåller personliga uppgifter om deltagarna, andra tabellen (historik) har endast historiska uppgifter, och den sista tabellen (antal lopp) innehåller antal genomförda lopp t.o.m. 2002 och vilka slags lopp som genomförts. Den ursprungliga tabellen för åren 1992 – 2002 innehöll 372 609 observationer. Eftersom vi inte vet om de som åkte 2002 kommer att återkomma eller ej, så har dessa tagits bort från följande analyser. Enligt Vasaloppsarrangörerna (Budh & Eriksson, 2003) är data av sämre kvalitet och inte fullständigt före 1992 och därför ingår inte åren 1989-1991 i analyserna. Undersökningsperioden innehåller därmed åren 1992 – 2001. För att förenkla analyserna gäller undersökningen bara svenska deltagare (89, 5 % av totala antalet deltagare är svenskar). Efter att observationerna där åkarna saknade antal tidigare lopp eller där de var anmälda utan att starta togs bort var det slutligen 236 887 observationer i tabellen. Dessa 236 887 observationer utgör råmaterialet för studien.

Jag har sammanställt dessa tre ovannämnda tabeller i en Minitab tabell, där alla observationer som saknade uppgift om antal lopp togs bort. Tabellen innehåller 236 887 observationer varav 33% är förstagångsåkare som är 78 558 observationer, och antalet återkommande deltagare påföljande år, av förstagångsåkarna, är 30 026 (Tabell 1).

Tabell 1: Antalet och andel observationer mellan 1992 – 2001

	Antal observationer	Andel observationer	
Ej förstagångs åkare	158 329	67 % av 236 887	
Återkommande	30 026	38 % av 78 558	
Förstagångs åkare	Ej återkommande	48 532	62 % av 78 558
Total	236 887		

Antal löpare som deltagit i hela Vasaloppsveckan har ökat från och med år 1992. Antalet förstagångsåkare har också ökat från 4444 personer 1992 till 8068 personer 2001 (Tabell 2).

Tabell 2: Antal löpare varav andel förstagångsåkare för varje år

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	Total
Alla åkare	14272	16288	21597	21817	24441	26916	28724	27868	29532	25432	236887
Förstagångsåkare	4444	4345	7768	7863	8297	9679	9599	8627	9868	8068	78558
Andel	0,31	0,27	0,36	0,36	0,34	0,36	0,33	0,31	0,33	0,32	

En jämställdhetslag mellan män och kvinnor genomfördes 1979 i Sverige (Ordbok,1992) och påverkade att kvinnor fick delta i Vasaloppet från år 1981. Eftersom intresset från tjejerna att delta i detta evenemang var stort, ökade trycket på deltagarna i det traditionella Vasaloppet, och på grund av detta intresse utvecklade Vasaloppet "Tjejvasan" som ett nytt skidlopp. Bland de återkommande mellan 1992-2001 var 37,28% kvinnor. Antalet förstagångsåkare var som högst 1997 med 9679 deltagare (Tabell 2).

Alla förstagångsåkare (78 558 observationer) under Vasaloppsveckan, från 1992 till och med 2001 som är undersökningsperiod, är i genomsnitt 4,5 år yngre än övriga åkare. Medelåldern på förstagångsåkare är 36,5 år (standard fel = 0,04) medan medelåldern för alla åkare (236 887 observationer) är 41 år (standard fel = 0,03).

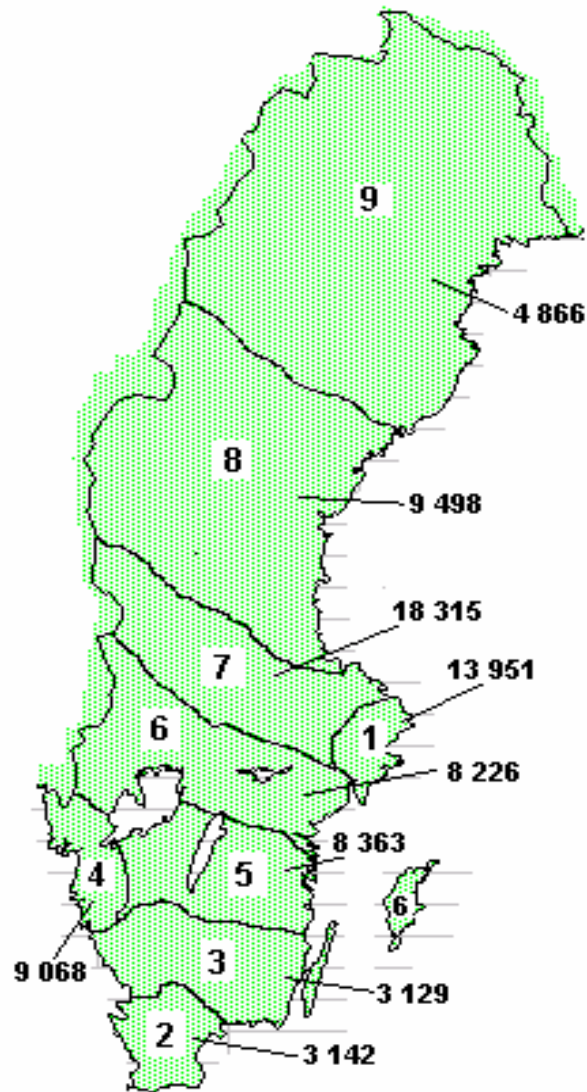
Trots att Vasaloppet är den mest kända tävlingen har Öppet Spår högsta antalet förstagångsåkare. Möjligen kan detta bero på att denna tävlingstyp ligger under två dagar samt flera starttillfällen. Trots att denna tävlingstyp inte är någon direkt tävling, som Vasaloppet ligger den genomsnittliga tiden för utfört lopp väldigt nära tiden för Vasaloppets tävlande. Antalet som bryter Vasaloppet och Öppet Spår är betydligt högre än antalet som bryter i de andra lopp. Sannolikt beror detta på loppens längd och antalet åkare (Tabell 3).

Tabell 3: Beskrivning av förstagångsåkare uppdelat på tävlingstyp.

Variabler\Tävlingstyp	Vasaloppet	Öppetspår	Tjejvasan	Halvvasan	Kortvasan
Antal Obs.	23320	24925	23241	2881	4191
Andel Kvinnor	8,2 %	10,3 %	100 %	11,1 %	48,6 %
medel	32,9	36,9	37,8	45,8	40,5
Ålder					
min	18(19)	18(19)	16(17)	17(17)	8(13)
max	76	81	84	85	82
medel	8T 57M 30S	8T 37M 17S	3T 20M 55S	4T 20M 59S	3T 26M 10S
Tid					
min	3T 46M 35S	2T 8M 17S	1T 25M 46S	2T 5M 45S	1T 28M 57S
max	12T 25M 23S	13T 50M	6T 38M 11S	10T 55M 25S	6T 48M 29S
Andel som bröt	3,9 %	3,8 %	2 %	0,2 %	0,2 %
Andel återkommande	40,2 %	36,4 %	39,4 %	37,4 %	31,7 %

Åldern vid anmälan var lägre än lägsta tillåtna för start vilket innebär att de personerna fyller år under perioden och uppfyller ålderskraven vid start. (siffror inom parantes anger ålderskrav vid start).

Antalet förstagångsåkare är överrepresenterade av deltagare från Dalarna (Figur 2). Detta efter de postnumren som är registrerade i Vasaloppet databas. Vilket kanske i sig har en naturlig förklaring: närheten till Vasaloppet. Därefter kommer Stockholmsregionen. Minsta antalet åkare kommer från norra och södra Sverige, vilket kanske kan förklaras med de långa avstånden, restiden och kostnader för resan.



Figur 2: Postnummerområden (definierad utifrån första siffran i hemadressens postnummer) och antal förstagångsåkare.

Jag har delat in åktiden för respektive lopp i tidsgrupper för att göra analysen lättare och för att kunna ta hänsyn till de som hade brutit. Tidsgrupper är tider avrundade uppåt till närmaste hel timme, förutom första gruppen som är de som har brutit lopp. De grupper som hade färre än 10 observationer har jag lagt ihop med närmaste tidsgrupp (Tabell 4).

Tabell 4: Fördelning för antal förstagångsåkare i varje tidgrupp efter tävlingstyp

Åktid	Vasaloppet	Öppetspår	Tjejvasan	Halvvasan	Kortvasan
Brutit	909	948	469	5**	7**
2 timmar	*	*	330	*	111
3 timmar	*	*	5351	173	1120
4 timmar	*	*	13314	950	2016
5 timmar	384	53	3518	1076	847
6 timmar	984	409	259	521	90
7 timmar	1661	1995	469	138	*
8 timmar	2909	4267	*	18	*
9 timmar	4429	5691	*	*	*
10 timmar	5485	5322	*	*	*
11 timmar	4394	3829	*	*	*
12 timmar	2165	2298	*	*	*
13 timmar	*	113	*	*	*

* Det fanns ingen med de tidsgrupperna.

** Togs bort i den statistiska analysen, eftersom det var för få observationer.

2.3 Statistiska metoder och modeller

Logistik regression analys används för att undersöka förhållandet mellan en binär responsvariabel och vissa förklarande variabler. Detta är den metod som jag har använt i min undersökning. För en binär responsvariabel kan responsen Y kan anta två möjliga värden, (tex i den här undersökningen har jag $Y=1$ om en förstagångsåkare återkommer följande år och $Y=0$ för ej återkommande).

Till exempel om x är en förklarande variabel och $p = \Pr(Y=1 \mid x)$ är responssannolikheten, då har linjer logistik modell formen

$$\log\left(\frac{p}{1-p}\right) = \alpha + \beta x$$

α = är konstant parameter

β = är skattningen för den förklarande parametern

Denna undersökning innehåller fem förklarande variabler, en kontinuerlig variabel och fyra klass variabler. Följande är grundmodellen med alla variabler:

$$\log(p/(1-p)) = \text{konstant} + b_1 * \text{ålder} + \text{kön}_j + \text{år}_k + \text{tid}_l + \text{postnr}_m + b_{21} * (\text{ålder} * \text{tid}) + (\text{kön} * \text{tid})_j$$

p = Sannolikheten för att återkomma

ålder = deltagares ålder, kontinuerlig variabel med regressionskoefficient b_1 .

kön = deltagares kön, och det är klass förklarande variabel ($j = \text{Man, Kvinna}$)

år = Vilket år åkte deltagare, klass variabel ($k = 1992-2001$)

tid = deltagares åktid, klass variabel ($l = \text{brutit, 1, 2, \dots, 13 timmar}$)

postnr = postområdes effekt (definierat utifrån första siffran i post numret),
klass variabel ($m = 1, 2, \dots, 9$)

ålder*tid = samspelseffekten mellan ålder och tid med regressionskoefficient b_{21} .

kön*tid = samspelseffekten mellan kön och tid.

Det finns andra parametrar som inte är med i den grundmodellen till exempel klubbkod och placering. Klubbkod har jag inte med eftersom många observationer saknade denna uppgift. Placeringen finns inte med för att tid och placering har nästan samma betydelse. Jag valde att ha med tid i stället för placering.

Jag har gjort individuella modeller för varje tävlingstyp för sig eftersom varje tävlingstyp har separata data typer. Till exempel så skiljer sig längden från en tävlingstyp till en annan, de infördes i olika år, och det finns kvinnor som åker Tjejvasan.

Skattningen av parametrarna gjordes i Minitab som använder *maximum likelihood* skattning. Maximum likelihood skattningarna är X^2 -fördelade under noll hypotesen (där alla förklarande effekter i modellen är noll) och den tekniken skaffar P -värden vid modellval.

En användbar metod för modellvalet var *stepwise backward elimination*. Med denna metod tar man med alla variabler i en grundmodell sedan tar man bort de förklarande variabler som har högst P -värde i turordning och som inte är signifikant på 20 % nivån. Med denna metod behåller man en variabel då $P > 0,2$ om samspelet där variabeln ingår har $P < 0,2$ (McCullagh & Nelder 1989 i Olsson 2002).

Det finns två olika mått att utvärdera modellens egenskaper som heter *Pearson Chi-2* och *Deviance*. Pearson Chi-2 beräknas ur kvadratsumman av residualerna medan Deviance beräknas från likelihood. Om skillnaden är stor mellan Pearson Chi-2 och Deviance så kan modellen vara dålig. För en exakt lämplig modell ska Pearson Chi-2 och Deviance vara lika med antalet frihetsgrader (*df*). Om kvoten Pearson Chi-2 / *df* eller Deviance / *df* är mycket större än ett, så kallas det problemet ”*overdispersion*”. Overdispersion händer när variansen för responsen är större än den som är förväntad. Till exempel i den här undersökningen har vi binomial fördelning, och förväntade variansen, $V(\mu)$, är en funktion av respons-sannolikhet p : $V(\mu) = np(1 - p)$.

3. RESULTAT

3.1 Resultat från analyser av sannolikhet att återkomma bland förstagångsåkare

I analysen har jag kommit fram till en slutgiltig modell för varje tävlingstyp (Tabell 5). Modellerna skiljer sig något åt mellan tävlingstyperna, speciellt Tjejvasan där endast kvinnor får delta.

Tabell 5: P-värde för förklarande variabel

	Vasaloppet	Öppetspår	Tjejvasan	Halvvasan	Kortvasan
År	<0,001	<0,001	<0,001	<0,001	<0,001
Kön	0,564	0,003	^b	0,803	0,764
Tid	<0,001	<0,001	<0,001	0,011 ^a	0,002 ^a
Ålder	<0,001	0,038	0,021	0,605	0,186
Post nummer	<0,001	0,001	<0,001	-	-
Tid* Kön	0,002	-	^b	0,032 ^a	0,160 ^a
Tid*Ålder	<0,001	0,003	0,009	0,064 ^a	0,024 ^a

^a De som brutit ej med i analysen, för få observationer.

^b Kön har bara en faktor, och det är kvinna.

- Ej med i slutlig modell för $p > 0.2$

Kvoten mellan Pearson Chi2 och antalet frihetsgrader för de olika modellerna i Tabell 6 var mellan 1,0 och 1,1, och för Deviance var kvoten mellan 1,0 och 1,3, vilket tyder på god anpassning för de fem olika modellerna. Detta tyder också på att det troligen inte är problem med overdispersion (Olsson, 2002). I appendix 1 ges ett exempel av hur skattningarna av effekterna (Tabell 6) används till att beräkna sannolikheten för återkommande deltagare.

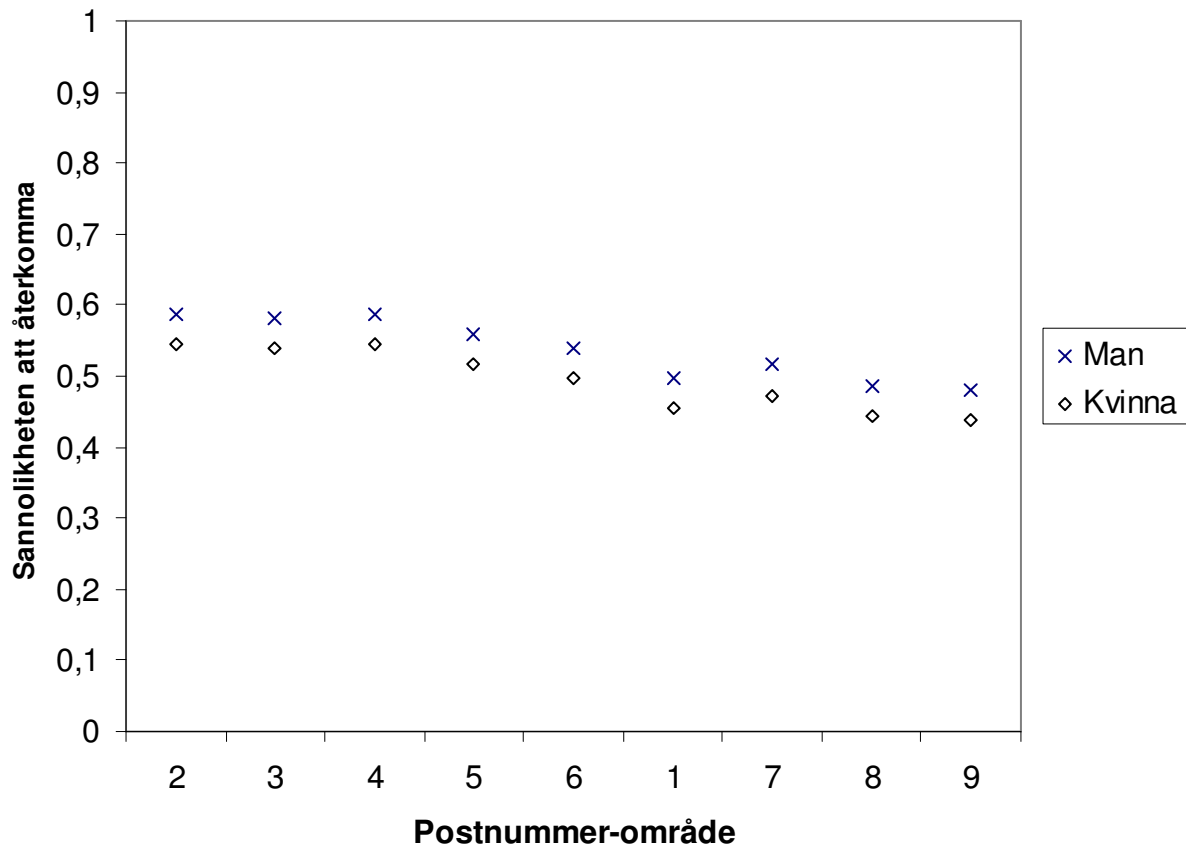
Tabell 6: Skattningar för de förklarande variablerna. Standardfel inom parentes

	Vasaloppet	Öppetspår	Tjevasan	Halvvasan	Kortvasan
Konstant	-0,901 (0,355)	0,257 (0,258)	-0,750 (0,334)	0,653 (1,454)	0,412 (1,047)
År					
1992	0,000	0,000	0,000	* *	* *
1993	-0,471 (0,062)	0,042 (0,095)	0,150 (0,098)	* *	* *
1994	-0,613 (0,061)	-1,389 (0,081)	-1,760 (0,087)	* *	* *
1995	-0,674 (0,063)	-1,047 (0,077)	-1,400 (0,086)	* *	* *
1996	-0,877 (0,061)	-1,068 (0,078)	-1,253 (0,085)	* *	* *
1997	-0,613 (0,063)	-0,965 (0,078)	-1,187 (0,084)	0,000	0,000
1998	-1,135 (0,064)	-1,191 (0,078)	-1,436 (0,086)	-0,510 (0,121)	0,006 (0,112)
1999	-0,473 (0,069)	-1,055 (0,080)	-1,590 (0,086)	-0,281 (0,120)	-0,022 (0,111)
2000	-0,840 (0,063)	-1,375 (0,079)	-1,607 (0,086)	-0,738 (0,125)	-0,953 (0,121)
2001	-0,776 (0,066)	-0,958 (0,082)	-1,316 (0,086)	-0,160 (0,135)	-0,176 (0,124)
Kön					
Kvinna	0,000	0,000	0,000	0,000	0,000
Man	0,150 (0,260)	0,143 (0,048)	* *	-0,359 (1,437)	-0,285 (0,950)
Tid					
Brutit	0,000	0,000	0,000	* *	* *
2 timmar	* *	* *	2,773 (0,541)	* *	0,000
3 timmer	* *	* *	1,374 (0,346)	0,000	-0,788 (1,055)
4 timmer	* *	* *	0,509 (0,340)	-2,687 (1,587)	-1,573 (1,053)
5 timmar	2,575 (1,151)	0,688 (0,929)	0,017 (0,365)	-0,912 (1,475)	-1,667 (1,077)
6 timmar	0,735 (0,610)	-0,056 (0,429)	-0,663 (0,767)	-2,538 (1,511)	-2,387 (1,459)
7 timmar	1,004 (0,427)	-0,082 (0,292)		-1,642 (1,681)	* *
8 timmar	0,123 (0,404)	-0,103 (0,270)	* *	-1,303 (2,855)	* *
9 timmar	-0,215 (0,390)	-0,662 (0,265)	* *	* *	* *
10 timmar	-0,294 (0,379)	-0,643 (0,267)	* *	* *	* *
11 timmar	-0,313 (0,382)	-1,206 (0,278)	* *	* *	* *
12 timmar	-1,252 (0,417)	-1,589 (0,304)	* *	* *	* *
13 timmar	* *	-1,003 (0,795)	* *	* *	* *
Ålder					
1	0,056 (0,008)	0,013 (0,006)	0,021 (0,009)	-0,005 (0,010)	-0,028 (0,021)
Posnr					
2	0,000	0,000	0,000	* *	* *
3	0,362 (0,076)	0,081 (0,068)	-0,103 (0,098)	* *	* *
4	0,342 (0,068)	0,184 (0,078)	-0,050 (0,093)	* *	* *
5	0,365 (0,054)	0,105 (0,051)	0,165 (0,054)	* *	* *
6	0,249 (0,054)	0,037 (0,052)	-0,047 (0,057)	* *	* *
7	0,167 (0,060)	0,079 (0,051)	0,018 (0,054)	* *	* *
8	0,075 (0,050)	0,069 (0,042)	0,053 (0,044)	* *	* *
9	-0,041 (0,055)	-0,082 (0,050)	0,147 (0,052)	* *	* *
9	-0,072 (0,057)	-0,120 (0,071)	-0,208 (0,079)	* *	* *
Kön *tid					
Man* Brutit	0,000	* *	* *	* *	* *
Man* 2 tim,	* *	* *	* *	* *	0,000
Man* 3 tim,	* *	* *	* *	0,000	0,530 (0,960)
Man* 4 tim,	* *	* *	* *	1,902 (1,566)	0,424 (0,955)
Man* 5 tim,	0,022 (1,051)	* *	* *	0,227 (1,452)	0,033 (0,967)
Man* 6 tim,	0,594 (0,539)	* *	* *	-0,029 (1,458)	1,083 (1,121)
Man* 7 tim,	-0,065 (0,336)	* *	* *	-0,244 (1,494)	* *
Man* 8 tim,	0,380 (0,311)	* *	* *	3,154 (2,297)	* *
Man* 9 tim,	0,515 (0,295)	* *	* *	* *	* *
Man* 10 tim,	-0,004 (0,281)	* *	* *	* *	* *
Man* 11 tim,	-0,241 (0,281)	* *	* *	* *	* *
Man* 12 tim,	0,073 (0,301)	* *	* *	* *	* *
Ålder *tid					
Brutit	0,000	0,000	0,000	* *	* *
2 tim,	* *	* *	-0,046 (0,020)	* *	0,000
3 tim,	* *	* *	-0,003 (0,009)	0,000	0,026 (0,021)
4 tim,	* *	* *	0,006 (0,009)	0,017 (0,011)	0,040 (0,021)
5 tim,	-0,092 (0,018)	-0,032 (0,025)	0,005 (0,010)	0,010 (0,012)	0,039 (0,022)
6 tim,	-0,052 (0,011)	0,000 (0,011)	0,007 (0,017)	0,037 (0,013)	0,041 (0,029)
7 tim,	-0,044 (0,010)	0,000 (0,007)	* *	0,019 (0,020)	* *
8 tim,	-0,032 (0,009)	-0,001 (0,007)	* *	-0,030 (0,065)	* *
9 tim,	-0,030 (0,009)	0,009 (0,007)	* *	* *	* *
10 tim,	-0,026 (0,009)	0,002 (0,007)	* *	* *	* *
11 tim,	-0,028 (0,009)	0,013 (0,007)	* *	* *	* *
12 tim,	-0,016 (0,010)	0,016 (0,008)	* *	* *	* *
13 tim	* *	0,007 (0,018)	* *	* *	* *

3.2 Illustration av resultat från själva Vasaloppet

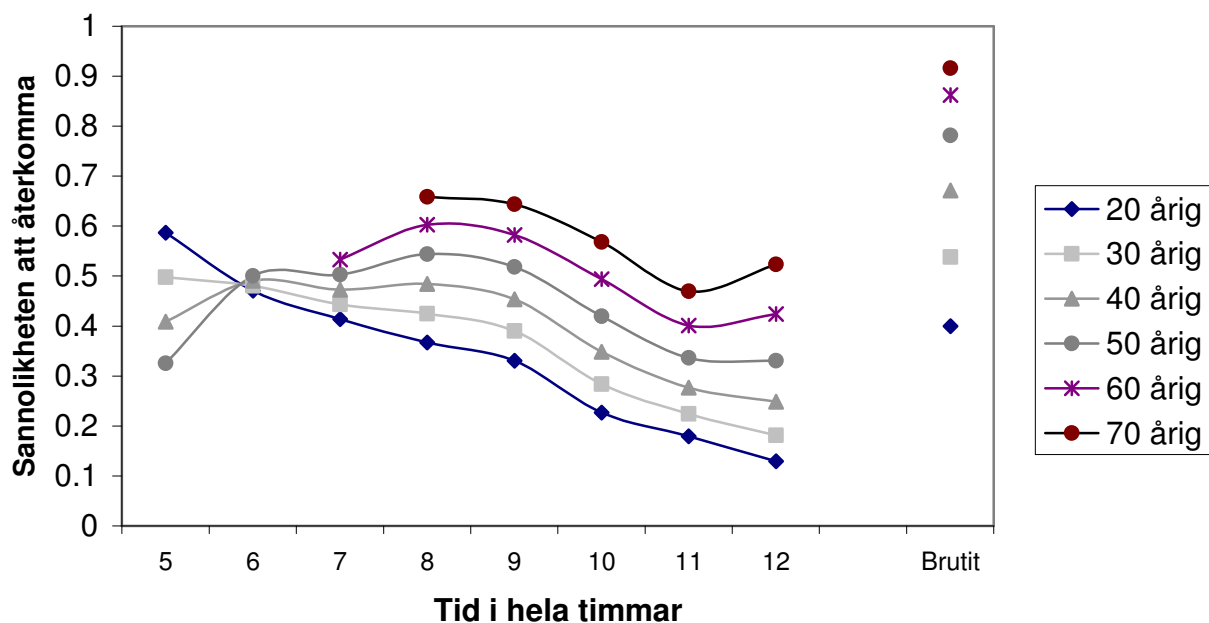
I detta avsnitt illustreras resultaten för själva Vasaloppet i Tabell 6.

En förstagångsåkare har olika sannolikheter att återkomma, beroende på varifrån han kommer (Figur 4). T ex en 30 årig man som åkte på 5 timmar, år 2001, har större sannolikhet att återkomma om han är från Skåne (59 %) än om han kommer från Norrbotten (48 %). Dessutom visar Figur 6 att män har större sannolikhet att återkomma än kvinnor.



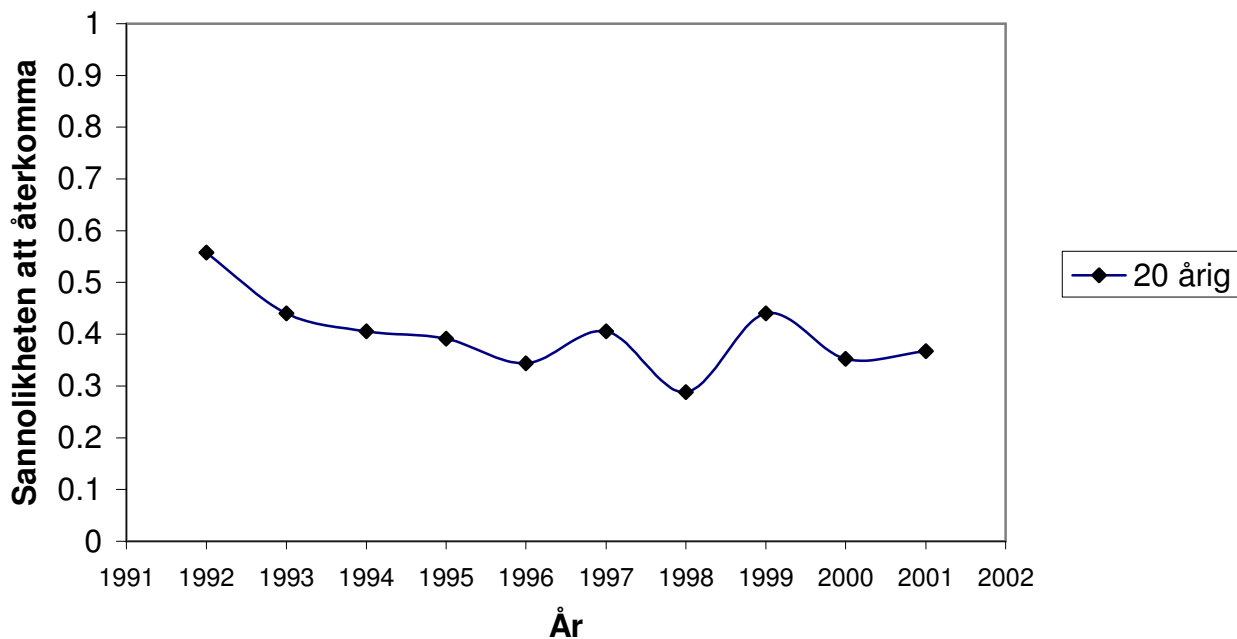
Figur 4: Sannolikheten att återkomma efter postnummer för en 30 åring man/kvinna, som åkte på 5 timmar år 2001. Postnummer-områden är ordnade från söder till norr på x-axeln (se Figur 2).

Analysen visar också att för unga åkare minskar sannolikheten att återkomma tydligt med åktiden (Figur 5). Medan sannolikheten att återkomma bland äldre åkare påverkas mindre av åktiden. Samma analys visar också att sannolikheten att återkomma bland dem som har brutit ökar med åldern. En 70-åring man från Stockholm som hade brutit under loppet 2001 har 91 % sannolikhet att återkomma. Medan motsvarande sannolikhet för en 20-åring bara är 40 %



Figur 5: Sannolikheten att återkomma efter antal åktimmar och ålder för en man från Stockholm, som har åkt eller brutit år 2001.

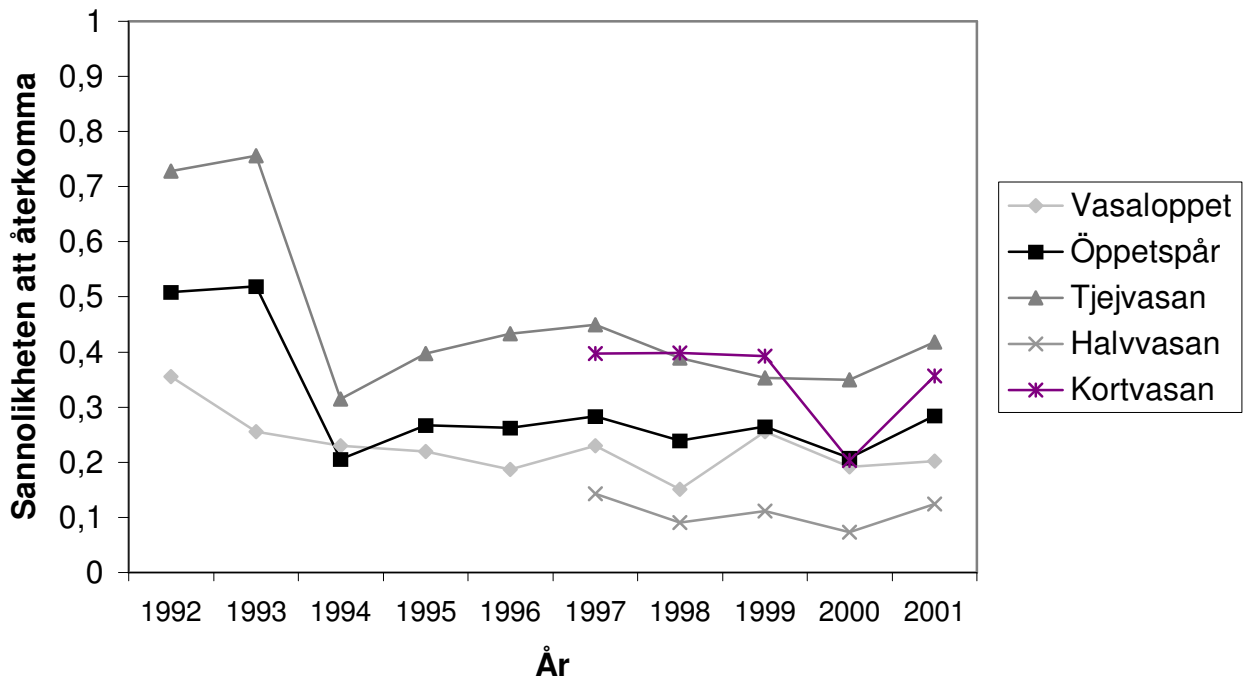
Sannolikheten att återkomma för en 20 årig man från Stockholm som åkte på 8 timmar, har varierat mellan år med en nedåtgående trend 1992 – 1996, och med stora fluktuationer 1997 – 1999 (Figur 6).



Figur 6: Sannolikheten att återkomma efter årtal för en 20 årig man från Stockholm som har åkt på 8 timmar.

3.3 Jämförelse mellan de fem tävlingstyperna

I Figur 7 visas jämförelse mellan de olika tävlingstyperna efter återkommande-sannolikhet. Denna är uppskattad utifrån de skattningarna som jag kom fram till i Tabell 6. Eftersom kvinnan kan åka alla de fem tävlingstyperna, så exempelvis tog jag en 20 åring kvinna från Stockholm som åkte på genomsnittlig åktid för varje tävlingstyp (se tabell 3 sidan 10). Resultaten visar att deltagarna som åker Tjejvasan har större sannolikhet för att återkomma än de som åker de andra tävlingstyperna.



Figur 7: återkommande sannolikhetsjämförelse mellan de olika tävlingstyper.

4.DISKUSSION

Utifrån de kunskaper jag har fått fram i min undersökning och analys, kommer jag att lägga fram några faktorer som jag ser kommer att påverka en förstagångsåkare för att återkomma följande år. Dessa faktorer har jag kommit fram till genom Vasaloppets databas, säkert finns det andra faktorer också, som kan påverka deltagare att återkomma. Men i så fall måste man göra en bredare analys som kräver mer tid och stöd.

Service är det viktigaste som påverkar kunder för att återkomma till ett företag. Dagens informationssamhälle med sin kunskap och teknologi kräver av företagen och evenemangen att anpassa sig mer efter kundernas behov och önskemål.

En viktig faktor som kan påverka en förstagångsåkare att återkomma är klimatförändringarna mot varmare vintrar. Vädret var orsak till att Vasaloppet ställdes in år 1990. Naturligtvis påverkar väderförhållandet om en förstagångsåkare skall återkomma. Ett lopp med lite snö och för mycket värme bäddar inte för att ett stort antal deltagare återkommer. Sannolikheten för en förstagångsåkare att återkomma har minskat från 1992 och framåt (*Figur 6*). Denna nedgång har en stor föränderlighet mellan 1997 – 1999. 1997 var det dåligt väder då Vasaloppet genomfördes och då minskade antalet återkommande bland förstagångsåkare året efter. 1998 var det extremt bra väder då Vasaloppet genomfördes och andelen återkommande ökade. 1999 var det dessutom 75-årsjubileum vilket ökade antalet återkommande. 1999 var det mycket dåligt väder då Vasaloppet genomfördes och antalet återkommande bland förstagångsåkare minskade ännu en gång.

Vasaloppsarrangörerna är oroliga för evenemangets framtid eftersom medelåldern ökar bland åkarna. Medelåldern bland alla åkare under perioden 1992-2002 har ökat med 2 år, och detta är samma ökning som för Sveriges befolknings medelålder under samma period som undersökningsperiod, se appendix 2 (SCB, 2003). Men en viktig punkt är att medelåldern ligger på 41 år och det är högre ålder än medelåldern inom andra sport evenemang.

Eftersom en förstagångs ungdom har större sannolikhet för att återkomma året efter än en vuxen när han/hon åker snabbt (*Figur 5*), så kan Vasaloppsarrangörerna satsa mer på att attrahera duktiga ungdomar för att åka Vasaloppet. Vasaloppsarrangörerna kan göra detta på olika sätt. Till exempel genom att erbjuda specialrabatter för ungdomar, särskilt för studentklasser som har låg inkomstnivå. Återkommande-sannolikheten ökar med åldern för dem som har brutit, t.ex. en 70 årig som har brutit har större sannolikhet för att återkomma än en ungdom (*Figur 5*). Detta kan Vasaloppsarrangörerna utnyttja genom att satsa på äldre som har brutit. Dvs Vasaloppsarrangörerna kan exempelvis skicka erbjudande hem till dem för att inte känna sig bortglömda.

En annan viktig faktor som jag tror kan påverka Vasaloppet och dess deltagare för att återkomma, är deltagaravgiften som många tycker är för hög (egna intervjuer under vasaloppsveckan 2003). Även transport och boende under Vasaloppsveckan kostar mycket för dem som kommer långt ifrån Dalarna till Vasaloppet. Exempelvis vasaloppet kanske kan sänka deltagaravgift med (10) kronor varje år för de som åker mer än engång.

Vasaloppsåkarna kommer i första hand från norra Sverige. Exakt vad detta beror på är svårt att säga (Henriksson, 2003). Detta mönster framkommer tydligt när man räknar antal deltagare i förhållande till antalet invånare per varje område. Men förstagångsåkare kommer inte främst ifrån norra Sverige (*Figur 2*).

En förstagångsåkare som kommer från Skåne har större sannolikhet för att återkomma än en ifrån Norrbotten, fast de som kommer ifrån Norrbotten har bättre träningsförhållanden än de ifrån södra Sverige (*Figur 4*). Personer från t ex Skåne tycker att Vasaloppet är exotiskt (Budh & Eriksson, 2003) vilket kan också vara en anledning till att de återkommer.

Eftersom varje tävlingstyp har enskilda krav, längd på lopp, t.o.m. deltagarnas kön, så skiljer sig sannolikheten också för en förstagångsåkare att återkomma. Tjejvasan har större återkommande sannolikhet. Detta beror kanske på att Tjejvasan är det enda skidtävling i Sverige som gäller endast kvinnor.

Den svenska arbetsmarknaden idag har i första hand ett stort problem med sjukskriven personal. Det finns en viktig ekonomisk effekt som detta skideevenemang kan tänkas ha för folkhälsan samt för landets ekonomi. Vasaloppet kan genom att introducera ett nytt nätverk, kommunikation som ett projekt mellan klubbar och försäkringskassan samt själva Vasaloppet. Då kan försäkringskassan delta med ett bidrag och Vasaloppet med att förmedla kontakter till aktiva klubbar där sjukskrivna kan finna ny motivation för att minska sjukskrivningarna.

Denna studie kan man utveckla genom att t.ex. undersöka dem som anmäler sig men inte återkommer, och ta reda på sannolikheten för att de återkommer påföljande år. Man kan också ta klubb kod med i resultaten för att undersöka om klubben har effekt för att dess deltagare återkommer till Vasaloppet. Detta har vi inte gjort för att många observationer inte hade klubb kod, men i så fall behöver Vasaloppet utveckla större noggrannhet av ifyllandet av koder vid registrering.

För att öka förståelsen om vilka förstagångsåkare som väljer att återkomma måste man göra en grundläggande undersökning genom intervjuer eller enkäter. Detta kan Vasaloppet samordna för att leta efter de orsaker som är anledning till att deltagarna inte kommer tillbaka eller har lite möjlighet att komma tillbaka.

5. SAMMANFATTNING

Ca 20 000 skidåkare deltar i Vasaloppsveckan varje år. Syftet med denna uppsats är att undersöka sannolikheten att förstagångsåkare i Vasaloppsevenemanget återkommer påföljande år. Vilka faktorer, som registrerats i Vasaloppets databas, påverkar denna sannolikhet? Resultaten kan bidra till nya perspektiv på deltagare och kan ge Vasaloppsevenemang viktiga erfarenheter om hur resultaten används som bas för fortsatt marknadsföring.

Sannolikheten att återkomma studerades med logistisk regression utifrån data i Vasaloppets databas 1992-2001.

De fem tävlingstyperna: Vasaloppet, Öppetspår, Tjejvasan, Halvvasan och Kortvasan. Eftersom själva Vasaloppet är den mest kända och viktiga loppet bland de andra löptävlingstyperna, så handlar resultaten mest om detta lopp. Resultaten visade att återkommande sannolikheten beror på geografiskt område. En förstagångsåkare från södra Sverige har t ex större chans att återkomma.

Ålder är också en viktig faktor som enligt mitt resultat har visat sig påverkat förstagångsåkare för att återkomma. Ungdomar bland förstagångsåkare har större sannolikhet att återkomma året efter om han/hon åker snabbt. Samma sannolikhet men för de som har brutit ökar med åldern, dvs äldre förstagångsåkare har större chans att återkomma året efter när han/hon har brutit än en yngre.

Jag tycker Vasaloppsarrangörerna kan använda resultaten i sin marknadsföring genom att satsa mer på de nya åkare som har stor chans att återkomma.

Appendix 1

Ett enkelt exempel

För att underlätta de siffrorna som finns i Tabell 6, så kommer jag med nedanstående exempel:

För själva Vasaloppet var den slutgiltiga modellen (Tabell 5):

$$\log(p/(1-p)) = \text{konstant} + b_1 * \text{ålder} + \text{kön}_j + \text{år}_k + \text{tid}_l + \text{postnr}_m + b_{2l} * (\text{ålder} * \text{tid}) + (\text{kön} * \text{tid})_j$$

Utifrån resultaten i Tabell 7 får man att för t ex en 30 åring man som åkte 2001 på 5 timmar från post nummer 2, blir:

$$\begin{aligned} \log(p/(1-p)) &= -0,901 + (30 * 0,056) + 0,15 - 0,776 + 2,575 + 0,362 - (30 * 0,092) + 0,022 \\ &= 0,352 \end{aligned}$$

Efter man har räknat $\log(p/(1-p))$, så kan man räkna p (sannolikheten för återkommande året efter) på följande sätt:

$$\begin{aligned} p &= \exp(0,352) / (1 + \exp(0,352)) \\ &= 0,59 \end{aligned}$$

Värdet (0,59) kan man tolka så här, att sannolikheten för den man som åkte år 2001 för att återkomma år 2002 är 59 %.

Eftersom man har skattningar för varje tävlingstyp i tabell 6, så på ovannämnda sätt kan man räkna återkommande sannolikheten för varje tävlingstyp med olika förklarande variabel.

Appendix 2

Medellivslängd för åren 1751-2002 (SCB källa).

År	Återstående medellivslängd i år					
	Män			Kvinnor		
	Vid födseln	Vid 50	Vid 65	Vid födseln	Vid 50	Vid 65
1751-1790	33,72	18,16	10,02	36,64	19,59	10,51
1791-1815	35,35	17,26	9,29	38,44	18,68	9,84
1816-1840	39,50	17,55	9,58	43,56	19,60	10,44
1841-1850	41,66	18,10	9,62	46,10	20,24	10,57
1851-1860	40,49	18,49	10,03	44,40	20,26	10,77
1861-1870	42,80	19,40	10,40	46,37	21,30	11,40
1871-1880	45,27	20,80	11,20	48,62	22,60	12,20
1881-1890	48,55	21,94	11,96	51,47	23,64	12,92
1891-1900	50,94	22,44	12,24	53,63	24,04	13,13
1901-1910	54,53	23,17	12,81	56,98	24,74	13,69
1911-1920	55,60	23,35	12,84	58,38	24,79	13,69
1921-1930	60,97	24,15	13,24	63,16	25,14	13,85
1931-1940	63,76	24,08	13,06	66,13	25,13	13,64
1941-1950	68,06	25,01	13,60	70,65	26,27	14,30
1951-1960	70,89	25,54	13,85	74,10	27,47	15,00
1961-1970	71,73	25,70	13,93	76,13	28,87	16,05
1971-1980	72,26	25,82	14,10	78,10	30,41	17,47
1981-1985	73,55	26,46	14,60	79,53	31,45	18,39
1986-1990	74,37	27,17	15,09	80,22	32,06	18,91
1991-1995	75,60	28,03	15,70	80,98	32,59	19,42
1996-2000	76,89	28,95	16,35	81,83	33,19	19,93
1982	73,43	26,37	14,56	79,38	31,34	18,32
1983	73,62	26,51	14,65	79,61	31,59	18,49
1984	73,84	26,73	14,81	79,89	31,72	18,65
1985	73,79	26,60	14,66	79,68	31,59	18,50
1986	73,97	26,83	14,80	79,99	31,84	18,69
1987	74,16	26,99	14,99	80,15	31,99	18,90
1988	74,15	26,99	14,95	79,96	31,85	18,70
1989	74,79	27,56	15,40	80,57	32,37	19,17
1990	74,81	27,50	15,30	80,41	32,20	19,04
1991	74,94	27,60	15,42	80,54	32,34	19,21
1992	75,35	27,82	15,55	80,79	32,42	19,27
1993	75,49	27,91	15,56	80,79	32,40	19,19
1994	76,08	28,43	16,03	81,38	32,92	19,75
1995	76,17	28,42	15,97	81,45	32,90	19,70
1996	76,51	28,61	16,10	81,53	32,95	19,73
1997	76,70	28,77	16,25	81,82	33,20	19,92
1998	76,87	28,94	16,34	81,94	33,30	20,03
1999	77,06	29,11	16,45	81,91	33,23	19,92
2000	77,38	29,41	16,69	82,03	33,30	20,08
2001	77,55	29,60	16,88	82,07	33,36	20,06
2002	77,73	29,64	16,90	82,11	33,37	20,01

Referenser

- Bohlin, M. (1994), *Vasaloppet – ett evenemang och en industri*, Lund SVEBIS, ”svensk förening för beteendevetenskaplig idrottsforskning”, s.64 – 88.
- Henriksson, C. (2003), *Vasaloppet 2003-deltagandet ur ett geografiskt perspektiv*, C-uppsats, Högskolan Dalarna.
- McCullagh, P. & Nelder J. A. (1989), *Generalized Linear Models*, London, Chapman and hall.
- Olsson, U. (2002), *Generalized linear models (An applied approach)*. Studentlitteratur, Lund
- Ordbok (1992), *Nationalencyklopedins ordbok*, 1992, band 10 ISSM-KIK.
- SCB (2003), http://www.scb.se/templates/tableOrChart_75552.asp , 2003-09-02
- Söderlund, G. (2000), *Skid-Sveriges framtid – mörk eller ljus?*. Vasalöpern, årgång 18, nr.1, Mora s.6-7
- Vasaloppet(2003), http://www.vasaloppet.se/wps/vasa_CNS/generated/resources.Generella_sidor/se/Bakat_i_tiden/se_bakat_i_tiden.html ,2003-08-14

Muntliga källor:

- Budh, B. (2002), Verkställande direktör på Vasaloppet, *intervju*, 021205
- Eriksson, M. (2002), Informationsansvarig på Vasaloppet, *intervju*, 021205
- Budh, B. (2003), Verkställande direktör på Vasaloppet, *intervju*, 031211
- Eriksson, M. (2003), Informationsansvarig på Vasaloppet, *intervju*, 031211

Övriga:

- Databas (2002), *Vasaloppet databas i Microsoft Access*, som är konstruerade 1989.
- Vasaloppet (2000), *Vasaloppet material och information år 2000*