

Metodproblem vid studier av Högskoleprovets prognosförmåga – och deras lösning!

JAN-ERIC GUSTAFSSON

Institutionen för pedagogik och didaktik, Göteborgs universitet

SVEN-ERIC REUTERBERG

Institutionen för pedagogik och didaktik, Göteborgs universitet

Sammanfattning: *Det klassiska problemet då man skall bestämma ett urvalstests prognostiska validitet är att kriterieresultat finns tillgängliga enbart för dem som blivit antagna och dessa utgör givetvis ett positivt gallrat urval av samtliga sökande. Den beskurna variationen leder då till att sambandet mellan urvalstestet och kriteriet utgör en underskattning av den verkliga validiteten, det vill säga den som gäller för den totala sökandegruppen. När det gäller Högskoleprovet kompliceras bilden ytterligare av det faktum att urvalet till högre utbildning är kompensatorisk. Man antas antingen på grundval av provresultat eller på grundval av höga gymnasiebetyg. I artikeln demonstreras att ett sådant urvalsförfarande kan leda till att man finner betydande negativa samband mellan provresultat och framgång i den högre utbildningen bland de antagna trots att motsvarande samband inom hela sökandegruppen är starkt positivt. Med hjälp av simulerade data visas att man med en statistisk metod kallad »modellering av inkompleta data» kan erhålla mycket säkra skattningar av provets sanna prognosvärde.*

Ett av den pedagogiska psykologins allra mest framgångsrika praktiska tillämpningsfält gäller användning av test för urval till utbildning och yrkesverksamhet. I en nyligen publicerad forskningsöversikt drar Schmidt och Hunter (1998) slutsatsen att test som mäter allmän intellektuell förmåga ger de högsta sambanden med en bred uppsättning kriterier på framgång i studier och arbete. För denna kategori av urvalsinstrument rapporterar Schmidt och Hunter att korrelationen med kriteriet i medeltal är 0,56. Även om korrelationer av denna storleksordning ger en viss osäkerhet vid förutsägelser i det enskilda fallet medför de mycket betydande effektivitetsvinster. Detta gäller särskilt i de fall då en liten andel av de sökande antas till utbildningen eller anställningen, och variationen i resultat är stor.

Till den kategori av urvalsinstrument som mäter allmän intellektuell förmåga hör det svenska *Högskoleprovet* (se Gustafsson, Wedman & Westerlund 1992). I den mån de internationella resultaten är generaliserbara till svenska förhållanden borde Högskoleprovet ha ett starkt samband med framgång i högre utbildning. Då vi nyligen genomförde en serie analyser av samband mellan Högskoleprovsresultat och studieframgång fann vi dock inte de förväntade resultaten. Sambanden var i allmänhet svagt positiva, men i vissa fall var de nära 0, och i några fall till och med negativa! Ett negativt samband betyder att personer med höga poäng på Högskoleprovet får sämre studieresultat än personer som har låga högskoleprovsresultat. Även andra studier har kommit fram till att sambandet mellan Högskoleprovet och studieframgång är svagt. Wolming drog sålunda slutsatsen efter att ha studerat utbildningar i företags ekonomi, socialt arbete, teknik och medicin att: »The association between SweSAT scores and academic performance is almost non-existent with some exceptions.» (Wolming 1999 s 346)

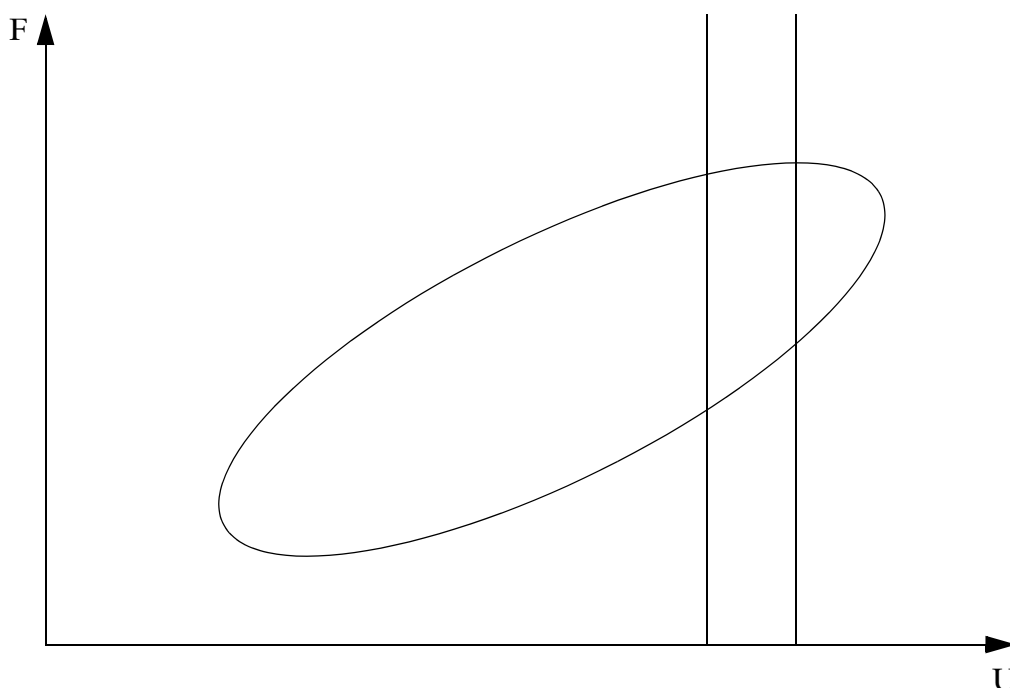
Betyder detta då att Högskoleprovet saknar förmåga att förutsäga framgång i högre studier? Det fullständiga svaret på frågan om Högskoleprovets prognosförmåga har vi ännu inte. Men vi vet nu att de svaga och negativa samband som vi fann i våra ursprungliga analyser ger en felaktig bild av Högskoleprovets prognosförmåga. Dessa artefakter orsakas bland annat av det sätt på vilket urvalet till högre utbildning sker i Sverige. Dessa metodproblem är av stort generellt intresse, och i denna artikel gör vi ett försök att reda ut dem.

ENKELT URVALSFÖRFARANDE

Det klassiska problemet när det gäller att avgöra hur väl ett urvalsprov kan förutsäga framgång, är att man får ett mått på graden av framgång enbart för dem som blir utvalda. Den utvalda gruppen har normalt ett högre medelvärde på urvalsprovet jämfört med samtliga provtagare och dessutom är variationen i provresultat avsevärt mindre. Den beskurna variationen medför alltid en sänkning av korrelationen mellan en sådan uppsättning av mätvärden och andra variabler.

När vi har ett enda urvalsinstrument utifrån vilket urvalet sker kan man tala om ett enkelt urvalsförfarande. När vi tolkar korrelationen mellan urvalsprovet och framgång i ett sådant fall måste vi beakta den beskurna variationens effekter. Ett par exempel ges nedan för att belysa vilken effekt den beskurna variationen får:

Antag att vi har en grupp sökande till en utbildning och att urvalet sker utifrån deras resultat på provet »U». Vi antar vidare att vi känner korrelationen mellan dessa provresultat och ett framgångsmått »F» för hela provtagargruppen. Sambandet är 0,66 (hela korrelationsytan i Figur 1). Utifrån provresultaten väljs de 10 procent bästa ut och vi beräknar korrelationen mellan U och F för den utvalda gruppen. Då blir den 0,30. Skulle urvalet begränsas till enbart de 5 procent bästa skulle korrelationen mellan U och F sjunka till 0,23.



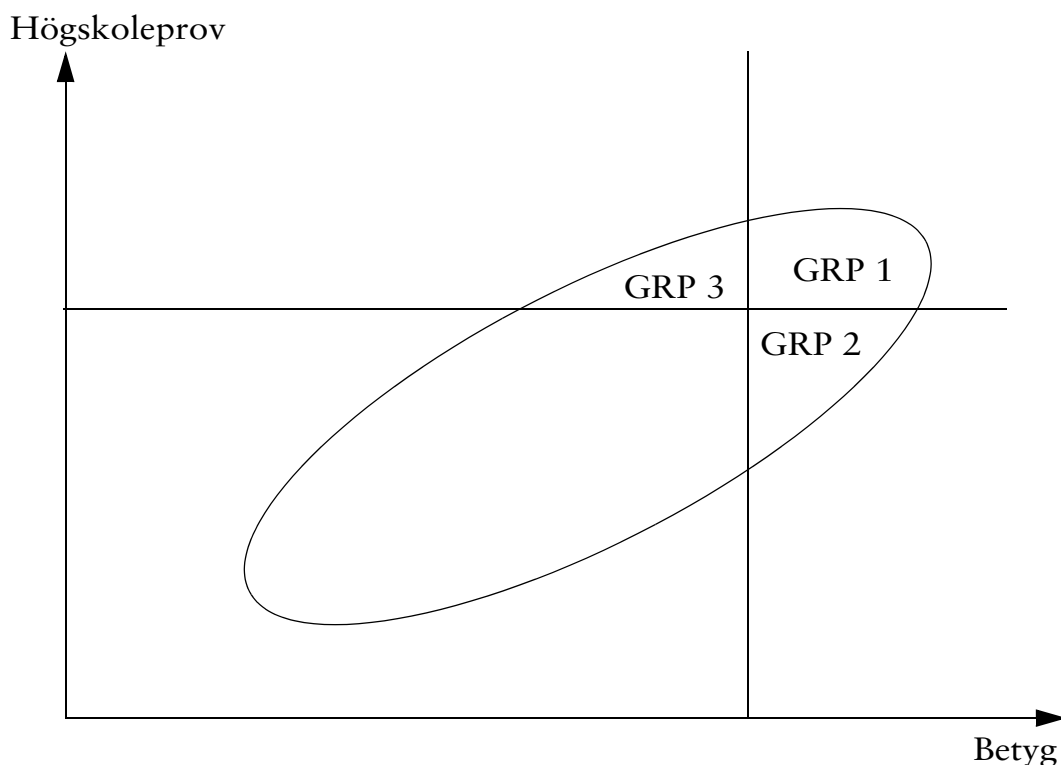
Figur 1. Hypotetiskt exempel på samband mellan variablerna U och F vid urval med avseende på U.

Eftersom urvalet sker bland samtliga sökande är provets verkliga prognosförmåga 0,66, men det värde vi kan beräkna utifrån tillgängliga data är det som gäller för den antagna gruppen. Som exemplet visar kan detta leda till en betydande underskattning av provets prognostiska värde.

Problemet att skatta provets verkliga prognosvärde är dock inte olösligt i detta fall. Man kan nämligen på statistisk väg göra en uppskattning av hur mycket korrelationen sänks av den beskurna variationen och därmed göra en skattning av vilken korrelation man skulle få om samtliga sökande blivit antagna.

KOMPENSATORISKT URVALSFÖRFARANDE

Problemet blir dock mer komplicerat om vi använder mer än ett urvalsinstrument. Enligt de regler som gäller sedan 1991 kan man antas till högre utbildning på grundval av antingen gymnasiebetyg eller högskoleprovsresultat, vilket innebär att vi har ett kompensatoriskt urvalsförfarande. Den sökande placeras i den urvalsgrupp där personen har bäst förutsättningar att bli antagen. Enligt antagningsreglerna skall minst en tredjedel antas utifrån gymnasiebetyg och minst en tredjedel utifrån högskoleprovsresultat. Varje utbildning har att bestämma sin egen urvalskvot mellan dessa båda gränser. Det faktum



Figur 2. Urvalsgrupper då samtidig hänsyn tas till de två variablerna högskoleprov och betyg.

att högskoleprovspoängen inte ensam avgör urvalet leder till att det blir avsevärt svårare att skatta dess prognostiska värde, vilket illustreras med ovanstående principiella figur.

Som Figur 2 visar kan de antagna grupperas i olika kategorier. En kategori utgörs av dem som kan bli antagna utifrån såväl betyg som högskoleprovsresultat (Grp 1). En annan kategori utgörs av dem som inte blir antagna utifrån högskoleprovsresultat, men som antas på grundval av goda betyg (Grp 2). Den tredje kategorien innefattar dem som antas utifrån sina provresultat men som inte skulle blivit antagna på grundval av betyg (Grp 3).

Om sambandet mellan provresultat och gymnasiebetyg varit mycket starkt skulle korrelationsytan varit smal, med få individer i kategorierna två och tre. Det faktiska sambandet är emellertid inte starkare än cirka 0,50 vilket ger en tämligen bred korrelationsyta. Detta gör i sin tur att vi har många sökande i kategorierna två och tre. Av det relativt låga sambandet följer också att de som väljs ut utifrån betyg har mycket varierande provresultat, liksom att de som blir antagna utifrån provresultat har mycket varierande gymnasiebetyg.

Det kompensatoriska urvalsförfarandet komplicerar följaktligen bilden på så sätt att somliga individer kan ha måttligt höga poäng på Högskoleprovet samtidigt som de har höga betyg och därmed bli antagna på grundval av dessa. Även om båda urvalsinstrumenten var för sig visar relativt starka samband

med framgångsmåttet kommer personer med höga betyg och låga högskoleprovsresultat att bidra till att sänka sambandet mellan framgångsmåttet och högskoleprovet. Vi belyser detta med några exempel baserade på fiktiva, men verklighetstroga, data. Data gäller för samtliga sökande och avser tre olika variabler: *Högskoleprov*, *Betyg* och *Studieresultat*. Högskoleprov representerar tänkta högskoleprovsresultat, Betyg gymnasiebetyg och Studieresultat framgång i den högre utbildningen i form av avklarade kurspoäng. De statistiska egenskaperna hos dessa data beskrivs i Tabell 1.

Tabell 1. Statistiska egenskaper för de tre variablerna bland samtliga sökande.

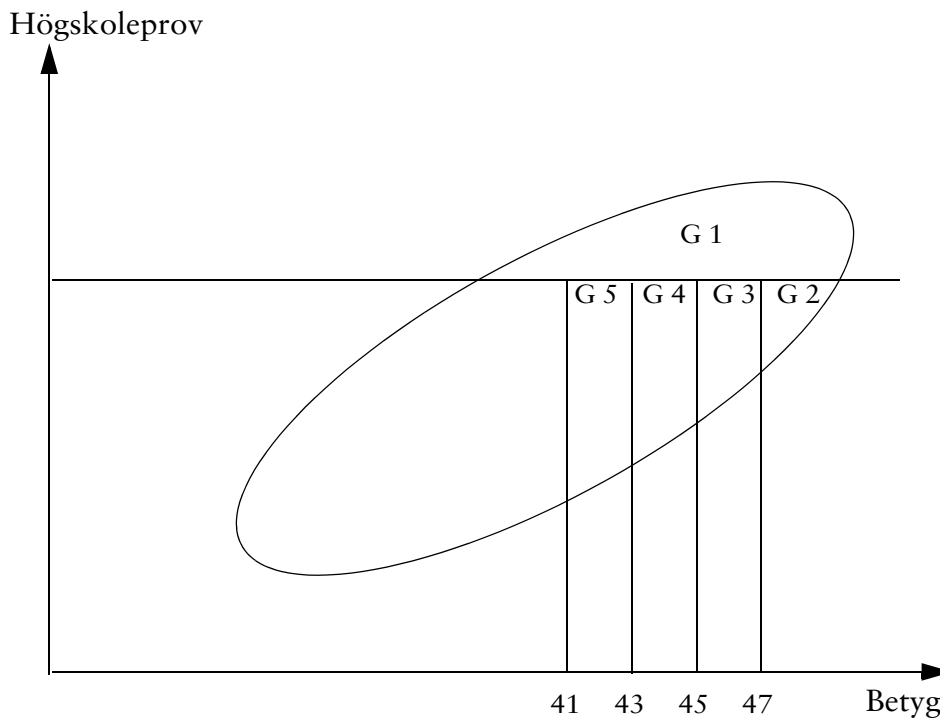
Variabel	Variationsområde	Medelvärde	Standardavvikelse
Högskoleprov	2 – 122	68,43	17,72
Betyg	10 – 50	35,13	4,26
Studieresultat	0 – 170	100,16	19,97

Totalt består sökandegruppen av 36 525 individer med uppgifter om samtliga tre variabler. Korrelationen mellan variablerna framgår av Tabell 2.

Tabell 2. Korrelationer mellan de tre variablerna.

	Högskoleprov	Betyg
Betyg	0,62	
Studieresultat	0,45	0,66

Vi startar med ett snävt urval utifrån enbart Högskoleprov och väljer ut de som har 116 poäng eller högre. Detta leder till att vi får 292 antagna (Grupp 1). Om vi först betraktar urvalsgränsen 47 på Betyg (Grupp 2 = G1 + G2) i Figur 3 på nästa sida, får vi ett tillskott av 27 individer, utöver de 292 som antas på Högskoleprov. Dessa 27 individer har alla en högskoleprovspoäng som är lägre än 116. Om vi sänker betygsgränsen till 45 (Grupp 3 = G1 + G2 + G3 i Figur 3) antas sammanlagt 513 personer.



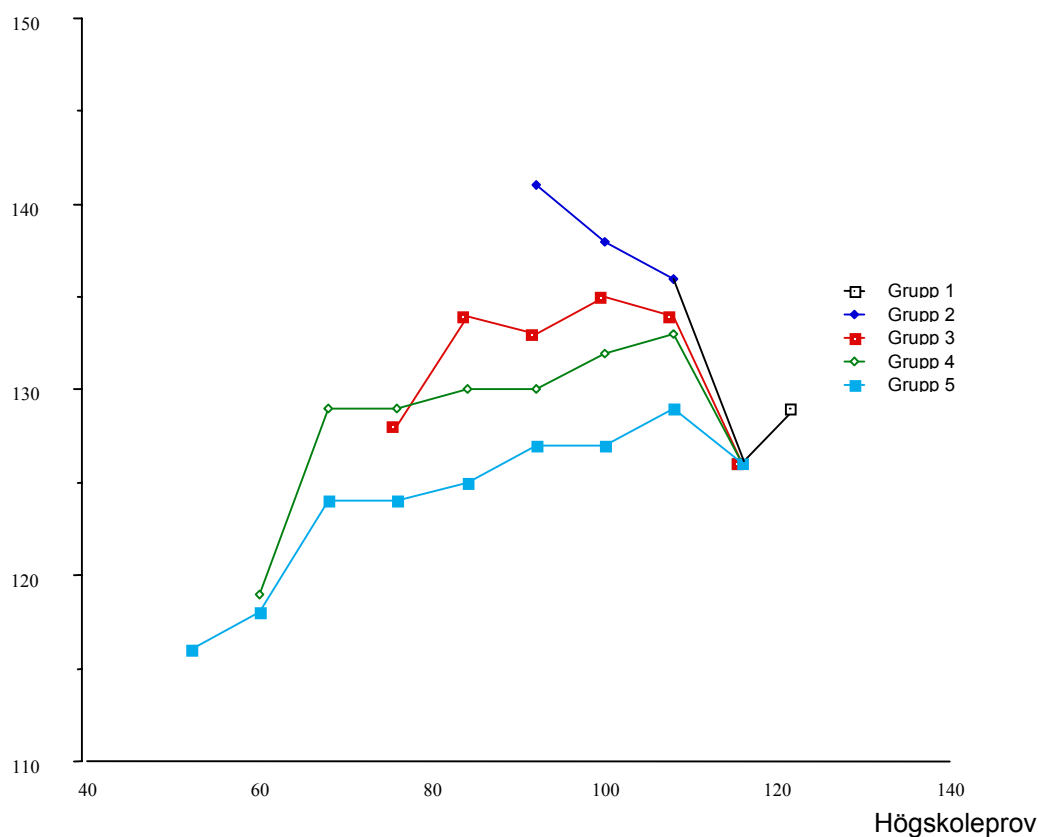
Figur 3. Urvalsgrupper vid olika urvalsgränser för Betyg och Högskoleprov.

I Figur 4 (på nästa sida) visas Studief framgång som en funktion av Högskoleprov, dels för de som antas vid betygsgränsen 47 (Grupp 2), dels för de som antas vid betygsgränsen 45 (Grupp 3). I båda fallen finner vi ett negativt samband mellan Högskoleprov och Studieresultat. De högsta medelvärdena för Studieresultat uppnås av de som har de lägsta Högskoleprovspoängen. Anledningen till att dessa personer trots den förhållandevis blygsamma Högskoleprovspoängen når så goda resultat är givetvis att de har mycket höga betyg.

Betraktar vi Figur 4 närmre finner vi att för personer med de allra högsta poängen på Högskoleprovet (dvs de som har 116 poäng eller högre) är sambandet svagt positivt (0,09), medan det är starkt negativt för de som har lägre Högskoleprovspoäng och som är betygsantagna. Den sammantagna effekten är dock att korrelationerna blir svagt negativa (-0,14 för Grupp 2 och -0,15 för Grupp 3). För Grupp 4 (betygsgräns 43 med 1.169 antagna) och Grupp 5 (betygsgräns 41 med 2 778 antagna) i Figur 3 ovan uppgår sambanden till 0,00 respektive 0,12.

Enligt Tabell 2 är sambandet mellan Högskoleprov och Studieresultat för den totala sökandegruppen 0,45. Detta är således det sanna prognosvärdet för Högskoleprovet. Har vi tillgång till data enbart för dem som blir antagna, visar de korrelationskoefficienter vi presenterat ovan, vilka varierar mellan -0,15 och 0,12, att vi kommer att grovt underskatta urvalsinstrumentets prognosvärde. För två av urvalsgrupperna blir sambandet till och med

Framgång



Figur 4. Samband mellan Studieframgång och Högskoleprov inom två grupper med högt Betyg.

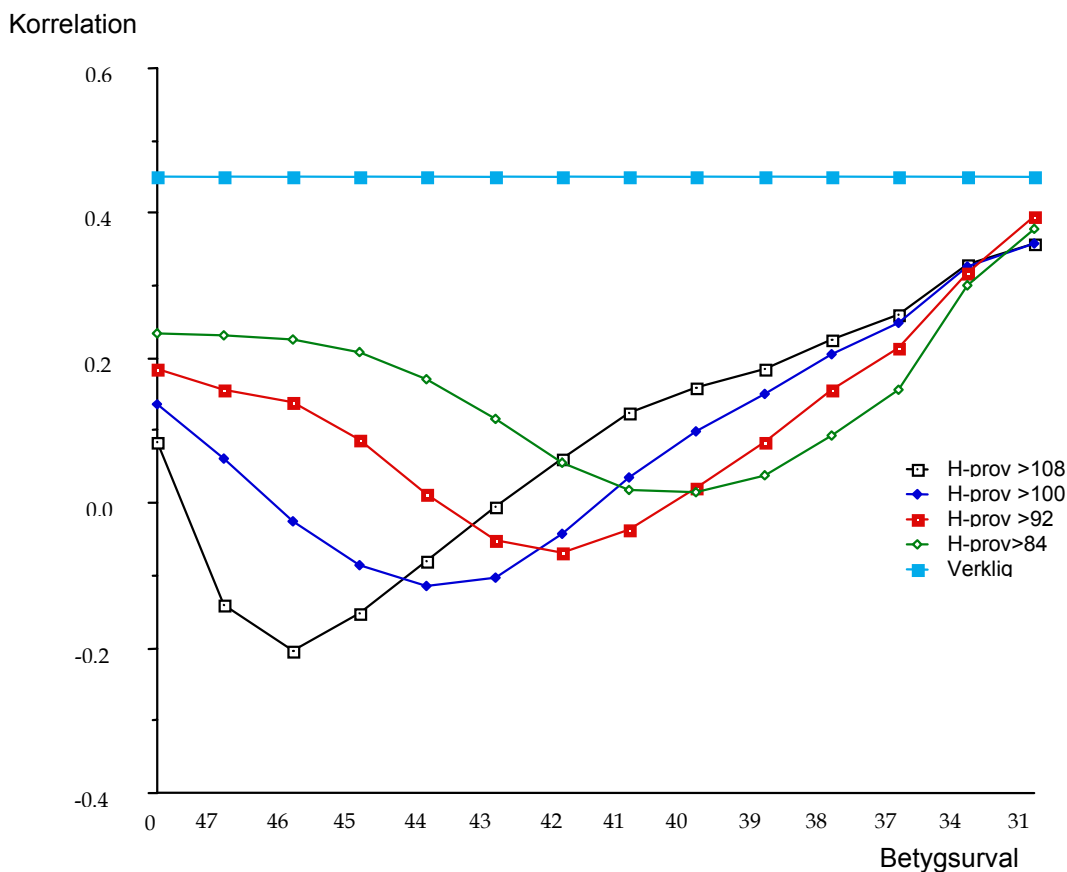
negativt, vilket kan föranleda den felaktiga slutsatsen att Högskoleprovet främst väljer ut sådana individer som kommer att misslyckas i sina studier. Så behöver emellertid inte alls vara fallet därför att sambanden baseras på enbart de antagna, medan högskoleprovet skall användas för urval bland samtliga sökande.

För Grupp 1 som valts ut enbart utifrån Högskoleprovsresultat sänks sambandet givetvis av den starkt beskurna variationen i Högskoleprov, men det är fortfarande svagt positivt. För Grupp 2 och Grupp 3 medför den kompensatoriska urvalsmodellen att åtskilliga sökande med mycket god förmåga enligt Betyg antagits. Samtidigt har dessa inte nått toppresultat på Högskoleprov, vilket sänker sambandet mellan Högskoleprov och Studieresultat. Går vi till grupp Grupp 4 har den antagna gruppen mer än fördubblats jämfört med Grupp 3. Det största tillskottet får vi bland dem som nått 60 poäng på Högskoleprovet, vilket är betydligt lägre än genomsnittsvärdet för samtliga sökande (jfr Tabell 1). De har samtidigt varit måttligt framgångsrika i sina studier, vilket påverkar sambandet mellan Högskoleprov och Studieresultat i positiv riktning och vi får ett totalt samband på 0,00. För Grupp 5 slutligen får

vi ett stort ytterligare tillskott av sökande med låga resultat på Högskoleprovet och som inte varit speciellt framgångsrika i studierna. Denna grupp påverkar ytterligare sambandet mellan Högskoleprov och Studieresultat i positiv riktning och för hela gruppen blir det 0,12.

Nu har vi med exemplen ovan beskrivit en ganska extrem situation, såttillvida att vi satt en mycket snäv gräns för urval utifrån Högskoleprov. För att göra bilden mer generell har vi simulerat fall också med andra gränsdragningar. Vi har sålunda liberaliserat gränserna för Högskoleprovsantagning genom att införa fyra nya antagningsgränser: 108, 100, 92 samt 84 poäng. Samtidigt har vi också lagt till nya och mer liberala gränser för betygsantagning.

Liksom tidigare redovisar vi först det fall, där urvalet sker helt utifrån Högskoleprov (värdet 0 på betygsskalan) och därefter det kompensatoriska urvalsförfarandet då även hänsyn tas till Betyg. Figur 5 visar hur korrelationen mellan Högskoleprov och Studieresultat förändras med de varierande antagningsgränserna. I figuren visas också det faktiska sambandet på 0,45, vilket ju är konstant över de olika antagningsgränserna.



Figur 5. Observerade samband mellan Studieresultat och Högskoleprov vid olika urvalsgränser för Betyg och Högskoleprov.

Kurvornas läge vid betygsgränsen 0 beskriver den urvalssituation då hela antagningen baseras på Högskoleprov. Då antagningsgränsen sänks från 108 högskoleprovspoäng till 84 poäng ökar sambandet mellan Högskoleprov och Studieresultat från 0,09 till 0,23. Denna ökning förklaras naturligtvis av att variationen i högskoleprovresultat blir allt mindre beskuren. När vi sedan går över till det kompensatoriska urvalsförfarandet och då även individer med höga betyg väljs ut, sjunker sambandet tämligen snabbt och för tre av de fyra fall som kurvorna beskriver övergår det till att bli negativt. Sambandet sjunker mest drastiskt då vi har den högsta antagningsgränsen för Högskoleprov och i detta fall finner vi också det starkaste negativa sambandet (-0,20). Negativa samband förekommer också då antagningsgränsen för Högskoleprov är 100 poäng respektive 92 poäng. Däremot blir sambandet aldrig negativt för den mest liberala antagningsgränsen (högre än 84 poäng), även om det ligger nära värdet 0 då betygsgränsen ligger omkring 40.

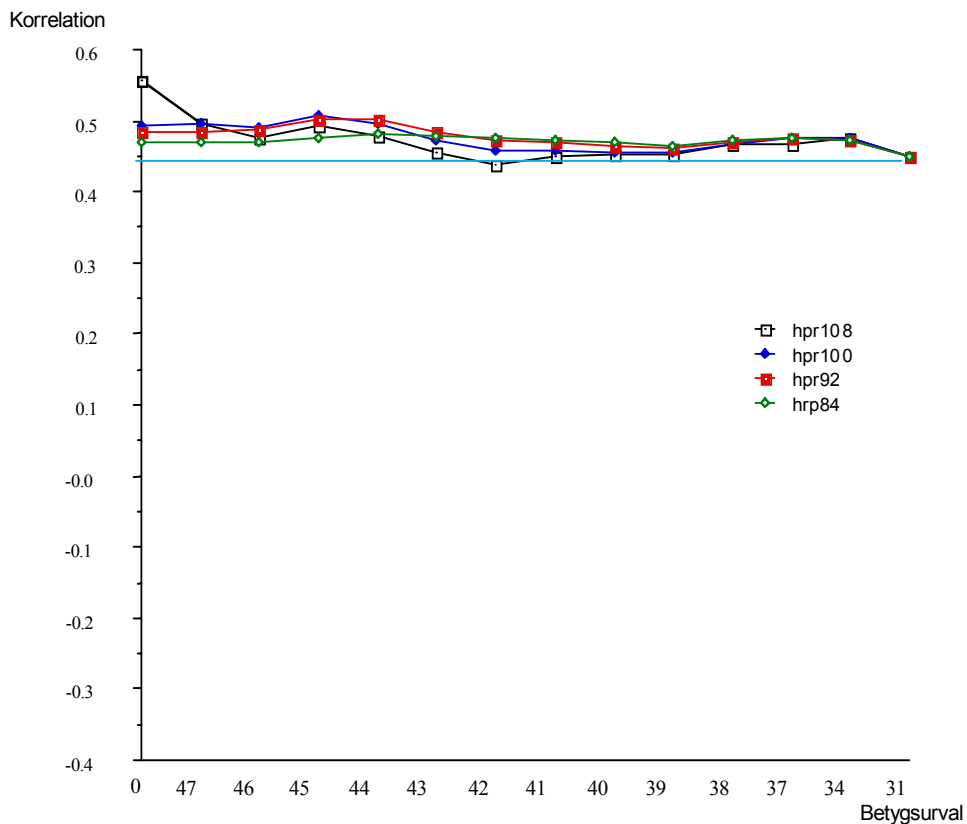
Om betygsgränsen blir mer liberal förändras sambandet i positiv riktning och vid den mest liberala gränsen, betyg över 31, närmar sig sambandet det korrekta prognosvärdet 0,45. Det når emellertid inte ända fram till 0,45 trots att urvalet här är så liberalt att inte mindre än 83 procent av de sökande blir antagna. Det enda fall där sambandet skulle nå upp till det faktiska värdet av 0,45 är då samtliga sökande antas, dvs inget urval sker. Därmed är ju också gruppen antagna identisk med sökandegruppen. Sambanden i Figur 5 visar således att vi aldrig kan avgöra ett urvalsinstruments prognostiska värde genom att studera enbart den grupp som blivit antagen. Sambanden inom gruppen antagna ger alltid en underskattning av det prognostiska värdet, dels som en följd av beskuren variation, dels som en följd av det kompensatoriska urvalsförfarandet.

LÖSNINGEN PÅ PROBLEMET

Urvalsinstrumentets prognostiska värde ligger naturligtvis inte i hur väl det kan förutsäga studieframgång bland dem som blivit antagna, utan i hur effektivt det bland alla sökande kan välja ut dem, som har de bästa förutsättningarna att bli framgångsrika. Man behöver således information om både de antagna och dem som inte antagits. Problemet är att de icke antagna saknar mått på framgång.

En statistisk metod som kallas »modellering av inkompleta data» gör det dock möjligt att utnyttja resultat från samtliga individer, även om de saknar uppgift på vissa variabler. Denna metod utvecklades under slutet av 1980-talet (se t ex Muthén, Kaplan & Hollis 1987) på grundval av den svenske statistikern Karl Jöreskogs arbete med så kallade strukturella ekvationsmodeller. När det gäller urvalet till högre studier har de allra flesta sökande såväl gymnasiebetyg som resultat på högskoleprovet. Utifrån dessa kända fakta gör metoden det möjligt att skatta hur väl de skulle ha lyckats inom den högre utbildningen om de hade blivit antagna och påbörjat studierna. Man kan säga att vi på statistisk väg skapar en situation där sambandet mellan Högskoleprov och Studieresultat beräknas bland samtliga sökande.

Det är i detta sammanhang endast möjligt att antyda hur beräkningarna genomförs för att erhålla de korrekta sambanden från de ofullständiga data. I kort sammanfattning kan beräkningarna sägas bestå av tre steg. I det första steget genomförs en multipel regressionsanalys där Studieresultat prediceras från Högskoleprov och Betyg bland de antagna. Denna analys ger samtidigt uppgift om sambandet mellan Studieresultat och Betyg för personer med samma högskoleprovsresultat, och om sambandet mellan Studieresultat och Högskoleprov för personer med samma betyg. De ostandardiserade partiella regressionskoefficienter som bestäms i denna analys gäller för såväl de antagna som de icke antagna. I det andra steget beräknas de standardiserade partiella regressionskoefficienterna med utgångspunkt i sambandet mellan Betyg och högskoleprovsresultat för såväl de antagna som de icke antagna. I det tredje steget, slutligen, transformeras de standardiserade partiella regressionskoefficienterna till korrelationer för Betyg och Högskoleprov var för sig. Vi har med hjälp av denna metod på nytt skattat sambanden mellan Högskoleprov och Studieresultat på samma datamaterial som låg till grund för sambanden i Figur 5. Resultaten återges i Figur 6 nedan. Liksom i Figur 5 har vi här också lagt in det faktiska sambandet som ju uppgick till 0,45.



Figur 6. Korrigerade samband mellan Studieresultat och Högskoleprov vid olika urvalsgränser för Betyg och Högskoleprov.

Vi får nu en helt annan resultatbild än då sambanden beräknades enbart bland dem som blivit antagna. För det första ligger kurvorna mycket nära varandra, vilket visar att resultatet blir i stort sett det samma oavsett vilken grupp det beräknas för. För det andra, och viktigast av allt, ligger sambanden mycket nära det faktiska prognosvärdet på 0,45. Genomsnittet för samtliga samband uppgår till 0,47 och enbart fyra av totalt 56 samband överstiger 0,50. Den största avvikelser får vi i det fall då urvalet skett enbart utifrån Högskoleprov och då det krävts en poäng överstigande 108, vilket är en extremt hård urvalsgräns med mindre än 1 procent antagna. Då blir sambandet 0,56. Detta värde kan för övrigt jämföras med det vi fick för gruppen antagna och som var 0,09.

SLUTSATS

Vår slutsats är att modellering av inkompleta data är en tillförlitlig metod för att skatta prognosvärdet hos urvalsinstrument. Användning av de enklare, traditionella, metoderna ger däremot uppenbart missvisande resultat om prognosvärdet hos de instrument som används vid antagning till den svenska högskolan. Det kan avslutningsvis noteras att våra resultat kring den kompensatoriska antagningsmodellens egenskaper sannolikt kan förklara varför det misstag hösten 1993 då en grupp studerande med låga Högskoleprovsresultat antogs till tandläkarlinjen kunnat betraktas som »en ovanligt lyckad tabbe» (Djurberg 2000). De goda studieresultaten för denna grupp torde kunna hänföras till det faktum att de hade såväl goda betyg som goda resultat i den antagningsintervju som i detta fall utgjorde en tredje urvalsgrund.

LITTERATUR

- Djurberg, S. 2000: En ovanligt lyckad tabbe. Felaktigt antagna tandläkarstudenter klarade sig bra. *Tandläkartidningen*, 92(2), 27–28.
- Gustafsson, J-E., Wedman, I. & Westerlund, A. 1992: The dimensionality of the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 36(1), 21–39.
- Muthén, B.O., Kaplan, D. & Hollis, M. 1987: On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52(4), 431–462.
- Schmidt, F.L. & Hunter, J.E. 1998: The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.
- Wolming, S. 1999: Validity issues in higher education selection: A Swedish example. *Studies in Educational Evaluation*, 25, 335–351.