

Avdelningen för social statistik och analys  
Sektionen för uppdragstjänster

## **Resursfördelningsmodellen**

*Region Gotlands skolor 2024*

### **Inledning**

Underlag för analyserna utgörs av ett register som innehåller elever som gått ut årskurs nio de senaste 5 åren. Registret innehåller uppgifter om skola och examensår, två variabler som beskriver elevens resultat och ett antal variabler som beskriver den socioekonomiska bakgrunden.

Analysen resulterar i ett index för var och en av skolorna som ingår i det inskickade materialet. En skola med en lägre andel elever som riskerar att inte klara målen än genomsnittet får ett index lägre än 100 och en skola med en högre beräknad andel som riskerar att inte klara målen får ett index som är högre än 100.

Indexet kan sedan användas som underlag för att fördela resurser mellan skolor i en kommun.

### **Variabler**

De variabler som beskriver elevens resultat är

1. eleven uppnår inte behörighet till gymnasieskolans yrkesprogram
2. eleven har fått icke-godkänt i minst två ämnen

En regressionsmodell har tillämpats på underlaget. Som beroende variabel används en variabel som beskriver om eleven anses ha uppnått målen eller inte. Denna variabel är en kombination av resultatvariablerna ovan. Om eleven uppfyller något av villkoren 1 eller 2 så anses eleven inte ha klarat målen.

Den beroende variabeln är vald för att de elever som inte uppnått behörighet eller har fått icke-godkänt i minst två ämnen antas vara de elever som har större behov av stöd och stimulans.

Som förklarande variabler används bakgrundsvariabler som beskriver egenskaper hos eleven och elevens hushåll, det är variabler som har visat sig vara viktiga i tidigare analyser. De variabler som används finns i Excel-filen ”\_Regression\_”.

Variablernas indelning och betydelse framgår av tabell 1.

## Analysmetod

Regressionsanalys är den generella metoden för att analysera samband mellan en beroende variabel och olika bakgrundsvariabler. En sådan analys ger underlag för att bedöma vilka variabler eller faktorer som är viktiga för att förklara den beroende variabeln och hur betydelsefulla faktorerna är i sig och i förhållande till varandra. Eftersom den beroende variabeln här kan anta två värden endast, används en variant av regressionsanalys - logistisk regression.

Vid logistisk regression kan regressionsmodellen beskrivas som

$$\log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

där  $p$  är sannolikheten att den beroende variabeln antar ett visst värde och  $x_1, x_2, \dots, x_k$  utgör  $k$  stycken förklarande variabler.  $\beta_1, \beta_2, \dots, \beta_k$  utgör regressionskoefficienterna och  $\alpha$  utgör interceptet.

Koefficienterna anger betydelsen av de olika bakgrunds faktorerna. Dessa bakgrunds faktorer har kodats så att olika grupper lätt kan jämföras med varandra. En grupp är referensalternativet för varje variabel eller faktor. För kön, till exempel, är flickor referensgruppen med en koefficient som är 0. Koefficienten för pojkar anger hur pojkar skiljer sig från flickor i genomsnitt, när vi kontrollerar för de övriga variablerna som ingår i regressionsmodellen. En positiv koefficient innebär att pojkar inte lyckats lika bra som flickor.

I logistisk regression räknas regressionskoefficienterna ofta om till så kallade oddskvoter. En oddskvot för könen är, till exempel, kvoten mellan oddset för att inte klara målen för pojkar och oddset att inte klara målen för flickor. Oddset motsvarar sannolikheten att inte klara målen dividerad med sannolikheten att klara målen ( $p/(1-p)$  där  $p$  är den skattade sannolikheten). En oddskvot större än 1 i exemplet anger då att pojkarna jämfört med flickorna har ett sämre utfall, en förhöjd risk att inte klara målen.

## Risksmodellen

Modellen är en standardiserad modell som är baserad på elever från hela riket. Riskmodellen rekommenderas för mindre samt medelstora kommuner på grund av att ett stort elevantal i modellen ger bättre skattningar.

I den bifogade (*Regression.xls*) filen presenteras antalet elever som modellen är baserat på under de senaste 5 åren, ni hitta mer information om modellens elevsammansättning i filen "FrekvensTabeller". Ni hittar även de samband som skattas mellan den beroende variabeln och de förklarande variablerna i filen

”ResultatTabell” detta presenteras i regressionskoefficienter och oddskvoter.

Vidare anges ”Percent concordant”, vilket är ett mått på hur väl modellen stämmer med data. Utifrån parvisa jämförelser beräknas hur ofta som modellen lyckas prediktera vilken elev som inte klarar målen. Parvisa jämförelser görs mellan elever som inte klarade målen och elever som klarade målen

En ”Percent concordant” på 75 procent innebär att modellen för 75 procent av paren predikterar en högre sannolikhet att inte klara målen för den elev som i själva verket inte klarade målen. I 25 procent av fallen tilldelar modellen en lägre (eller samma) sannolikhet att inte klara målen för den elev som i själva verket inte klarade målen. Måttet är inte jämförbart med förklaringsgraden i linjär regressionsanalys.

Ett lågt P-värde betyder att modellen kan lättare kategorisera om eleven är behörig till gymnasiet eller ej, ett högt värde betyder motsatsen. Samtliga variabler är signifikanta på 5 % nivå.

Tillämpning av modellen för att skatta variationen i behov

Modellen har tillämpats på elever som går i skolan i er kommun, för att få en uppfattning om deras förväntade risk att inte nå målen. Utifrån de skattade regressionskoefficienterna samt uppgifter om den socioekonomiska bakgrunden för en elev så erhålls en skattad sannolikhet för att en elev inte ska uppnå målen.

De skattade sannolikheterna på elevnivå summeras till skolenivå. För varje skola skattas antal och andel elever som inte förväntas nå målen. Utifrån de skattade andelarna beräknas ett index för varje skolenhet. Indexet kan sedan användas för att få fram hur behovet av stöd och stimulans varierar mellan kommunens skolor.

När man använder indexen bör man beakta att de är modellskattningar och att individer i själva verket kan prestera annorlunda än vad modellen säger. Modellen säger hur en elev i genomsnitt borde prestera. Indexen ska ses som en vägledning om hur det kommer gå för eleverna i en skola. Särskild försiktighet bör iaktas om det handlar om små skolenheter.

Resultaten för skolenheter med färre än 5 elever anses osäkra och resultatet bör användas med försiktighet.