

Logistisk regression av bortfall

Logistisk regression används då vi vill undersöka samband mellan en dikotom responsvariabel Y och en eller fler förklarande X -variabler. I vårt fall är responsvariabeln olika uppdelningar av urvalsmängden och de förklarande variablerna är registervariabler från folkbokföringsregistret. Här låter vi X -variablerna vara dummy-variabler som indikatorer för respektive kategori av registervariabeln.

Modeller

Urvalet delas upp i kategorierna *ej anträffad (I)* och *anträffad (II)*, som delas upp i *svarande (IIa)*, *avböjd medverkan (IIb)* och *övrigt bortfall (IIc)*, samt *övertäckning (III)*. Vi skapar fyra modeller där första responsvariablerna definieras som

$$Y_1 = \begin{cases} 1, & \text{om ej svarande/bortfall (I, IIb, IIc)} \\ 0, & \text{om svarande (IIa)} \end{cases}$$

$$Y_{2a} = \begin{cases} 1, & \text{om ej anträffad (I)} \\ 0, & \text{om anträffad (II)} \end{cases}$$

$$Y_{2b} = \begin{cases} 1, & \text{om övrigt bortfall (IIc)} \\ 0, & \text{om svarande eller avböjd medverkan (IIa, IIb)} \end{cases}$$

$$Y_{2c} = \begin{cases} 1, & \text{om avböjd medverkan (IIb)} \\ 0, & \text{om svarande (IIa)} \end{cases}$$

Notera att övertäckningen tas bort från alla modeller och att Y_{2a-c} hör ihop hierarkiskt så att objekten i $Y_{2a}=0$ utgör samtliga objekt för Y_{2b} osv. Mängden i Y_{2c} kan sägas utgöra de som har haft möjlighet att svara då *övrigt bortfall* innebär *förhindrad medverkan*.

Förklarande variabler

Följande registervariabler är använda som förklarande variabler (se tabell 1):

Tabell 1 Förklarande variabler

Variabel (benämning)	Kategorier (koder)
Kön	1 = Man, 0= Kvinna
Ålder	(Ålder: år) 1 = 18 – 24 2 = 25 – 34 3 = 35 – 44 4 = 45 – 54 5 = 55 – 64 6 = 65 – 79 Dummy-variabler för respektive kategori

Födelseland	1 = Födda i Sverige 2 = Födda i Norden 3 = Födda i övriga världen Dummy-variabler för respektive kategori
Utbildningsnivå	1 = Låg 2 = Medel 3 = Hög Dummy-variabler för respektive kategori

Resultat

I detta avsnitt redovisas resultaten av SAS-körningarna med modellerna Y1-Y2c. För att tolka modellerna används oddskvoten som antar värden större än 0. Värden större än 1 innebär att oddset, dvs att sannolikheten för det inträffade, ökar för den aktuella gruppen när övriga grupper är konstanta. Värden mindre än 1 motsvarar ett minskat odds eller sannolikhet.

Bortfallet fördelat på förklaringsvariabler

I tabell 2 visas antalet och andelen för de olika bortfallsvariablerna i urvalet fördelat på kategorier för förklaringsvariablerna. Det totala bortfallet (minus övertäckning är 45,2 procent. För de enskilda förklaringsvariablerna minskar bortfallsandelen med ökad utbildningsnivå och i viss mån för ökande ålder. Något större bortfall för män än för kvinnor. Det förekommer också relativt stor skillnad i svarsfrekvens för de olika födelselandsgrupperna. Liknande mönster finns då vi jämför de ej anträffade mot de anträffade.

Kategorin *Övrigt bortfall* är relativt liten hos de flesta grupper. Dock sticker grupperna *låg utbildning* och *född i övriga världen* ut.

Av de som man kom i kontakt med avböjden 28,8 procent medverkan. Mönstret för del olika förklaringsvariablerna liknar de andra bortfallsvariablerna förutom att *Födda i övriga Norden* har den högsta andelen 35,9 procent bland födelselandsgrupperna. Skillnaden mellan de olika åldersgrupperna är också mindre.

Tabell 2. Bortfallet fördelat på förklaringsvariabler

Bortfallet fördelat på förklaringsvariabler, Michael Franzén, SCB 2014-01-23, GGS_analys.sas

The FREQ Procedure

Frequency Row Pct		Totalt bortfall			Ej anträffad			Övrigt bortfall			Avböjd medverkan		
		0	1	Total	0	1	Total	0	1	Total	0	1	Total
Kön	Kvinnor	4991	3804	8795	7283	1512	8795	6843	440	7283	4991	1852	6843
		56,8	43,3		82,8	17,2		94,0	6,0		72,9	27,1	
	Män	4697	4182	8879	7190	1689	8879	6773	417	7190	4697	2076	6773
		52,9	47,1		81,0	19,0		94,2	5,8		69,4	30,7	
Åldersgrupp	18-24	1152	1096	2248	1652	596	2248	1578	74	1652	1152	426	1578
		51,3	48,8		73,5	26,5		95,5	4,5		73,0	27,0	
	25-34	1393	1435	2828	2009	819	2828	1859	150	2009	1393	466	1859
		49,3	50,7		71,0	29,0		92,5	7,5		74,9	25,1	
	35-44	1745	1418	3163	2572	591	3163	2441	131	2572	1745	696	2441
		55,2	44,8		81,3	18,7		94,9	5,1		71,5	28,5	
	45-54	1764	1417	3181	2661	520	3181	2554	107	2661	1764	790	2554
		55,5	44,6		83,7	16,4		96,0	4,0		69,1	30,9	
	55-64	1681	1238	2919	2554	365	2919	2413	141	2554	1681	732	2413
		57,6	42,4		87,5	12,5		94,5	5,5		69,7	30,3	
	65-79	1953	1382	3335	3025	310	3335	2771	254	3025	1953	818	2771
		58,6	41,4		90,7	9,3		91,6	8,4		70,5	29,5	
Födelselandsgrupp	Sverige	8521	6177	14698	12292	2406	14698	11948	344	12292	8521	3427	11948
		58,0	42,0		83,6	16,4		97,2	2,8		71,3	28,7	
	Övr Norden	257	273	530	432	98	530	401	31	432	257	144	401
		48,5	51,5		81,5	18,5		92,8	7,2		64,1	35,9	
	Övr världen	910	1536	2446	1749	697	2446	1267	482	1749	910	357	1267
37,2		62,8		71,5	28,5		72,4	27,6		71,8	28,2		
Utbildningsnivå	Låg	1831	2520	4351	3382	969	4351	2898	484	3382	1831	1067	2898
		42,1	57,9		77,7	22,3		85,7	14,3		63,2	36,8	
	Mellan	4140	3631	7771	6294	1477	7771	6067	227	6294	4140	1927	6067
		53,3	46,7		81,0	19,0		96,4	3,6		68,2	31,8	
	Hög	3717	1835	5552	4797	755	5552	4651	146	4797	3717	934	4651
67,0		33,1		86,4	13,6		97,0	3,0		79,9	20,1		
Total	Total	9688	7986	17674	14473	3201	17674	13616	857	14473	9688	3928	13616
		54,8	45,2		81,9	18,1		94,1	5,9		71,2	28,8	

Modeller

Nedan följer resultat från SAS för de fyra olika modellerna.

Figur 1. SAS-utskrift för modell Y1

Logistisk regression modell Y1, Michael Franzén, SCB 2014-01-22

The LOGISTIC Procedure

Model Information	
Data Set	WORK.GGS_ANALYS
Response Variable	Totalt bortfall
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	17674
Number of Observations Used	17674

Response Profile		
Ordered Value	Totalt bortfall	Total Frequency
1	1	7986
2	0	9688

Probability modeled is Totalt bortfall=1.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	24339.210	23253.585
SC	24346.990	23339.164
-2 Log L	24337.210	23231.585

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1105.6249	10	<.0001
Score	1076.4253	10	<.0001
Wald	1007.3449	10	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1974	0.0443	19.8351	<.0001
Man	1	0.1069	0.0313	11.6643	0.0006
18-24 år	1	-0.1274	0.0587	4.7074	0.0300
25-34 år	1	0.2145	0.0538	15.8732	<.0001
45-54 år	1	-0.0364	0.0521	0.4874	0.4851
55-64 år	1	-0.1718	0.0538	10.1977	0.0014
65-79 år	1	-0.3678	0.0539	46.6251	<.0001
Född i Övr Norden	1	0.4005	0.0910	19.3719	<.0001
Född i Övr världen	1	0.7602	0.0469	262.8245	<.0001
Låg utbildning	1	0.4680	0.0407	132.1762	<.0001
Hög utbildning	1	-0.6500	0.0376	299.5782	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Man	1.113	1.047	1.183
18-24 år	0.880	0.785	0.988
25-34 år	1.239	1.115	1.377
45-54 år	0.964	0.871	1.068
55-64 år	0.842	0.758	0.936
65-79 år	0.692	0.623	0.769
Född i Övr Norden	1.493	1.249	1.784
Född i Övr världen	2.139	1.951	2.345
Låg utbildning	1.597	1.474	1.729
Hög utbildning	0.522	0.485	0.562

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	62.7	Somers' D	0.278
Percent Discordant	34.9	Gamma	0.285
Percent Tied	2.4	Tau-a	0.138
Pairs	77368368	c	0.639

Av oddskvoterna att döma ökar sannolikheten för bortfall för män jämfört med kvinnor något. Sannolikheten för bortfall minskar med ökande åldergrupp i jämförelse med referensgruppen 35-44 år, och ökar för åldergruppen under 25-34 år. Den yngsta åldergruppen har dock mindre odds för bortfall. Sannolikheten för bortfall ökar avsevärt för de som är födda i övriga världen. Sannolikheten för bortfall är nästan 60 procent högre för lågutbildade och 50 procent lägre för högutbildade jämfört med de med medelhög utbildningsnivå.

Figur 2. SAS-utskrift för modell Y2a

Logistisk regression modell Y2a, Michael Franzén, SCB 2014-01-22

The LOGISTIC Procedure

Model Information	
Data Set	WORK.GGS_ANALYS
Response Variable	Ej anträffad
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	17674
Number of Observations Used	17674

Response Profile		
Ordered Value	Ej anträffad	Total Frequency
1	1	3201
2	0	14473

Probability modeled is Ej anträffad=1.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	16724.352	15854.444
SC	16732.132	15940.023
-2 Log L	16722.352	15832.444

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	889.9077	10	<.0001
Score	893.9010	10	<.0001
Wald	829.9379	10	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4896	0.0562	703.2313	<.0001
Man	1	0.0798	0.0403	3.9177	0.0478
18-24 år	1	0.2918	0.0692	17.7660	<.0001
25-34 år	1	0.5594	0.0626	79.7441	<.0001
45-54 år	1	-0.1749	0.0669	6.8279	0.0090
55-64 år	1	-0.5178	0.0734	49.7506	<.0001
65-79 år	1	-0.9374	0.0782	143.7007	<.0001
Född i Övr Norden	1	0.3733	0.1172	10.1491	0.0014
Född i Övr världen	1	0.5497	0.0527	108.7837	<.0001
Låg utbildning	1	0.2474	0.0503	24.1510	<.0001
Hög utbildning	1	-0.4773	0.0503	90.0889	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Man	1.083	1.001	1.172
18-24 år	1.339	1.169	1.533
25-34 år	1.750	1.547	1.978
45-54 år	0.840	0.736	0.957
55-64 år	0.596	0.516	0.688
65-79 år	0.392	0.336	0.457
Född i Övr Norden	1.453	1.154	1.828
Född i Övr världen	1.733	1.563	1.921
Låg utbildning	1.281	1.160	1.413
Hög utbildning	0.620	0.562	0.685

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	65.4	Somers' D	0.333
Percent Discordant	32.1	Gamma	0.341
Percent Tied	2.6	Tau-a	0.099
Pairs	46328073	c	0.666

Det finns i princip ingen skillnad mellan kvinnor och män i regressionen för *ej anträffade*. Jämfört med åldersklassen 35-44 år har de lägre åldersklasserna högre oddskvoter och högre sannolikhet att vara ej anträffade, medan de högre åldersklasserna har lägre odds och sannolikhet att vara ej anträffade. Sannolikheten för ej anträffad ökar för de som inte är födda i Sverige. Oddskvoten är större för lågutbildade och lägre för högutbildade jämfört med de med medelhög utbildningsnivå.

Figur 3. SAS-utskrift för modell Y2b

Logistisk regression modell Y2b, Michael Franzén, SCB 2014-01-22

The LOGISTIC Procedure

Model Information	
Data Set	WORK.GGS_ANALYS
Response Variable	Övrigt bortfall
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	14473
Number of Observations Used	14473

Response Profile		
Ordered Value	Övrigt bortfall	Total Frequency
1	1	857
2	0	13616

Probability modeled is Övrigt bortfall=1.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	6509.018	5007.516
SC	6516.598	5090.897
-2 Log L	6507.018	4985.516

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1521.5012	10	<.0001
Score	2139.1656	10	<.0001
Wald	1323.8948	10	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.1144	0.1312	983.3432	<.0001
Man	1	-0.0308	0.0779	0.1561	0.6928
18-24 år	1	-0.3979	0.1664	5.7169	0.0168
25-34 år	1	0.2087	0.1388	2.2608	0.1327
45-54 år	1	-0.0831	0.1463	0.3230	0.5698
55-64 år	1	0.2776	0.1398	3.9448	0.0470
65-79 år	1	0.6641	0.1302	25.9946	<.0001
Född i Övr Norden	1	0.7423	0.1997	13.8170	0.0002
Född i Övr världen	1	2.6400	0.0849	966.5761	<.0001
Låg utbildning	1	1.2370	0.0919	181.3831	<.0001
Hög utbildning	1	-0.3731	0.1132	10.8678	0.0010

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Man	0.970	0.832	1.130
18-24 år	0.672	0.485	0.931
25-34 år	1.232	0.939	1.617
45-54 år	0.920	0.691	1.226
55-64 år	1.320	1.004	1.736
65-79 år	1.943	1.505	2.508
Född i Övr Norden	2.101	1.420	3.107
Född i Övr världen	14.013	11.865	16.550
Låg utbildning	3.445	2.878	4.125
Hög utbildning	0.689	0.552	0.860

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	83.4	Somers' D	0.693
Percent Discordant	14.2	Gamma	0.710
Percent Tied	2.4	Tau-a	0.077
Pairs	11668912	c	0.846

Det finns ingen skillnad mellan könen vad det gäller odds för *övrigt bortfall*. Sannolikheten för *övrigt bortfall* är lågt för unga 18-24 år och högt för äldre 65-79 år. Det förefaller inte vara några större skillnader mellan de andra åldersgrupperna. Det är stora skillnader mellan de olika födelselandsgrupperna. Sannolikheten för *övrigt bortfall* är dubbelt så stor för personer födda i övriga nordnorden och mångdubbelt större för de födda i övriga världen. En stor skillnad finns också för de lågutbildade jämfört med de med medelhög utbildningsnivå.

Figur 4. SAS-utskrift för modell Y2c

Logistisk regression modell Y2c, Michael Franzén, SCB 2014-01-22
The LOGISTIC Procedure

Model Information	
Data Set	WORK.GGS_ANALYS
Response Variable	Avböjd medverkan
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	13616
Number of Observations Used	13616

Response Profile		
Ordered Value	Avböjd medverkan	Total Frequency
1	1	3928
2	0	9688

Probability modeled is Avböjd medverkan=1.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	16362.681	16017.338
SC	16370.200	16100.047
-2 Log L	16360.681	15995.338

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	365.3435	10	<.0001
Score	356.8244	10	<.0001
Wald	347.8682	10	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7605	0.0541	197.3013	<.0001
Man	1	0.1289	0.0385	11.1928	0.0008
18-24 år	1	-0.3788	0.0757	25.0163	<.0001
25-34 år	1	-0.1501	0.0708	4.4886	0.0341
45-54 år	1	0.0733	0.0629	1.3567	0.2441
55-64 år	1	-0.0159	0.0644	0.0609	0.8051
65-79 år	1	-0.1909	0.0648	8.6735	0.0032
Född i Övr Norden	1	0.2724	0.1082	6.3384	0.0118
Född i Övr världen	1	-0.0184	0.0672	0.0749	0.7843
Låg utbildning	1	0.3056	0.0502	37.0949	<.0001
Hög utbildning	1	-0.6295	0.0464	183.9439	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Man	1.138	1.055	1.227
18-24 år	0.685	0.590	0.794
25-34 år	0.861	0.749	0.989
45-54 år	1.076	0.951	1.217
55-64 år	0.984	0.867	1.117
65-79 år	0.826	0.728	0.938
Född i Övr Norden	1.313	1.062	1.623
Född i Övr världen	0.982	0.861	1.120
Låg utbildning	1.357	1.230	1.498
Hög utbildning	0.533	0.487	0.584

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	58.6	Somers' D	0.204
Percent Discordant	38.2	Gamma	0.211
Percent Tied	3.2	Tau-a	0.084
Pairs	38054464	c	0.602

För de som har haft möjlighet att svara (avböjd medverkan och svarande) avböjde män i större utsträckning än kvinnor. Sannolikheten var 13 procent högre för män att avböja medverkan än för kvinnor. Det verkar inte finnas något tydligt mönster mellan de olika åldersgrupperna vad det gäller benägenhet att avböja medverkan. Den yngsta åldersgruppen har dock lägre oddskvot. Det var omkring 30 procent högre sannolikhet för avböjd medverkan för födda i övriga Norden. Det var tydlig skillnad mellan utbildningsnivåerna. Lågutbildade hade klart högre sannolikhet att avböja medverkan och högutbildade klart lägre sannolikhet att avböja medverkan.

Bilaga 1. SAS-program för körning av logit-modeller

```

/*=====
Bereder data
-----*/

data GGS_analys;
  set ggs.D3;
  resk=substr(resultatkod,1,1);
  if SUN2000Niva_old=' ' then SUN2000Niva_old=1; /* om
uppgift saknas imputerar vi förgymnasial */

  *Kön;
  if kon=2 then kon=0;
  kon_=kon*1;

  *Utbildningsnivå;
  if      1<=SUN2000Niva_old<=2 then utb=1;
  else if 3<=SUN2000Niva_old<=4 then utb=2;
  else                                     utb=3;

  if 1<=SUN2000Niva_old<=2 then utb_1=1; else utb_1=0;
  if 3<=SUN2000Niva_old<=4 then utb_2=1; else utb_2=0;
  if SUN2000Niva_old>=5     then utb_3=1; else utb_3=0;

  *Åldersgrupp;
  if      18<=alder<= 24 then aldergrupp=1;
  else if 25<=alder<= 34 then aldergrupp=2;
  else if 35<=alder<= 44 then aldergrupp=3;
  else if 45<=alder<= 54 then aldergrupp=4;
  else if 55<=alder<= 64 then aldergrupp=5;
  else                                     aldergrupp=6;

  aldergrupp_1=(aldergrupp=1);
  aldergrupp_2=(aldergrupp=2); aldergrupp_3=(aldergrupp=3);
  aldergrupp_4=(aldergrupp=4); aldergrupp_5=(aldergrupp=5);
  aldergrupp_6=(aldergrupp=6);

  *Födelselandgrupp;
  if Fodelselandnamn in ('SVERIGE') then FodelselandGrp=1;
  else if Fodelselandnamn in
('NORGE', 'DANMARK', 'FINLAND', 'ISLAND') then
FodelselandGrp=2;
  else FodelselandGrp=3;

  if Fodelselandnamn in ('SVERIGE')
then FodelselandSverige=1; else FodelselandSverige=0;
  if Fodelselandnamn in ('NORGE', 'DANMARK', 'FINLAND',
'ISLAND') then FodelselandNorden=1; else
FodelselandNorden=0;
  if Fodelselandnamn in ('SVERIGE' 'NORGE' 'DANMARK'
'FINLAND' 'ISLAND') then FodelselandOvriga=0; else
FodelselandOvriga=1;

  if resk=5 then ejant=1; else ejant=0;           *Ej
anträffad;
  if resk=6 then ejmed=1; else ejmed=0;         *Avböjd
medverkan;

```



```

        if resk=4 then ovrbf=1; else ovrbf=0;           *Övrigt
    bortfall;
        if 1<resk<9 then ej svar=1; else ej svar=0;   *Totalt
    bortfall;

    format kon_ KonFmt. aldergrupp Alder6k. utb   UtbNiva.
    FodselandGrp Fodland.;
run;

/*=====
Bortfallet fördelat på förklaringsvariabler
-----*/

%let set = GGS_analys.sas;
data _null_;
    call symput ('dag',put (date (), yymmdd10.));
run;
%put &set;
%put &dag;

ods html path="P:\Data\PCA\MIH\MichaelF\GGS\Logistisk
Regression Bortfall"
body="ObservationerPerGrupp &dag..html"
style=Styles.Minimal ;
title "Bortfallet fördelat på förklaringsvariabler, Michael
Franzén, SCB &dag, &set";

proc freq data=GGS_analys (where=(resk<'9')
rename=(ej svar='Totalt bortfall'n kon_='Kön'n
aldergrupp='Åldersgrupp'n
FodselandGrp='Födelselandsgrupp'n
utb='Utbildningsnivå'n));
tables 'Kön'n*'Totalt bortfall'n 'Åldersgrupp'n*'Totalt
bortfall'n 'Födelselandsgrupp'n*'Totalt bortfall'n
'Utbildningsnivå'n*'Totalt bortfall'n / nocol nopercnt ;
run;

proc freq data=GGS_analys (where=(resk<'9')
rename=(ej ant='Ej anträffad'n kon_='Kön'n
aldergrupp='Åldersgrupp'n
FodselandGrp='Födelselandsgrupp'n
utb='Utbildningsnivå'n));
tables 'Kön'n*'Ej anträffad'n 'Åldersgrupp'n*'Ej
anträffad'n 'Födelselandsgrupp'n*'Ej anträffad'n
'Utbildningsnivå'n*'Ej anträffad'n / nocol nopercnt ;
run;

proc freq data=GGS_analys (where=(resk<'9' and ej ant=0)
rename=(ovrbf='Övrigt bortfall'n kon_='Kön'n
aldergrupp='Åldersgrupp'n
FodselandGrp='Födelselandsgrupp'n
utb='Utbildningsnivå'n));
tables 'Kön'n*'Övrigt bortfall'n 'Åldersgrupp'n*'Övrigt
bortfall'n 'Födelselandsgrupp'n*'Övrigt bortfall'n
'Utbildningsnivå'n*'Övrigt bortfall'n / nocol nopercnt;
run;

proc freq data=GGS_analys (where=(resk<'9' and ej ant=0 and
ovrbf=0) rename=(ej med='Avböjd medverkan'n kon_='Kön'n

```

```

aldergrupp='Åldersgrupp'n
FodselandGrp='Födelselandsgrupp'n
utb='Utbildningsnivå'n));
tables 'Kön'n*'Avböjd medverkan'n 'Åldersgrupp'n*'Avböjd
medverkan'n 'Födelselandsgrupp'n*'Avböjd medverkan'n
'Utbildningsnivå'n*'Avböjd medverkan'n / nocol nopercnt;
run;

ods html close;
title; footnote;
ods listing;

/*=====
Modeller - Skriver ut resultat till html-filer via ODS
-----*/

ods html path="P:\Data\PCA\MIH\MichaelF\GGS\Logistisk
Regression Bortfall"
body="Logit Y0 &dag;.html" style=Styles.Minimal ;
title "Logistisk regression modell Y1, Michael Franzén, SCB
&dag";
proc logistic data=ggs_analys (where=(resk<'9')
rename=(ejsvar='Totalt bortfall'n kon_='Man'n
aldergrupp_1='18-24 år'n aldergrupp_2='25-34 år'n
aldergrupp_4='45-54 år'n
aldergrupp_5='55-64 år'n aldergrupp_6='65-79 år'n
FodselandNorden='Född i Övr Norden'n
FodselandOvriga='Född i Övr världen'n
utb_1='Låg utbildning'n utb_3='Hög utbildning'n))
descending;

    model 'Totalt bortfall'n = 'Man'n '18-24 år'n '25-34
år'n '45-54 år'n '55-64 år'n '65-79 år'n 'Född i Övr
Norden'n 'Född i Övr världen'n 'Låg utbildning'n 'Hög
utbildning'n;

run;
ods html close;
title; footnote;
ods listing;

ods html path="P:\Data\PCA\MIH\MichaelF\GGS\Logistisk
Regression Bortfall"
body="Logit Y1 &dag;.html" style=Styles.Minimal ;
title "Logistisk regression modell Y2a, Michael Franzén,
SCB &dag";
proc logistic data=ggs_analys (where=(resk<'9')
rename=(ejant='Ej anträffad'n kon_='Man'n
aldergrupp_1='18-24 år'n aldergrupp_2='25-34 år'n
aldergrupp_4='45-54 år'n
aldergrupp_5='55-64 år'n aldergrupp_6='65-79 år'n
FodselandNorden='Född i Övr Norden'n
FodselandOvriga='Född i Övr världen'n
utb_1='Låg utbildning'n utb_3='Hög utbildning'n))
descending;

    model 'Ej anträffad'n = 'Man'n '18-24 år'n '25-34 år'n
'45-54 år'n '55-64 år'n '65-79 år'n 'Född i Övr Norden'n

```

```
'Född i Övr världen'n 'Låg utbildning'n 'Hög utbildning'n;
*/rsq lackfit;
```

```
run;
ods html close;
title; footnote;
ods listing;
```

```
ods html path="P:\Data\PCA\MIH\MichaelF\GGS\Logistisk
Regression Bortfall"
```

```
body="Logit Y2 &dag..html" style=Styles.Minimal ;
title "Logistisk regression modell Y2b, Michael Franzén,
SCB &dag";
```

```
proc logistic data=ggs_analys (where=(resk<'9' and ejant=0)
rename=(ovrbf='Övrigt bortfall'n kon_='Man'n
aldergrupp_1='18-24 år'n aldergrupp_2='25-34 år'n
aldergrupp_4='45-54 år'n
aldergrupp_5='55-64 år'n aldergrupp_6='65-79 år'n
FodselandNorden='Född i Övr Norden'n
FodselandOvriga='Född i Övr världen'n
utb_1='Låg utbildning'n utb_3='Hög utbildning'n))
descending;
```

```
model 'Övrigt bortfall'n = 'Man'n '18-24 år'n '25-34
år'n '45-54 år'n '55-64 år'n '65-79 år'n 'Född i Övr
Norden'n 'Född i Övr världen'n 'Låg utbildning'n 'Hög
utbildning'n;
```

```
run;
ods html close;
title; footnote;
ods listing;
```

```
ods html path="P:\Data\PCA\MIH\MichaelF\GGS\Logistisk
Regression Bortfall"
```

```
body="Logit Y3 &dag..html" style=Styles.Minimal ;
title "Logistisk regression modell Y2c, Michael Franzén,
SCB &dag";
```

```
proc logistic data=ggs_analys (where=(resk<'9' and ejant=0
and ovrbf=0) rename=(ejmed='Avböjd medverkan'n kon_='Man'n
aldergrupp_1='18-24 år'n aldergrupp_2='25-34 år'n
aldergrupp_4='45-54 år'n
aldergrupp_5='55-64 år'n aldergrupp_6='65-79 år'n
FodselandNorden='Född i Övr Norden'n
FodselandOvriga='Född i Övr världen'n
utb_1='Låg utbildning'n utb_3='Hög utbildning'n))
descending;
```

```
model 'Avböjd medverkan'n = 'Man'n '18-24 år'n '25-34
år'n '45-54 år'n '55-64 år'n '65-79 år'n 'Född i Övr
Norden'n 'Född i Övr världen'n 'Låg utbildning'n 'Hög
utbildning'n;
```

```
run;
ods html close;
title; footnote;
ods listing;
```