

Likelihood ratio test.....	2
Collinearity	4
collin	4
Standardized residuals and leverage	4
Standardized residuals	5
Leverage	5
listcoef.....	6
prtab.....	7
prchange.....	8
prvalue.....	10
Plots.....	11
postgr3	11
Plot with confidence	12
Plot with confidence for groups.....	13

Likelihood ratio test

tabulate result

Result	Freq.	Percent	Cum.
died	63	48.09	48.09
survived	68	51.91	100.00
Total	131	100.00	

logistic result gender age age height cause duration oxy_conc intens

note: age dropped because of collinearity

Logistic regression
 Log likelihood = -75.740674

Number of obs = 131
 LR chi2(7) = 29.93
 Prob > chi2 = 0.0001
 Pseudo R2 = 0.1650

result	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
gender	.8081534	.3814021	-0.45	0.652	.3204612 2.038037
age	.9642082	.0167886	-2.09	0.036	.9318583 .9976812
height	1.036385	.0199122	1.86	0.063	.9980839 1.076157
cause	1.285372	.3947316	0.82	0.414	.704091 2.346544
duration	.9349991	.0283636	-2.22	0.027	.8810278 .9922766
oxy_conc	.0807813	.1109453	-1.83	0.067	.0054735 1.192214
intens	.965642	.0223696	-1.51	0.131	.9227789 1.010496

estimates store full

logistic result age height duration oxy_conc

Logistic regression
 Log likelihood = -77.316925

Number of obs = 131
 LR chi2(4) = 26.78
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.1476

result	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age	.9634765	.0166204	-2.16	0.031	.9314457 .9966088
height	1.044679	.0177059	2.58	0.010	1.010546 1.079965
duration	.9397322	.0271166	-2.15	0.031	.8880596 .9944113
oxy_conc	.0483873	.0630016	-2.33	0.020	.0037709 .6208969

estimates store short

lrtest full short < lrtest full . > < lrtest full >

Likelihood-ratio test
 (Assumption: . nested in full)

LR chi2(3) = 3.15
 Prob > chi2 = 0.3687

```
gen oxygt08=oxy_conc>0.8          /* ←-- in order to have a categorical variable */
```

```
label define oxygt08 1 "0.8+" 0 "<=0.8"  
label value oxygt08 oxygt08
```

```
tabulate oxygt08 result, row exact
```

oxygt08	result		Total
	died	survived	
<=0.8	14 31.82	30 68.18	44 100.00
0.8+	49 56.32	38 43.68	87 100.00
Total	63 48.09	68 51.91	131 100.00

```
Fisher's exact = 0.010  
1-sided Fisher's exact = 0.007
```

```
logistic result age height duration oxygt08
```

```
Logistic regression          Number of obs = 131  
                             LR chi2(4) = 29.56  
                             Prob > chi2 = 0.0000  
Log likelihood = -75.925772   Pseudo R2 = 0.1630
```

result	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age	.9637965	.0168108	-2.11	0.035	.9314047 .9973148
height	1.044083	.0182429	2.47	0.014	1.008933 1.080458
duration	.9373933	.0269203	-2.25	0.024	.886088 .9916693
oxygt08	.2945295	.1285616	-2.80	0.005	.1251932 .6929099

Collinearity collin

collin age height duration oxygt08

Collinearity Diagnostics

Variable	VIF	SQRT VIF	Tolerance	R- Squared
age	1.09	1.04	0.9171	0.0829
height	1.12	1.06	0.8966	0.1034
duration	1.08	1.04	0.9301	0.0699
oxygt08	1.01	1.01	0.9873	0.0127
Mean VIF	1.07			

	Eigenval	Cond Index
1	4.3499	1.0000
2	0.3428	3.5623
3	0.2147	4.5007
4	0.0896	6.9681
5	0.0029	38.4516

Condition Number 38.4516
 Eigenvalues & Cond Index computed from scaled raw sscp (w/ intercept)
 Det(correlation matrix) 0.8755

Standardized residuals and leverage

Leverage = dbeta = Pregibon Delta-Beta influence statistic = counterpart to Cook'distance in linear regression

"We have seen quite a few logistic regression diagnostic statistics. Now how large does each one have to be, to be considered influential?"

First of all, we always have to make our judgment based on our theory and our analysis.
 Secondly, there are some rule-of-thumb cutoffs when the sample size is large. These are shown below.

When the sample size is large, the asymptotic distribution of some of the measures would follow some standard distribution.

That is why we have these cutoff values, and why they only apply when the sample size is large enough.

Usually, we would look at the relative magnitude of a statistic an observation has compared to others.

That is, we look for data points that are farther away from most of the data points. "

Measure	Value
leverage (hat value)	>2 or 3 times of the average of leverage
abs(Pearson Residuals)	> 2
abs(Deviance Residuals)	> 2

<http://www.ats.ucla.edu/stat/Stata/webbooks/logistic/chapter3/statalog3.htm>

Standardized residuals

```
predict prob
label var prob "Prob(survive)"
```

```
predict stdres, res
label var std "Std residual"
```

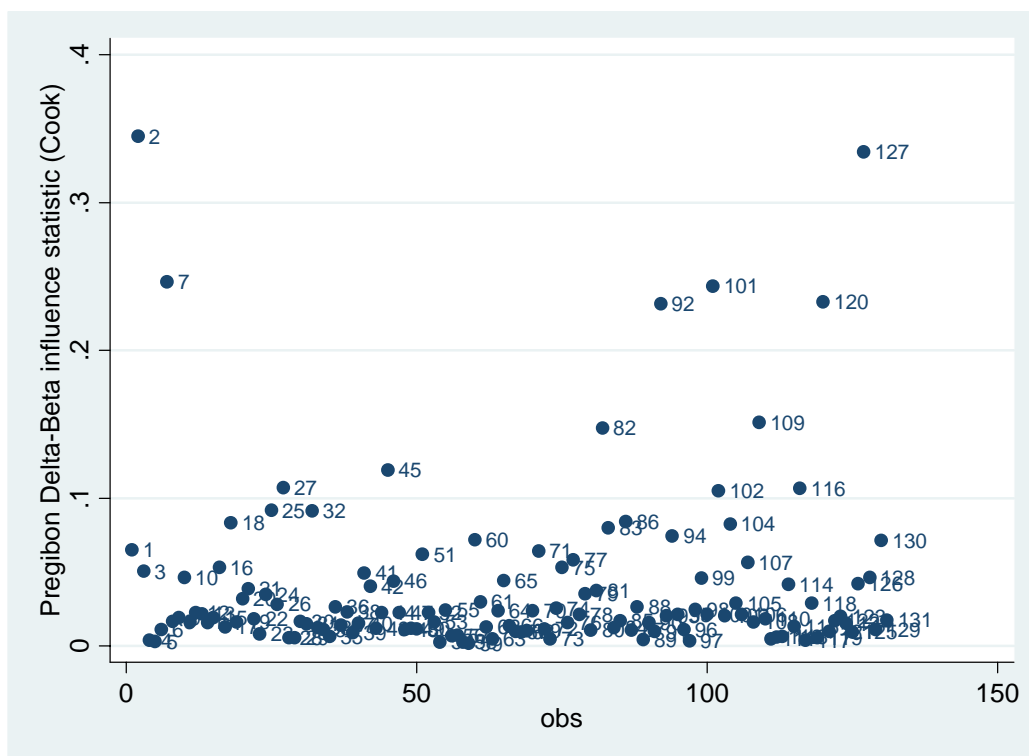
```
gen predict="survived" if prob>0.5
replace predict="died" if prob<=0.5
```

```
list prob predict result stdres age height duration oxygt08 if abs(stdres)>2
```

	prob	predict	result	stdres	age	height	duration	oxygt08
2.	.1455186	died	survived	2.423215	35	170	33	0.8+
7.	.8008081	survived	died	-2.005065	21	185	21	<=0.8
92.	.1790529	died	survived	2.141248	23	140	16	0.8+

Leverage

```
predict cook, dbeta
label var cook "Pregibon Delta-Beta influence statistic (Cook)"
gen obs=_n
scatter cook obs, mlabel(obs)
```



listcoef

listcoef, help

logistic (N=131): Factor Change in Odds
 Odds of: survived vs died

result	b	z	P> z	e^b	e^bStdX	SDofX
age	-0.03688	-2.114	0.035	0.9638	0.6310	12.4865
height	0.04314	2.469	0.014	1.0441	1.8123	13.7834
duration	-0.06465	-2.251	0.024	0.9374	0.5794	8.4412
oxygt08	-1.22238	-2.800	0.005	0.2945	0.5602	0.4741

b = raw coefficient
 z = z-score for test of b=0
 P>|z| = p-value for z-test
 e^b = exp(b) = factor change in odds for unit increase in X
 e^bStdX = exp(b*SD of X) = change in odds for SD increase in X
 SDofX = standard deviation of X

listcoef, reverse

logistic (N=131): Factor Change in Odds
 Odds of: died vs survived

result	b	z	P> z	e^b	e^bStdX	SDofX
age	-0.03688	-2.114	0.035	1.0376	1.5848	12.4865
height	0.04314	2.469	0.014	0.9578	0.5518	13.7834
duration	-0.06465	-2.251	0.024	1.0668	1.7259	8.4412
oxygt08	-1.22238	-2.800	0.005	3.3952	1.7852	0.4741

listcoef, percent help

logistic (N=131): Percentage Change in Odds
 Odds of: survived vs died

result	b	z	P> z	%	%StdX	SDofX
age	-0.03688	-2.114	0.035	-3.6	-36.9	12.4865
height	0.04314	2.469	0.014	4.4	81.2	13.7834
duration	-0.06465	-2.251	0.024	-6.3	-42.1	8.4412
oxygt08	-1.22238	-2.800	0.005	-70.5	-44.0	0.4741

b = raw coefficient
 z = z-score for test of b=0
 P>|z| = p-value for z-test
 % = percent change in odds for unit increase in X
 %StdX = percent change in odds for SD increase in X
 SDofX = standard deviation of X

listcoef, reverse percent

logistic (N=131): Percentage Change in Odds
 Odds of: died vs survived

result	b	z	P> z	%	%StdX	SDofX
age	-0.03688	-2.114	0.035	3.8	58.5	12.4865
height	0.04314	2.469	0.014	-4.2	-44.8	13.7834
duration	-0.06465	-2.251	0.024	6.7	72.6	8.4412
oxygt08	-1.22238	-2.800	0.005	239.5	78.5	0.4741

prtab

prtab duration oxygt08

logistic: Predicted probabilities of positive outcome for result

```
-----  
duration |      oxygt08  
          |      <=0.8      0.8+  
-----+-----  
    3 | 0.8215  0.5754  
    4 | 0.8118  0.5595  
    5 | 0.8017  0.5435  
    6 | 0.7912  0.5275  
    7 | 0.7803  0.5113  
    8 | 0.7691  0.4952  
    9 | 0.7574  0.4790  
   10 | 0.7453  0.4629  
   11 | 0.7328  0.4469  
   12 | 0.7200  0.4309  
   13 | 0.7068  0.4152  
   14 | 0.6932  0.3996  
   15 | 0.6793  0.3842  
   16 | 0.6650  0.3690  
   17 | 0.6505  0.3541  
   18 | 0.6356  0.3394  
   19 | 0.6205  0.3251  
   20 | 0.6052  0.3110  
   21 | 0.5896  0.2974  
   22 | 0.5739  0.2840  
   23 | 0.5580  0.2711  
   24 | 0.5420  0.2585  
   25 | 0.5260  0.2463  
   26 | 0.5098  0.2345  
   30 | 0.4454  0.1913  
   33 | 0.3981  0.1631  
   35 | 0.3676  0.1462  
   36 | 0.3527  0.1383  
   38 | 0.3238  0.1236  
   62 | 0.0921  0.0290  
-----
```

```
          age      height  duration  oxygt08  
x= 29.870229  168.73282  12.992366  .66412214
```

prchange

```
summarize age height duration oxygt08
. summarize age height duration oxygt08
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	131	29.87023	12.48654	2	61
height	131	168.7328	13.78339	90	190
duration	131	12.99237	8.441196	3	62
oxygt08	131	.6641221	.4741095	0	1

prchange, help

logistic: Changes in Probabilities for result

	min->max	0->1	-+1/2	-+sd/2	MargEfct ¹
age	-0.4959	-0.0067	-0.0092	-0.1145	-0.0092
height	0.6936	0.0000	0.0108	0.1474	0.0108
duration	-0.6283	-0.0134	-0.0161	-0.1354	-0.0161
oxygt08	-0.2916	-0.2916	-0.2961	-0.1437	-0.3052

```
Pr(y|x)      died  survived
             0.4829  0.5171
```

Proof that
 $x=(age=mean\ height=mean\ duration=mean\ oxygt08=mean)$

```
          age  height  duration  oxygt08
x=  29.8702  168.733  12.9924  .664122
sd(x)=  12.4865  13.7834  8.4412  .474109
```

Pr(y|x): probability of observing each y for specified x values
 Avg|Chg|: average of absolute value of the change across categories
 Min->Max: change in predicted probability as x changes from its minimum to its maximum
 0->1: change in predicted probability as x changes from 0 to 1
 -+1/2: change in predicted probability as x changes from 1/2 unit below base value to 1/2 unit above
 -+sd/2: change in predicted probability as x changes from 1/2 standard dev below base to 1/2 standard dev above
 MargEfct: the partial derivative of the predicted probability/rate with respect to a given independent variable

```
prvalue, x(age=29.87023 height=168.7328 duration=12.99237 oxygt08=.6641221)
```

logistic: Predictions for result
 Confidence intervals by delta method

		95% Conf. Interval	
Pr(y=survived x):	0.5171	[0.4202,	0.6140]
Pr(y=died x):	0.4829	[0.3860,	0.5798]

```
          age  height  duration  oxygt08
x=  29.87023  168.7328  12.99237  .6641221
```

```
prchange, x(age=20 height=170) rest(median)
```

logistic: Changes in Probabilities for result

	min->max	0->1	-+1/2	-+sd/2	MargEfct
age	-0.4915	-0.0075	-0.0091	-0.1134	-0.0091
height	0.7068	0.0000	0.0107	0.1461	0.0107
duration	-0.6297	-0.0134	-0.0160	-0.1342	-0.0160
oxygt08	-0.2554	-0.2554	-0.2936	-0.1424	-0.3024

```
Pr(y|x)      died  survived
             0.4489  0.5511
```

```
          age  height  duration  oxygt08
x=         20    170         11         1
sd(x)=  12.4865  13.7834  8.4412  .474109
```

¹ Derivative of propability curve (which is equal to a one unit change if the curve is linear) at the value of x (either mean for all variables or given by x() and rest)
 Guido Lüchters 12/07

This is the suggested and correct method to calculate probabilities for a 10 years step at the mean of age.
 All other values were also recalculated for the new variable ~ age/10
 which is = probability(age)*10 only if age is uniformly distributed

prchange age, delta(10) help

logistic: Changes in Probabilities for result

(Note: d = 10)

	min->max	0->1	++d/2	++sd/2	MargEfct
age	-0.4959	-0.0067	-0.0918	-0.1145	-0.0092

	died	survived
Pr(y x)	0.4829	0.5171

	age	height	duration	oxygt08
x=	29.8702	168.733	12.9924	.664122
sd(x)=	12.4865	13.7834	8.4412	.474109

Pr(y|x): probability of observing each y for specified x values
 Avg|Chg|: average of absolute value of the change across categories
 Min->Max: change in predicted probability as x changes from its minimum to its maximum
 0->1: change in predicted probability as x changes from 0 to 1
 ++d/2: change in predicted probability as x changes from 5 units below base value to 5 units above
 ++sd/2: change in predicted probability as x changes from 1/2 standard dev below base to 1/2 standard dev above
 MargEfct: the partial derivative of the predicted probability/rate with respect to a given independent variable

Still a 1 year step from 0 to 1
 A 10 years step at the mean

Compare:

gen age10=age/10

logistic result age10 height duration oxygt08

Logistic regression

Number of obs	=	131
LR chi2(4)	=	29.56
Prob > chi2	=	0.0000
Pseudo R2	=	0.1630

Log likelihood = -75.925772

result	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age10	.6915977	.1206307	-2.11	0.035	0.450888 1.082306
height	1.044083	.0182429	2.47	0.013	0.998893 1.089273
duration	.9373933	.0269203	-2.25	0.024	0.886088 0.9916693
oxygt08	.2945295	.1285616	-2.80	0.006	0.038893 0.550165

A 10 years step from 0 to 10
 A 10 years step at the mean

prchange age10, help

logistic: Changes in Probabilities for result

	min->max	0->1	++1/2	++sd/2	MargEfct
age10	-0.4959	-0.0729	-0.0918	-0.1145	-0.0921

	died	survived
Pr(y x)	0.4829	0.5171

	age10	height	duration	oxygt08
x=	2.98702	168.733	12.9924	.664122
sd(x)=	1.24865	13.7834	8.4412	.474109

Pr(y|x): probability of observing each y for specified x values
 Avg|Chg|: average of absolute value of the change across categories
 Min->Max: change in predicted probability as x changes from its minimum to its maximum
 0->1: change in predicted probability as x changes from 0 to 1
 ++1/2: change in predicted probability as x changes from 1/2 unit below base value to 1/2 unit above
 ++sd/2: change in predicted probability as x changes from 1/2 standard dev below base to 1/2 standard dev above
 MargEfct: the partial derivative of the predicted probability/rate with respect to a given independent variable

prvalue

logistic result age height duration oxygt08

Restore the original regression

prvalue

logistic: Predictions for result

Confidence intervals by delta method

		95% Conf. Interval	
Pr(y=survived x):	0.5171	[0.4202,	0.6140]
Pr(y=died x):	0.4829	[0.3860,	0.5798]

	age	height	duration	oxygt08
x=	29.870229	168.73282	12.992366	.66412214

prvalue, x(duration=6 height=mean oxygt08=1)

logistic: Predictions for result

Confidence intervals by delta method

		95% Conf. Interval	
Pr(y=survived x):	0.5275	[0.3861,	0.6688]
Pr(y=died x):	0.4725	[0.3312,	0.6139]

	age	height	duration	oxygt08
x=	29.870229	168.73282	6	1

prvalue, x(duration=24 oxygt08=0 age=60) save

logistic: Predictions for result

Confidence intervals by delta method

		95% Conf. Interval	
Pr(y=survived x):	0.2804	[0.0347,	0.5261]
Pr(y=died x):	0.7196	[0.4739,	0.9653]

	age	height	duration	oxygt08
x=	60	168.73282	24	0

prvalue, x(duration=24 oxygt08=0 age=30) dif

logistic: Change in Predictions for result

Confidence intervals by delta method

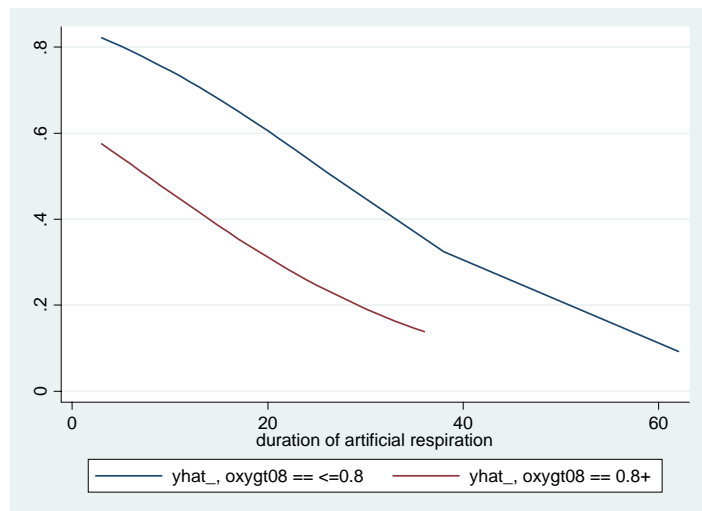
	Current	Saved	Change	95% CI for Change	
Pr(y=survived x):	0.5409	0.2804	0.2605	[0.0404,	0.4805]
Pr(y=died x):	0.4591	0.7196	-0.2605	[-0.4805,	-0.0404]

	age	height	duration	oxygt08
Current=	30	168.73282	24	0
Saved=	60	168.73282	24	0
Diff=	-30	0	0	0

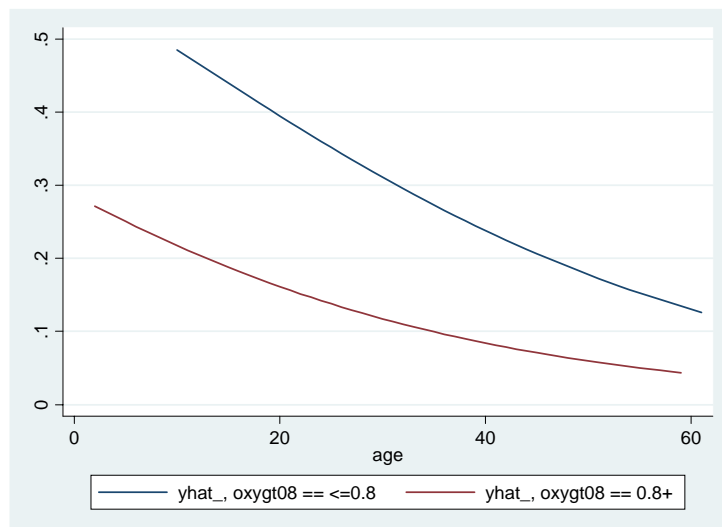
Plots

postgr3

postgr3 duration, by(oxygt08)



postgr3 age, by(oxygt08) x(height=130)



Plot with confidence

```
prgen duration, from (2) to (62) gap(2) generate (dur) rest(mean) ci
```

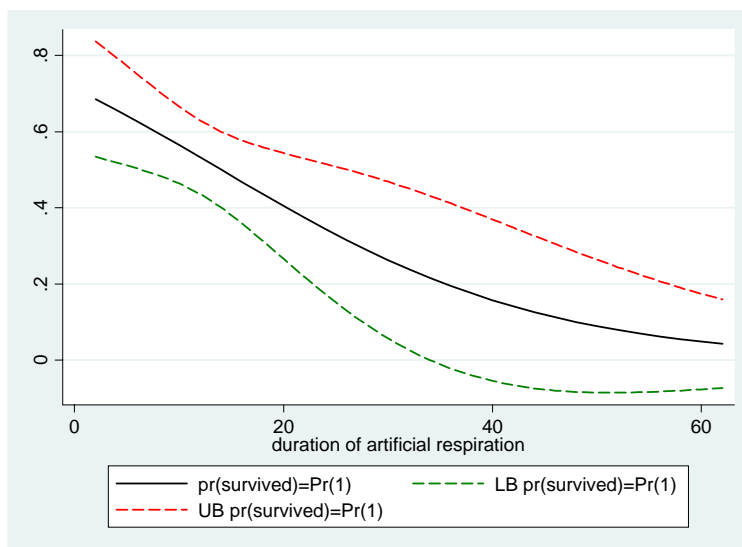
logistic: Predicted values as duration varies from 2 to 62.

```
age      height  duration  oxygt08
x= 29.870229 168.73282 12.992366 .66412214
```

New variables

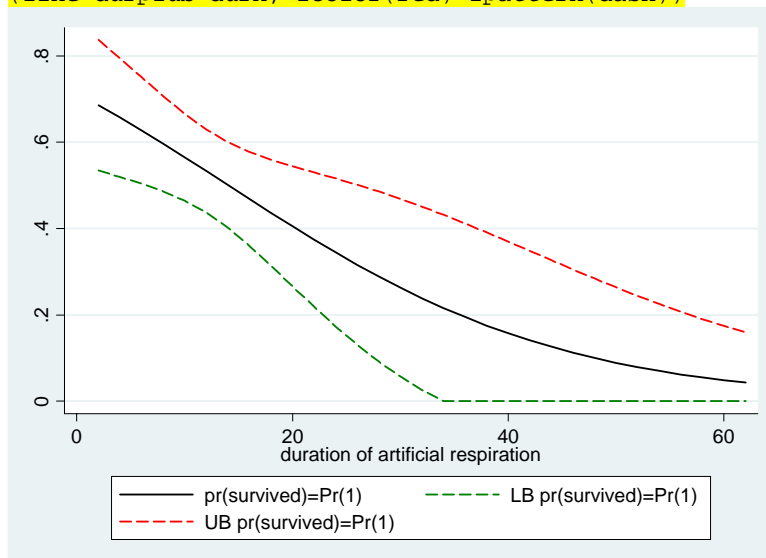
Variable	Type	Format	Description
durx	float	%9.0g	Duration of artificial respiration
durp0	float	%9.0g	pr(died)=Pr(0)
durp1	float	%9.0g	pr(survived)=Pr(1)
durp0lb	float	%9.0g	LB pr(died)=Pr(0)
durp1lb	float	%9.0g	LB pr(survived)=Pr(1)
durp0ub	float	%9.0g	UB pr(died)=Pr(0)
durp1ub	float	%9.0g	UB pr(survived)=Pr(1)

```
twoway (line durp1 durx, lcolor(black)) (line durp1lb durx, lcolor(green) lpattern(dash))
(line durp1ub durx, lcolor(red) lpattern(dash))
```



```
replace durp1lb=0 if durp1lb<0
```

```
twoway (line durp1 durx, lcolor(black)) (line durp1lb durx, lcolor(green) lpattern(dash))
(line durp1ub durx, lcolor(red) lpattern(dash))
```



Plot with confidence for groups

```
prgen duration, from (2) to (62) gap(2) generate (oxygt08) x( oxygt08=1) rest(mean) ci
```

logistic: Predicted values as duration varies from 2 to 62.

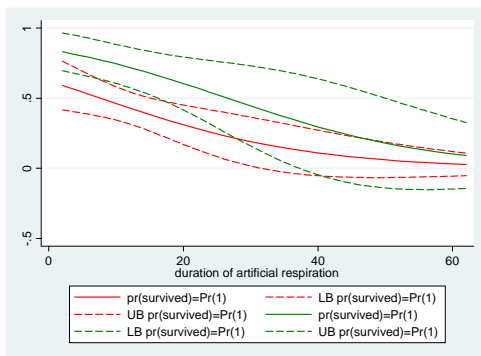
	age	height	duration	oxygt08	
x=	29.870229	168.73282	12.992366	1	
oxygt08x	float	%9.0g			Duration of artificial respiration
oxygt08p0	float	%9.0g			pr(died)=Pr(0)
oxygt08p1	float	%9.0g			pr(survived)=Pr(1)
oxygt08p0lb	float	%9.0g			LB pr(died)=Pr(0)
oxygt08p1lb	float	%9.0g			LB pr(survived)=Pr(1)
oxygt08p0ub	float	%9.0g			UB pr(died)=Pr(0)
oxygt08p1ub	float	%9.0g			UB pr(survived)=Pr(1)

```
prgen duration, from (2) to (62) gap(2) generate (oxyle08) x( oxygt08=0) rest(mean) ci
```

logistic: Predicted values as duration varies from 2 to 62.

	age	height	duration	oxygt08	
x=	29.870229	168.73282	12.992366	0	
oxyle08x	float	%9.0g			Duration of artificial respiration
oxyle08p0	float	%9.0g			pr(died)=Pr(0)
oxyle08p1	float	%9.0g			pr(survived)=Pr(1)
oxyle08p0lb	float	%9.0g			LB pr(died)=Pr(0)
oxyle08p1lb	float	%9.0g			LB pr(survived)=Pr(1)
oxyle08p0ub	float	%9.0g			UB pr(died)=Pr(0)
oxyle08p1ub	float	%9.0g			UB pr(survived)=Pr(1)

```
twoway (line oxygt08p1 oxygt08x, lcolor(red)) (line oxygt08p1lb oxygt08x, lcolor(red) lpattern(dash)) (line oxygt08p1ub oxygt08x, lcolor(red) lpattern(dash)) (line oxyle08p1 oxyle08x, lcolor(green)) (line oxyle08p1lb oxyle08x, lcolor(green) lpattern(dash)) (line oxyle08p1ub oxyle08x, lcolor(green) lpattern(dash))
```



```
twoway (rarea oxygt08p1lb oxygt08p1ub oxygt08x, color(sandb)) (line oxygt08p1 oxygt08x, lcolor(red)) (line oxygt08p1lb oxygt08x, lcolor(red) lpattern(dash)) (line oxygt08p1ub oxygt08x, lcolor(red) lpattern(dash)) (line oxyle08p1 oxyle08x, lcolor(green)) (line oxyle08p1lb oxyle08x, lcolor(green) lpattern(dash)) (line oxyle08p1ub oxyle08x, lcolor(green) lpattern(dash))
```

