

* STATA.OUTPUT -- Chapter 11

. clear

. infile v0 age ht wt sys dias chol v1 smk ab0 chd0 v2 v3 v4 using wgsnew0.data

. summarize age

Variable	Obs	Mean	Std. Dev.	Min	Max
age	3153	46.27846	5.524906	39	59

. summarize wt

Variable	Obs	Mean	Std. Dev.	Min	Max
wt	3153	169.9505	21.09835	78	320

. summarize ht

Variable	Obs	Mean	Std. Dev.	Min	Max
ht	3153	69.77704	2.528785	60	78

. summarize sys

Variable	Obs	Mean	Std. Dev.	Min	Max
sys	3153	128.6007	15.01191	98	212

. summarize dias

Variable	Obs	Mean	Std. Dev.	Min	Max
dias	3153	82.00412	9.707279	58	150

. summarize chol

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	3153	226.3584	43.34188	103	645

. summarize smk

Variable	Obs	Mean	Std. Dev.	Min	Max
smk	3153	11.59467	14.47087	0	80

. summarize ab

Variable	Obs	Mean	Std. Dev.	Min	Max
ab0	3153	.5039645	.5000636	0	1

```
. gen chd = chd0-1
. gen ab = 1
. replace ab = 0 if ab0==0
. cor chd age wt ht sys dias chol smk ab
```

(obs=3153)

	chd	age	wt	ht	sys	dias	chol
chd	1.0000						
age	0.1193	1.0000					
wt	0.0637	-0.0344	1.0000				
ht	0.0189	-0.0954	0.5329	1.0000			
sys	0.1347	0.1667	0.2541	0.0166	1.0000		
dias	0.1016	0.1393	0.2960	0.0093	0.7722	1.0000	
chol	0.1630	0.0889	0.0084	-0.0887	0.1218	0.1276	1.0000
smk	0.1044	-0.0057	-0.0818	0.0149	0.0324	-0.0586	0.0957
ab	0.1124	0.0884	0.0427	0.0327	0.0793	0.0592	0.0571

	smk	ab
smk	1.0000	
ab	0.0923	1.0000

```
. logit chd chol
```

```
Logistic regression                               Number of obs   =       3153
                                                    LR chi2(1)      =       77.07
                                                    Prob > chi2     =       0.0000
Log likelihood = -851.99957                       Pseudo R2      =       0.0433
```

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
chol	.0124863	.0014399	8.67	0.000	.0096642 .0153084
_cons	-5.371807	.3598052	-14.93	0.000	-6.077013 -4.666602

```
. predict pchol
```

```
. logistic chd chol
```

```
Logistic regression                               Number of obs   =       3153
                                                    LR chi2(1)      =       77.07
                                                    Prob > chi2     =       0.0000
Log likelihood = -851.99957                       Pseudo R2      =       0.0433
```

chd	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
chol	1.012565	.001458	8.67	0.000	1.009711 1.015426

```
. quietly centile chol, centil(10 20 30 40 50 60 70 80 90)
. sort chol
. gen chol_cat = group(10)
. sort chol_cat
. by chol_cat: summarize chd
```

-> chol_cat = 1					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	316	.028481	.1666064	0	1
-> chol_cat = 2					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.0349206	.1838709	0	1
-> chol_cat = 3					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.0571429	.2324847	0	1
-> chol_cat = 4					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.0539683	.2263146	0	1
-> chol_cat = 5					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	316	.0443038	.2060956	0	1
-> chol_cat = 6					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.0952381	.2940106	0	1
-> chol_cat = 7					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.1015873	.3025855	0	1
-> chol_cat = 8					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.0920635	.2895756	0	1
-> chol_cat = 9					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	315	.1396825	.3472089	0	1
-> chol_cat = 10					
Variable	Obs	Mean	Std. Dev.	Min	Max
chd	316	.1677215	.3742114	0	1

```
. by chol_cat: egen temp1=mean(chd)
. by chol_cat: egen temp2=mean(1-chd)
. gen lodds=log(temp1/temp2)
. duplicates examples lodds
```

Duplicates in terms of lodds

group:	#	e.g. obs	lodds
1	316	1	-3.529623
2	315	317	-3.319132
3	315	632	-2.803360
4	315	947	-2.863880
5	316	1262	-3.071370
6	315	1578	-2.251292
7	315	1893	-2.179711
8	315	2208	-2.288696
9	315	2523	-1.817929
10	316	2838	-1.601862

```
. quietly centile chol, centil(10 20 30 40 50 60 70 80 90)
. sort chol
. gen chol_cat = group(10)
. sort chol_cat
. by chol_cat: summarize chol
```

-> chol_cat = 1

chol	Obs	Mean	Std. Dev.	Min	Max
chol	316	157.0823	14.51916	103	175

-> chol_cat = 2

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	182.9937	4.613048	175	190

-> chol_cat = 3

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	197.346	3.847939	190	203

-> chol_cat = 4

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	208.4571	3.100324	203	214

-> chol_cat = 5

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	316	218.4968	2.830108	214	223

-> chol_cat = 6

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	228.4889	3.032053	223	234

-> chol_cat = 7

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	239.5587	3.544828	234	246

-> chol_cat = 8

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	252.8889	3.988483	246	260

-> chol_cat = 9

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	315	268.9968	5.639656	260	280

-> chol_cat = 10

Variable	Obs	Mean	Std. Dev.	Min	Max
chol	316	309.2563	30.8193	280	645

. *probabilities
. by chol_cat: summarize pchol

-> chol_cat = 1					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	316	.0324207	.0051972	.0165331	.0396691
-> chol_cat = 2					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.043714	.0024001	.0396691	.0474525
-> chol_cat = 3					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.0518255	.0023545	.0474525	.0553526
-> chol_cat = 4					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.0590619	.0021477	.0553526	.0629887
-> chol_cat = 5					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	316	.0664186	.0021897	.0629887	.0699562
-> chol_cat = 6					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.0745908	.0026114	.0699562	.0794376
-> chol_cat = 7					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.0847265	.0034382	.0794376	.0911082
-> chol_cat = 8					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.0985741	.0044262	.0911082	.1066554
-> chol_cat = 9					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	315	.1180383	.0073801	.1066554	.13289
-> chol_cat = 10					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchol	316	.1855996	.0662522	.13289	.9359474

```
. gen lodds = log(pchol/(1-pchol))  
. sort lodds  
. sort chol_cat  
. *log-odds  
. by chol_cat: summarize lodds
```

-> chol_cat = 1					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	316	-3.410431	.1812905	-4.085719	-3.186705
-> chol_cat = 2					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-3.086894	.0575999	-3.186705	-2.999411
-> chol_cat = 3					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-2.907686	.0480465	-2.999411	-2.837089
-> chol_cat = 4					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-2.768949	.0387116	-2.837089	-2.699739
-> chol_cat = 5					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	316	-2.643591	.0353376	-2.699739	-2.587363
-> chol_cat = 6					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-2.518827	.0378591	-2.587363	-2.450014
-> chol_cat = 7					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-2.380606	.0442618	-2.450014	-2.300178
-> chol_cat = 8					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-2.214161	.0498014	-2.300178	-2.12537
-> chol_cat = 9					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	315	-2.013033	.0704185	-2.12537	-1.875644
-> chol_cat = 10					
Variable	Obs	Mean	Std. Dev.	Min	Max
lodds	316	-1.510341	.384819	-1.875644	2.681855

. logit chd age

Logistic regression

Number of obs = 3153
LR chi2(1) = 42.88
Prob > chi2 = 0.0000
Pseudo R2 = 0.0241

Log likelihood = -869.09461

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	.0744097	.0113009	6.58	0.000	.0522603	.0965591
_cons	-5.938553	.5492503	-10.81	0.000	-7.015063	-4.862042

. logit chd wt

Logistic regression

Number of obs = 3153
LR chi2(1) = 12.32
Prob > chi2 = 0.0004
Pseudo R2 = 0.0069

Log likelihood = -884.37663

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
wt	.0104287	.0029191	3.57	0.000	.0047074	.0161501
_cons	-4.21512	.511984	-8.23	0.000	-5.21859	-3.211649

. logit chd ht

Logistic regression

Number of obs = 3153
LR chi2(1) = 1.13
Prob > chi2 = 0.2867
Pseudo R2 = 0.0006

Log likelihood = -889.96941

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ht	.0275348	.0259013	1.06	0.288	-.0232308	.0783005
_cons	-4.345331	1.812284	-2.40	0.016	-7.897342	-.7933199

. logit chd sys

Logistic regression

Number of obs = 3153
LR chi2(1) = 50.28
Prob > chi2 = 0.0000
Pseudo R2 = 0.0282

Log likelihood = -865.3978

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
sys	.0274045	.0037187	7.37	0.000	.0201159	.034693
_cons	-6.022447	.5035263	-11.96	0.000	-7.009341	-5.035554

. logit chd dias

Logistic regression

Number of obs = 3153
LR chi2(1) = 29.88
Prob > chi2 = 0.0000
Pseudo R2 = 0.0168

Log likelihood = -875.59735

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dias	.0338237	.0059958	5.64	0.000	.0220721	.0455753
_cons	-5.243111	.5128964	-10.22	0.000	-6.248369	-4.237852

. logit chd chol

Logistic regression

Number of obs = 3153
LR chi2(1) = 77.07
Prob > chi2 = 0.0000
Pseudo R2 = 0.0433

Log likelihood = -851.99957

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
chol	.0124863	.0014399	8.67	0.000	.0096642	.0153084
_cons	-5.371807	.3598052	-14.93	0.000	-6.077013	-4.666602

. logit chd smk

Logistic regression

Number of obs = 3153
LR chi2(1) = 31.67
Prob > chi2 = 0.0000
Pseudo R2 = 0.0178

Log likelihood = -874.70099

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
smk	.0235297	.0040759	5.77	0.000	.015541	.0315183
_cons	-2.746229	.0925568	-29.67	0.000	-2.927637	-2.564821

. logit chd ab

Logistic regression

Number of obs = 3153
LR chi2(1) = 40.83
Prob > chi2 = 0.0000
Pseudo R2 = 0.0229

Log likelihood = -870.1203

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ab	.8634518	.1402078	6.16	0.000	.5886495	1.138254
_cons	-2.933722	.1154623	-25.41	0.000	-3.160024	-2.70742

. logit chd age wt ht sys dias chol smk ab

```

Logistic regression                               Number of obs   =       3153
                                                    LR chi2(8)      =       198.93
                                                    Prob > chi2     =       0.0000
Log likelihood = -791.07076                       Pseudo R2       =       0.1117

```

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	.0649012	.0120767	5.37	0.000	.0412314	.0885711
wt	.007658	.0038739	1.98	0.048	.0000653	.0152507
ht	.0184704	.0329786	0.56	0.575	-.0461664	.0831072
sys	.0186586	.0064272	2.90	0.004	.0060615	.0312557
dias	-.0017547	.0107152	-0.16	0.870	-.0227562	.0192468
chol	.0111349	.0015224	7.31	0.000	.0081512	.0141187
smk	.0208314	.0042956	4.85	0.000	.0124121	.0292506
ab	.6532634	.1451742	4.50	0.000	.3687271	.9377996
_cons	-13.71567	2.314528	-5.93	0.000	-18.25206	-9.179276

. est store additive

. logistic chd age wt ht sys dias chol smk ab

```

Logistic regression                               Number of obs   =       3153
                                                    LR chi2(8)      =       198.93
                                                    Prob > chi2     =       0.0000
Log likelihood = -791.07076                       Pseudo R2       =       0.1117

```

chd	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age	1.067054	.0128864	5.37	0.000	1.042093	1.092612
wt	1.007687	.0039037	1.98	0.048	1.000065	1.015368
ht	1.018642	.0335934	0.56	0.575	.9548831	1.086658
sys	1.018834	.0065482	2.90	0.004	1.00608	1.031749
dias	.998247	.0106965	-0.16	0.870	.9775008	1.019433
chol	1.011197	.0015394	7.31	0.000	1.008184	1.014219
smk	1.021050	.004386	4.85	0.000	1.012489	1.029683
ab	1.921802	.2789961	4.50	0.000	1.445893	2.554355

. *standardized data -- commensurate beta-coefficients

. egen zage = std(age)
. egen zwt = std(wt)
. egen zht = std(ht)
. egen zsys = std(sys)
. egen zdias = std(dias)
. egen zchol = std(chol)
. egen zsmk = std(smk)
. egen zab = std(ab)

. logit chd zage zwt zht zsys zdias zchol zsmk zab

```
Logistic regression                Number of obs   =       3153
                                   LR chi2(8)         =       198.93
                                   Prob > chi2        =       0.0000
Log likelihood = -791.07076        Pseudo R2      =       0.1117
```

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
zage	.3585733	.0667224	5.37	0.000	.2277997	.4893468
zwt	.1615715	.0817326	1.98	0.048	.0013786	.3217644
zht	.0467077	.0833957	0.56	0.575	-.1167449	.2101602
zsys	.2801015	.0964845	2.90	0.004	.0909953	.4692077
zdias	-.0170332	.1040159	-0.16	0.870	-.2209006	.1868343
zchol	.4826085	.0659819	7.31	0.000	.3532863	.6119307
zsmk	.3014483	.0621613	4.85	0.000	.1796143	.4232822
zab	.3266732	.0725963	4.50	0.000	.184387	.4689594
_cons	-2.774984	.0844624	-32.85	0.000	-2.940528	-2.609441

. *spline analysis

. mkspline sage = age, cubic nknots(4)
 . mkspline swt = wt, cubic nknots(4)
 . mkspline sht = ht, cubic nknots(4)
 . mkspline ssys = sys, cubic nknots(4)
 . mkspline sdias = dias, cubic nknots(4)
 . mkspline schol = chol, cubic nknots(4)

. logit chd sage* swt* sht* ssys* sdias* schol* smk ab

```

Logistic regression                               Number of obs   =       3153
                                                  LR chi2(20)    =       216.09
                                                  Prob > chi2    =       0.0000
Log likelihood = -782.49313                    Pseudo R2      =       0.1213
    
```

chd	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
sage1	-.1329371	.0872062	-1.52	0.127	-.3038582 .0379839
sage2	1.044346	.4368531	2.39	0.017	.1881297 1.900562
sage3	-2.207283	.9231329	-2.39	0.017	-4.01659 -.3979754
swt1	.0132742	.0159522	0.83	0.405	-.0179915 .0445399
swt2	.0158485	.0510689	0.31	0.756	-.0842447 .1159417
swt3	-.0954586	.1645081	-0.58	0.562	-.4178886 .2269714
sht1	.1167998	.0961435	1.21	0.224	-.0716379 .3052376
sht2	-.3346443	.2799024	-1.20	0.232	-.883243 .2139544
sht3	1.222719	1.098748	1.11	0.266	-.9307881 3.376227
ssys1	-.0119124	.0312045	-0.38	0.703	-.0730722 .0492474
ssys2	.1804732	.1520856	1.19	0.235	-.117609 .4785555
ssys3	-.3805943	.3169247	-1.20	0.230	-1.001755 .2405668
sdias1	.0770599	.0442372	1.74	0.082	-.0096434 .1637633
sdias2	-.2747369	.1451275	-1.89	0.058	-.5591816 .0097078
sdias3	.8779016	.4814752	1.82	0.068	-.0657724 1.821576
schol1	.0169226	.0099323	1.70	0.088	-.0025444 .0363896
schol2	-.0064873	.0274069	-0.24	0.813	-.0602038 .0472292
schol3	.0052337	.0849552	0.06	0.951	-.1612755 .1717428
smk	.0205975	.0043138	4.77	0.000	.0121427 .0290523
ab	.6753959	.1458911	4.63	0.000	.3894546 .9613373
_cons	-16.60102	8.374004	-1.98	0.047	-33.01377 -.1882755

. predict pchd

. est store spline
 . lrtest additive spline

```

Likelihood-ratio test                               LR chi2(12) =       17.16
(Assumption: additive nested in spline)            Prob > chi2 =       0.1438
    
```

```
. quietly centile pchd, centil(10 20 30 40 50 60 70 80 90)
. sort pchd
. gen cat_pchd = group(10)
. sort cat_pchd
. by cat_pchd: summarize pchd
```

-> cat_pchd = 1					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	316	.010046	.0030935	.0019929	.0148132
-> cat_pchd = 2					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.0186996	.0023117	.0148238	.0230551
-> cat_pchd = 3					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.0283678	.0029754	.0230972	.0331453
-> cat_pchd = 4					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.03825	.003131	.0331628	.0437167
-> cat_pchd = 5					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	316	.0493724	.0033107	.0437315	.0551959
-> cat_pchd = 6					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.0623804	.0044679	.0552336	.0705931
-> cat_pchd = 7					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.0808903	.0057977	.0707122	.0920707
-> cat_pchd = 8					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.1081906	.0096525	.0923631	.1261055
-> cat_pchd = 9					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	315	.1519289	.0167349	.1261056	.1835481
-> cat_pchd = 10					
Variable	Obs	Mean	Std. Dev.	Min	Max
pchd	316	.2667116	.0809755	.1836586	.814405

. by cat_pchd: summarize chd

-> cat_pchd = 1						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	316	0	0	0	0	

-> cat_pchd = 2						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.015873	.1251831	0	1	

-> cat_pchd = 3						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.047619	.2132977	0	1	

-> cat_pchd = 4						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.0380952	.1917308	0	1	

-> cat_pchd = 5						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	316	.028481	.1666064	0	1	

-> cat_pchd = 6						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.047619	.2132977	0	1	

-> cat_pchd = 7						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.0984127	.2983458	0	1	

-> cat_pchd = 8						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.1460317	.3536999	0	1	

-> cat_pchd = 9						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	315	.1365079	.3438734	0	1	

-> cat_pchd = 10						
Variable	Obs	Mean	Std. Dev.	Min	Max	
chd	316	.2563291	.4372981	0	1	