

---

# 21

---

## *The Statistical Regression Model: An Easy Way to Understand the Model*

---

---

### **Appendix 21.A M65-Spread Base Means X10–X14**

---

```
libname c15c 'c:\0-chap15c';
```

```
%let spread=65;
```

```
title "Base means with M-spread&spread";
```

```
PROC RANK data=c15c.ezway_LRM groups=100 out=OUT;
```

```
var X10-X14;
```

```
ranks X10r X11r X12r X13r X14r;
```

```
run;
```

```
data spread&spread._X10;
```

```
set out;
```

```
rhp=(100-&spread)/2;
```

```
if X10r=>(rhp-1) and X10r<=(99-rhp);
```

```
keep ID X10 X10r;
```

```
run;
```

```
data spread&spread._X11;
```

```
set out;
```

```
rhp=(100-&spread)/2;
```

```
if X11r=>(rhp-1) and X11r<=(99-rhp);  
keep ID X11 X11r;  
run;
```

```
data spread&spread._X12;  
set out;  
rhp=(100-&spread)/2;  
if X12r=>(rhp-1) and X12r<=(99-rhp);  
keep ID X12 X12r;  
run;
```

```
data spread&spread._X13;  
set out;  
rhp=(100-&spread)/2;  
if X13r=>(rhp-1) and X13r<=(99-rhp);  
keep ID X13 X13r;  
run;
```

```
data spread&spread._X14;  
set out;  
rhp=(100-&spread)/2;  
if X14r=>(rhp-1) and X14r<=(99-rhp);  
keep ID X14 X14r;  
run;
```

```
PROC SORT DATA=spread&spread._X10; by ID;  
PROC SORT DATA=spread&spread._X11; by ID;  
PROC SORT DATA=spread&spread._X12; by ID;  
PROC SORT DATA=spread&spread._X13; by ID;  
PROC SORT DATA=spread&spread._X14; by ID;  
run;
```

```
data spread&spread._X10X11X12X13X14;  
merge  
spread&spread._X10 (in=var_X10)  
spread&spread._X11 (in=var_X11)  
spread&spread._X12 (in=var_X12)
```

```

spread&spread._X13 (in=var_X13)
spread&spread._X14 (in=var_X14);
by ID;
if var_X10=1 and var_X11=1 and var_X12=1 and var_X13=1 and var_X14=1;
run;

PROC MEANS data=spread&spread._X10X11X12X13X14 mean n maxdec=4;
var X10-X14;
run;

```

---

## Appendix 21.B Create Ten Datasets for Each Decile

```

libname c15c 'c:\0-chap15c';
title' X10-X14 ';

PROC LOGISTIC data=c15c.ezway_LRM nosimple des outest=coef;
model RESPONSE = X10-X14;
run;

PROC SCORE data=c15c.ezway_LRM predict type=parms score=coef out=score;
var X10-X14;
run;

data score;
set score;
estimate=response2;
run;

data notdot;
set score ;
if estimate ne .;
PROC MEANS data=notdot noprint sum; var wt;
output out=samsize (keep=samsize) sum=samsize;
run;

```

```
data scoresam (drop=samsize);
set samsize score;
retain n;
if _n_=1 then n=samsize;
if _n_=1 then delete;
run;

PROC SORT data=scoresam; by descending estimate;
run;
```

```
data score;
set scoresam;
if estimate ne . then cum_n+wt;
if estimate = . then dec=.;
else dec=floor(cum_n*10/(n+1));
prob_hat=exp(estimate)/(1+ exp(estimate));
logit=estimate;
run;
```

```
data c15c.ezway_probs;
set scoresam;
if estimate ne . then cum_n+wt;
if estimate = . then dec=.;
else dec=floor(cum_n*10/(n+1));
prob_complete=exp(estimate)/(1+ exp(estimate));
keep ID response estimate dec X10-X14 wt;
run;
```

```
data c15c.dec0 c15c.dec1 c15c.dec2 c15c.dec3 c15c.dec4
c15c.dec5 c15c.dec6 c15c.dec7 c15c.dec8 c15c.dec9;
set c15c.ezway_probs;
if dec=0 then output c15c.dec0;
if dec=1 then output c15c.dec1;
if dec=2 then output c15c.dec2;
if dec=3 then output c15c.dec3;
if dec=4 then output c15c.dec4;
if dec=5 then output c15c.dec5;
if dec=6 then output c15c.dec6;
```

```
if dec=7 then output c15c.dec7;
if dec=8 then output c15c.dec8;
if dec=9 then output c15c.dec9;
run;
```

---

## Appendix 21.C Indexed Profiles of Deciles

---

```
libname c15c 'c:\0-chap15c';
options pageno=1;

%macro doMIDSPREAD;
%do dec= 0 %to 9;
%let spread=65;

PROC RANK data=c15c.dec&dec. groups=100 out=OUT;
var X10-X14;
ranks X10r X11r X12r X13r X14r;
run;

title1 "dec=&dec ";
title2 "midspread=&spread";
run;

data midspread&spread._X10;
set out;
dec=&dec;
rhp=(100-&spread)/2;
if X10r=>(rhp-1) and X10r<=(99-rhp);
keep ID dec X10 X10r;
run;

data midspread&spread._X11;
set out;
dec=&dec;
rhp=(100-&spread)/2;
```

```
if X11r=>(rhp-1) and X11r<=(99-rhp);  
keep ID dec X11 X11r;  
run;
```

```
data midsread&spread._X12;  
set out;  
dec=&dec;  
rhp=(100-&spread)/2;  
if X12r=>(rhp-1) and X12r<=(99-rhp);  
keep ID dec X12 X12r;  
run;
```

```
data midsread&spread._X13;  
set out;  
dec=&dec;  
rhp=(100-&spread)/2;  
if X13r=>(rhp-1) and X13r<=(99-rhp);  
keep ID dec X13 X13r;  
run;
```

```
data midsread&spread._X14;  
set out;  
dec=&dec;  
rhp=(100-&spread)/2;  
if X14r=>(rhp-1) and X14r<=(99-rhp);  
keep ID X14 X14r;  
run;
```

```
PROC SORT data=midsread&spread._X10; by ID;  
PROC SORT data=midsread&spread._X11; by ID;  
PROC SORT data=midsread&spread._X12; by ID;  
PROC SORT data=midsread&spread._X13; by ID;  
PROC SORT data=midsread&spread._X14; by ID;  
run;
```

```
data midsread&spread._dec&dec._X10X11X12X13X14;  
merge  
midsread&spread._X10 (in=var_X10)  
midsread&spread._X11 (in=var_X11)
```

```
midsread&spread._X12 (in=var_X12)
midsread&spread._X13 (in=var_X13)
midsread&spread._X14 (in=var_X14);
by ID;
if var_X10=1 and var_X11=1 and var_X12=1 and var_X13=1 and var_X14=1;
run;
```

```
PROC MEANS data=midsread&spread._dec&dec._X10X11X12X13X14
mean n MAXDEC=4;
var dec X10-X14;
%end;
%mend;
%doMIDSPREAD
quit;
```

```
%let spread=65;
data midsread&spread._X10X11X12X13X14;
set midsread&spread.;
Decile=dec;
keep ID Decile X10-X14;
run;
```

```
PROC FORMAT;
value Decile
  0 = 'top'
  1 = ' 2 '
  2 = ' 3 '
  3 = ' 4 '
  4 = ' 5 '
  5 = ' 6 '
  6 = ' 7 '
  7 = ' 8 '
  8 = ' 9 '
  9 = 'bot';
run;
```

```
title '';
*21.3 Predictor Means by Deciles;
PROC TABULATE data=midsread&spread._X10X11X12X13X14;
class Decile;
var X10-X14;
table Decile, ((X10-X14) *((mean)*f=12.1));
format Decile Decile.;
run;
```

```
*21.4 Indexed Profiles of Deciles;
PROC SUMMARY data=midsread&spread._X10X11X12X13X14;
class DECILE;
var X10-X14;
output out=DECILE_means mean=;
run;
```

```
data DECILE_means;
set DECILE_means;
k=1;
run;
```

```
PROC SUMMARY data=spread&spread._X10X11X12X13X14;
var X10-X14;
output out=BASE_means mean=;
run;
```

```
data BASE_means;
set BASE_means;
drop _TYPE_ _FREQ_;
k=1;
rename
X10 = X10_mean
X11 = X11_mean
X12 = X12_mean
X13 = X13_mean
X14 = X14_mean;
```



```

PROC PRINT;
title' BASE_MEANS';
format X10_mean X11_mean X12_mean X13_mean X14_mean 12.1;
run;

PROC SORT data=BASE_MEANS; by k;
PROC SORT data=DECILE_MEANS; by k;
run;

data INDEX;
merge
BASE_MEANS DECILE_means; by k;
array DECILE_MEANS X10-X14;
array BASE_MEANS X10_mean X11_mean X12_mean X13_mean X14_mean;
array INDEX X10X X11X X12X X13X X14X;
do over INDEX;
INDEX=(DECILE_MEANS-BASE_MEANS)/BASE_MEANS;
end;

label
X10X= 'X10 indexed over X10_mean'
X11X= 'X11 indexed over X11_mean'
X12X= 'X12 indexed over X12_mean'
X13X= 'X13 indexed over X13_mean'
X14X= 'X14 indexed over X14_mean';
if DECILE=' ' then delete;
run;

PROC PRINT data=INDEX label;
var DECILE X10X X11X X12X X13X X14X;
format X10X X11X X12X X13X X14X PERCENT8.1;
format DECILE DECILE.;
title' Indexed Profiles of Deciles ';
run;

```