

ese14_5_09. sas

```
*ese14_5_09;
*soluzione esercizio regressione del 14 maggio 2009;
libname dati 'V:\carla\sas\';
options nocenter nodate nonumber;

*controllo contenuto sas data set;
proc contents data=dati.hsb2;
run;
proc means data=dati.hsb2 maxdec=0 fw=6 n min max;
run;

*definizione formato per il punteggio;
proc format;
value punti 26-36='26-36' 37-46='37-46'
            47-56='47-56' 57-66='57-66' 66-75='66-75';
value sex 1='1 F' 0='0 M';
run;

*confronto di distribuzione di scienze per sesso;
title1 'frequenza relativa per classi di punteggio e sesso';
proc freq data=dati.hsb2;
tables science*female/norow nopercnt ;
format science punti. female sex.;
run;
title1 'BOX-PLOT';
proc sort data=dati.hsb2 out=hsb2;
by female;
proc boxplot data=hsb2;
plot science*female;
run;
title1 'medie e varianze condizionate';
proc means data=dati.hsb2 mean var vardef=n maxdec=2;
var science;
class female;
format female sex.;
run;

title1 'tabella bivariata';
proc freq data=dati.hsb2;
*solo freq assolute e condizionate di scienze rispetto a math
*opzione chi sq per indice di associazione;
table science*math/norow nopercnt chi sq;
format science math punti.;
run;

title "matrici di covarianza e di correlazione";
proc corr data=dati.hsb2 cov;
var read write math science socst;
run;

title "scatter plot scienze per matematica";
symbol1 v=star color=blue;
proc gplot data=dati.hsb2;
plot science*math;
run;
```

ese14_5_09. sas

```
title "modello di regressione";
proc reg data=dati.hsb2 outest=stime;
model science=math;
run;
*valori previsti verso i residui del modello;
plot r.*p.;
run;
quit;

*calcola il punteggio stimato a scienze
per uno studente con un punteggio
a matematica pari a 30;
title 'valore previsto';
data previsione;
set stime;
voto_math=30;
ypred=interc+math*voto_math;
run;
proc print data=previsione;
run;

*modello di regressione multipla per scienze
in funzione di math, female, socst, read;
title "modello di regressione multipla";
proc reg data=dati.hsb2 outest=stime;
model science=math female socst read;
run;
*Come mai il coefficiente di math
è diverso da quello stimato nel modello precedente?;

*risposta: il coeff di math ora e' al netto
delle altre variabili;
```