TUTORIAL SCRIPT R PER L'APPLICAZIONE DEL MODELLO GAM PER IL CALCOLO DEL LEG

- 1. Installare il software statistico R (<u>http://cran.r-project.org/bin/windows/base/</u>) in Windows e l'interfaccia *user friendly* R studio (<u>https://www.rstudio.com/products/rstudio/download/</u>).
- 2. Creare una cartella dove saranno inseriti tutti i file contenenti i dati e lo script "comandi.R". Questa sarà la vostra Working directory.
- 3. Creare un foglio Excel con due variabili per ogni sostanza: la prima colonna sarà la variabile binaria con valori "1" se il campione è giudicato tossico, "0" se il campione è giudicato non tossico; la seconda colonna con i valori delle concentrazioni della sostanza analizzata.
- 4. Salvare nella Working directory il foglio in formato .csv delimitato da separatore di elenco (File- Salva con nome Salva come CSV (delimitato da separatore di elenco) (*.csv)). Aprire il file così creato (.CSV) con un editor di testo (es. blocco note), eliminando l'eventuale presenza di spazi o altri caratteri estranei ai valori numerici.
- 5. Avviare R studio con doppio click sull'icona.
- 6. Aprire lo script "comandi.R" (File-Open file). Si aprirà una finestra contenente il codice con i comandi.
- Nella seconda riga del codice, posizionandosi col puntatore, specificare l'esatto percorso della Working directory dove sono stati salvati tutti i file. Ad esempio: setwd ("C://Users//Utente//Desktop")
- 8. Nella quinta riga del codice specificare il nome del file .csv da elaborare. Ad esempio: data<-read.csv("Pb.csv",header=T,sep=";")
- Nella riga 31 del codice scegliere il livello di effetto p digitando il comando p<-livello. Ad esempio: p<-0.95 per stimare il Livello di Effetto Grave (LEG).
- 10. Posizionarsi col cursore sulla riga 1 e premere contemporaneamente Ctrl+Alt+R per lanciare l'elaborazione per intero. E' anche possibile eseguire il codice passo-passo premendo ripetutamente il pulsante "Run" localizzato nella finestra dei comandi in alto a destra.

Al termine dell'elaborazione, sulla R Console (finestra in basso), oltre i vari passaggi, verrà visualizzato il valore del livello di effetto.

Nella finestra in alto a destra (Global Environment) viene riportato un riepilogo dei dati di input e del modello GAM risultante.

Nella finestra in basso a destra (selezionando la cartella "plot", se non già selezionata) è possibile visualizzare il grafico della relazione tra concentrazione della sostanza e probabilità di effetto. Il grafico può essere esportato come immagine o in .pdf cliccando su "export".

Esempio 1 (con R Studio): Pb.csv

🗷 RStud	0		1 . Mar.	•	- 0 ×
<u>F</u> ile <u>E</u> d	it <u>C</u> ode <u>View Plots</u> <u>Session</u> <u>Build</u> <u>Debug</u> <u>Tools</u> <u>H</u> elp				
🔍 🗸 🔐 🚽 🔚 🔛 🚔 🗌 🦽 Go to file/function 🔄 🔡 🔹 Addins 👻					🙁 Project: (None) 👻
🙆 con	handi.R × 🔤 🗖	Env	ironment History		
	🗐 📄 Source on Save 🛛 🔍 🎽 🗐 💿 🕀 Run 🖙 🕞 Source 🔹 😑	1 🕣	🔒 📑 Import D	ataset 🗸 🔬	📃 List 🕶 🛛 😅
1	#cambia]a cartella di lavoro		Global Environment	-	Q,
2	Setwd(D:\\ARCHIVIO FURVIETTO\\SIN-SIR\\TOSCana\\P10mb1no\\Approtond1ment1_2015\\GAM_Pb1r	Dat	ta		
4	#carica i dati dal csv	O d	lata	277 obs. of 2 variables	
5	data<-read.csv("Pb_Pbino.csv",header=T,sep=";")	O d	lata_sup	139 obs. of 2 variables	
7		O d	lata0	144 obs. of 2 variables	
8		O d	lata0_sup	72 obs. of 2 variables	
10	# esegue 11 t test Execution Executi	O d	lata1	133 obs. of 2 variables	
11	Y<-data[,2]	O d	lata1_sup	67 obs. of 2 variables	
12	data0 < cubect (data y "0")	0 n	iewd	10000 obs. of 1 variable	
14	datal<-subset(data, y== 0) datal<-subset(data, y== "1")	val	lues		
15		L	P	238.481721846387	
16	<pre>data0_sup<-subset(data0,data0[,1]>=median(data0[,1])) data1_sup<_subset(data1_data1[_1]>=median(data1[_1]))</pre>	0 m	od	List of 49	
18	datar_sup(-subset(datar,datar[,r]))	р)	0.85	
19	data_sup<-rbind(data0_sup,data1_sup)	Op	red	List of 2	
20	t test/data sup[1] data sup[2] var equal=r)	р	red2	num [1:10000(1d)] 0.942 0	.967 1 0.996
22	t.test(uata_sup[,1]~uata_sup[,2],vai.equal==)	S	eq	num [1:10000] 421 525 127	8 875 291
23		X		num [1:277] 1.4 1.5 1.6 1	.6 1.6 1.9 2
24	# carica il package mgcv	X	_tossici	num [1:8243] 421 525 1278	875 291
26		Y		INT [1:2//] 0 1 0 1 1 1 0	000
27	# chiama l'help per il package				
28	? mgcv				
30	۰ III کې د د د د د د د د د د د د د د د د د د	File	s Plots Package	es Help Viewer	
1:10	(Top Level) 🗘 R Script 🗘		i Dom	🗷 Export - 🝳 🍕	💁 Publish 🛛 🤤
Console D:/ARCHIVIO FURVIETTO/SIN-SIR/Toscana/Piombino/Approfondimenti 2015/GAM Pbino H0 1/					
> pred<-predict(mod,type="response",se=T) ###stimare i valori di p in corrispondenza dei valor i osservati del contaminante					
> seq<-runif(10000,0,max(X)) ###Generare una variabile aleatoria uniforme nell'intervallo [0,m ax(X)]			ç: _		
> newd<-data.frame(X=seq)				/	
<pre>> pred2<-predict(mod,newd,type="response") ###Stimare i valori di p per ogni valore della vari abile uniforme</pre>			°] /	/	
<pre>> plot(seq,pred2,xlab="X (contaminante)",ylab="p (probabilità di tossicità)",cex=0.2)</pre>			0.6		
<pre>> abline(p,0,col="red")</pre>					
> X_tossici<-seq[pred2>p] ###Tutti i valori di X con probabilità di tossicità >p			8 - [/		
> LP<-min(X_tossici) ### Stima del Livello di Pericolo			0 2	200 400 600 800 100	0 1400
> LP ###Valore del Livello di Pericolo			N (and a family)		
[1] 2	38.481/			X (contaminante)	
		1			

Il codice può essere eseguito anche direttamente su R senza installare R studio (o nel caso di malfunzionamento di quest'ultimo).

Un volta aperto R, aprire lo script "comandi.R" (File-Apri script).

Eseguire i passaggi da 2 a 9 di cui sopra, poi premere Ctrl A e poi Ctrl R, facendo attenzione che il cursore sia posizionato sullo script e non sulla Console R.

Sarà visualizzato il grafico della relazione e sulla R Console il valore del livello di effetto (LE). Cliccando con il tasto destro sul grafico, potrà essere copiato, salvato o stampato.

Esempio 2 (con R): Zn.csv

