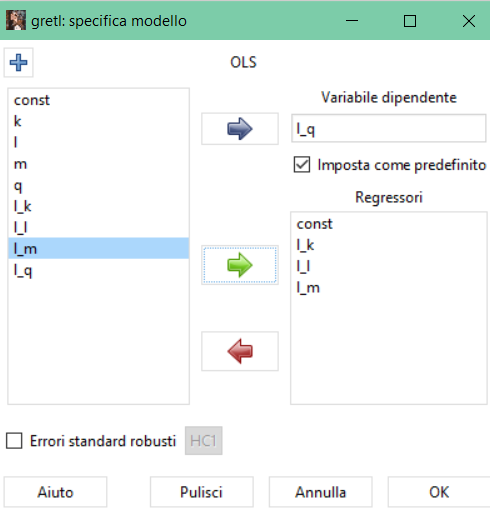
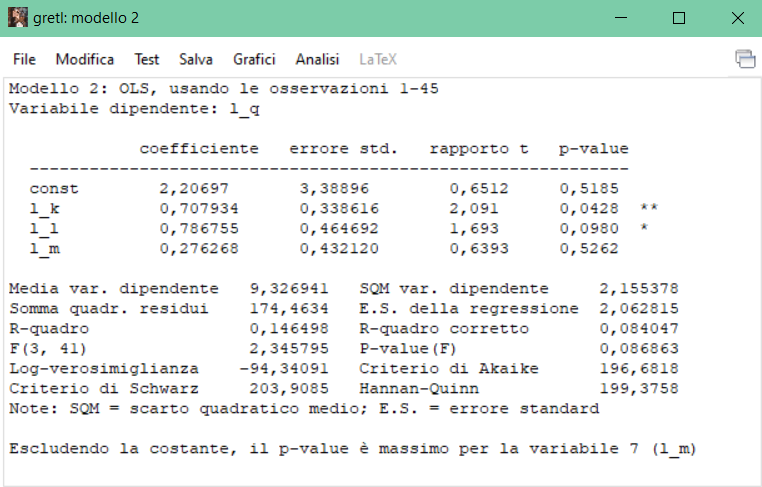
1. Il file dati **ols\_300123.gdt** contiene le informazioni riguardanti le unità di produzione (**q**) di 45 stabilimenti produttivi, le unità di capitale utilizzate (**k**) insieme alle unità di lavoro (**l**) e a quelle di beni intermedi (**m**). Sapendo che la funzione di produzione è di tipo Cobb-Douglas, cioè q=c kα lβ mδ, dove c, α, β, δ sono dei parametri da stimare, (I risultati delle stime **vanno riportati nella allegata tabella** in formato Word):
2. Si stimi il miglior modello che descrive le unità di produzione dei 45 stabilimenti
3. Si verifichi la presenza di eteroschedasticità degli errori
4. Si calcoli l’elasticità delle unità di produzione rispetto ai beni intermedi

1.a stimare il miglior modello

- siccome la Cobb Douglas è una funzione non lineare, devo calcolare il logaritmo delle variabili.

Seleziono le variabili > AGGIUNGI > LOGARITMI DELLE VARIABILI SELEZIONATE

MODELLO > MINIMI QUADRATI ORDINARI > (var dipendente: log quantità prodotta) > spunto IMPOSTA COME PREDEFINITO > (regressori: inserisco const e le altre tre variabili log)

Non è presente il DW perché i dati sono in sezione e non in serie storica.

Rq è abbastanza basso (0.14), ma nel caso dei dati in sezione sono spesso piu piccoli

La costante ha un p-value di 0.52, quindi accetto l’H0 e non sono presenti gli asterischi.

Il log del capitale ha un p value del 4.28%, quindi accetto l’H1. Questo viene confermato dai due asterischi

Nel caso del lavoro c’è un solo asterisco, quindi la probabilità di commettere un errore di tipo 1 è inferiore al 10%.

Il coefficiente dei prodotti medi ha un p-value molto alto, quindi accetto l’H0.

Ora andiamo a vedere i residui. Facciamo un grafico per analizzarli rispetto alla quantità prodotta.

Nel modello 2 > GRAFICI > RESIDUI > RISPETTO A L\_Q

C’è una concentrazione dei residui attorno all’origine. Ma c’è anche un valore molto alto che mi fa pensare ad un errore. Il modello spiega bene la parte in cui essi oscillano attorno allo 0 e poi un valore che incide in maniera pesante. In questo caso provo ad aggiungere una variabile dummy, che sarà 0 eccetto in corrispondenza di questa osservazione.

Immagine che contiene grafico

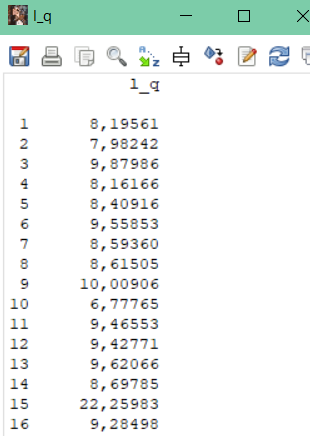
Descrizione generata automaticamente

Seleziono l-q > guardo la serie storica dei valori cercando un dato basso o elevato che può generare un errore grande.

Immagine che contiene testo

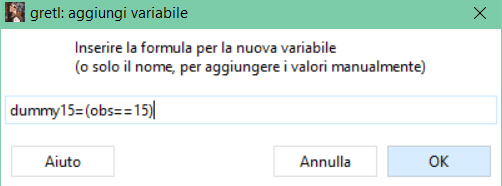
Descrizione generata automaticamente> visualizza > statistiche descrittive

Vedo che la media delle osservazioni di q è 9.32. lo scarto quadratico medio è 2,155. Posso costruire un intervallo di confidenza prendendo la media, aggiungendo e togliendo due volte lo scarto quadratico medio. Significa che l’intervallo contiene il 95% delle osservazioni, e quello che è fuori sono valori che hanno una probabilità molto bassa di essere osservati.

Nella serie cerchiamo i valori che stanno fuori da questo intervallo di confidenza

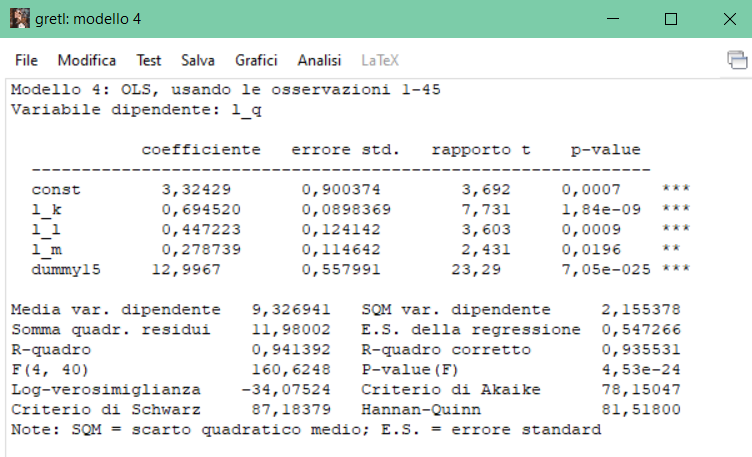
L’azienda 15 è un outlier, ecco perché dobbiamo costruire una variabile dummy.

AGGIUNGI > DEFINISCI NUOVA VARIABILE > scrivo > dummy nome dell’osservazione anormale =(obs==n osservazione)



Ora faccio un nuovo modello aggiungendo la variabile dummy tra i residui.

MODELLO > MINIMI QUADRATI ORDINARI >



Ora Rq è molto più alto. Vuol dire che la scarsa spiegazione del modello senza dummy era dovuta al fatto che molta della variabilità della nostra variabile dipendente dipendeva dall’osservazione anormale. Anche la costante è ora statisticamente significativa come le altre variabili (ci sono tutti gli asterischi). Solo il log di m presenta due asterischi, e questo mi porta ad accettare l’H1 con la probabilità di errore di tipo 1 inferiore al 5%.

CONFRONTO I DUE MODELLI IN BASE AI 3 CRITERI (SCHWARZ, AKAIKE E HANNAN-QUEEN):

Se considero i due modelli OLS calcolati, senza e con dummy, cerco di capire quale modello è migliore. Guardo il CRITERIO DI AKAIKE e vedo in quale modello la statistica è minore. In più, nel primo modello ci sono delle statistiche non significative

2.a VERIFICARE LA PRESENZA DI ETEROSCHEDACISTICITA DEGLI ERRORI

Devo verificare che gli errori siamo omoschedastici:

* Prendo il modello scelto > TEST > LMTEST- ETEROSCHEDASTICITA > TEST DI WHITE

Immagine che contiene tavolo

Descrizione generata automaticamente

Guardo la statistica test = 5.0557 e il pvalue =0.8874, accettando l’H0. Qui gli errori sono omoschedastici.

Nel test di eteroschedasticità l’H0 dice che non è presente eteroschedasticità, quindi gli errori sono omoschedastici.

Siccome non c’è eteroschedasticità, calcolare il modello spuntando ERRORI STANDARD ROBUSTI non migliora il test, perché le differenze sono minime.

3.a calcolare elasticità delle unità di produzione rispetto

Elasticità: rapporto tra due variazioni % var q/var m.

Quando le variabili sono espresse tutte in logaritmi, il coefficiente di l\_m è già l’elasticità, quindi non va calcolata. Una curva di domanda è considerata elastica quando l’elasticità è maggiore di 1. In questo caso l’elasticità è bassa (0.2787), quindi la produzione è rigida rispetto alle variazioni dei beni intermedi. Significa che l’aumento della quantità di m incide solamente del 27.87% sulla variazione di q.

1. Il file dati **ols\_140622.gdt** contiene le informazioni riguardanti l’indice di democrazia (*dem\_ind*) di un insieme di paesi per il periodo 1960 2000. Il dataset è un insieme di osservazioni in sezione e in serie storica (tecnicamente si dice che i dati sono stati “raggruppati”). Trattate il dataset come se fosse composto da osservazioni in sezione e non preoccupatevi se mancano diverse osservazioni. **Gretl** automaticamente scarterà le osservazioni con valori mancanti.
2. Si stimi il miglior modello che spiega le differenze osservate in *dem\_ind*. Si verifichi mediante l’impego di una variabile dummy se il livello di *dem\_ind* per l’Italia è superiore o inferiore alla media degli altri paesi (I risultati delle stime **vanno riportati nella allegata tabella** in formato Word)
3. Si calcoli l’elasticità di *dem\_ind* al livello di educazione dei paesi (*educ*)
4. Si calcoli l’intervallo di confidenza del coefficiente di *educ*.

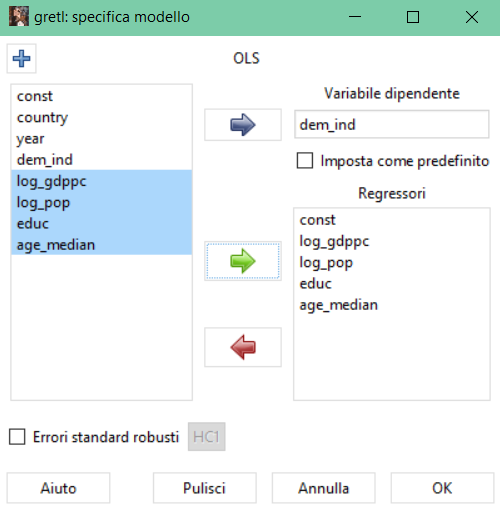
1. a stima miglior modello

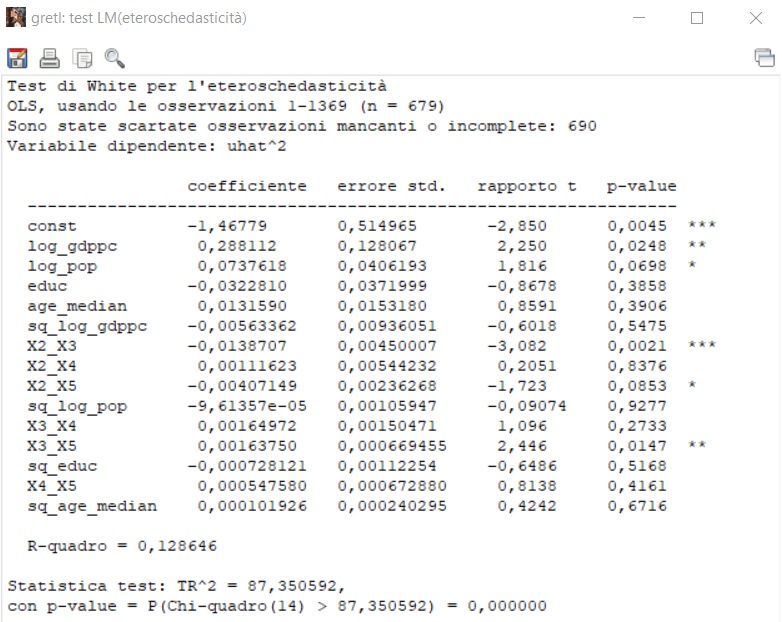
Immagine che contiene testo, tavolo

Descrizione generata automaticamenteMODELLO > MINIMI QUADRATI ORDINARI > var dip: *dem\_ind;* regr:

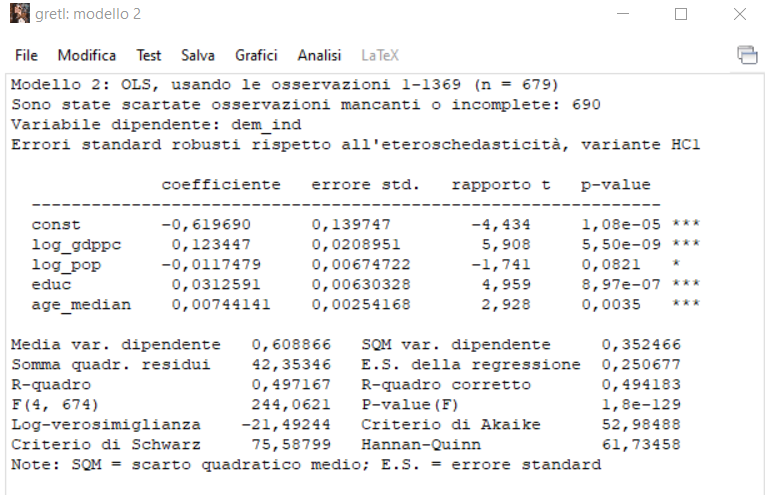
Vedendo il log\_gdppc, che è statisticamente significativo, si può affermare che nei paesi democratici il PIL è più elevato. Il log della popolazione ha un coefficiente negativo e c’è solo un asterisco, ha una probabilità inferiore al 10% di essere diverso da 0. Nei paesi più grandi, dove la popolazione è maggiore, l’indice di democrazia è più basso. L’educazione ha un coefficiente positivo, è statisticamente significativo, quindi nei paesi dove il livello di educazione è maggiore, l’indice di democrazia è più elevato. L’età mediana è anch’essa statisticamente significativa e nei paesi dove questa è alta si riscontra un indice di democrazia elevato.

Quando ci sono i dati in sezione (cross-section) è buona cosa fare il test di eteroschedasticità

Prendo il modello scelto > TEST > LMTEST- ETEROSCHEDASTICITA > TEST DI WHITE

H0: c’è omoschedasticità

Qui p-value = 0; accetto H1

Gli errori sono eteroschedastici, quindi è opportuno calcolare gli errori standard. L’eteroschedasticità non ha effetto sui coefficienti stimanti, quindi mediamente è corretto. Ciò che caambia è il valore dell’errore standard dei coefficienti e non sappiamo a priori se aumentano o diminuiscono.

Quindi questo è il caso in cui devo rifare il test spuntando la casella ERRORI STANDARD ROBUSTI

\*\*\* Probabilità di commettere un errore di tipo 1 inferiore all’1%

\* Probabilità inferiore al 10% di essere diverso da 0

Se andiamo a vedere gli errori, vediamo che è diminuito nel caso della costante, è aumentato nel caso dell’indice gdppc e più o meno è rimasto stabile nelle altre variabili. Visto che il modello è stato corretto, preferiamo il modello 2.

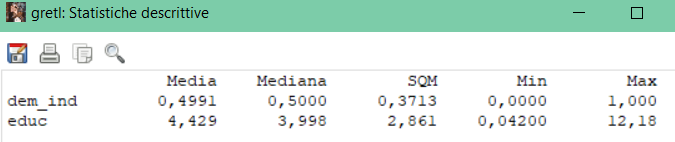
RIPORTARE NELLA TABELLA I DUE MODELLI (QUELLO STANDARD E QUELLO CON GLI ERRORI ROBUSTI)

1.b calcolare l’elasticità

Elasticità= variazione percentuale dem\_ind/variazione percentuale educ

=(var assoluta dem\_ind/var assoluta eudc)\*(media(educ)/media(dem\_ind))

VISUALIZZA > STATISTICHE DESCRITTIVE > aggiungo dem\_ind e educ



=0,031\*(4,429/0,499)= 0.275

L’indice di democrazia ha un’elasticità inferiore a 1 rispetto all’educazione.

1.c calcolare l’intervallo di confidenza

0,031 +/- 2\*0,006=0,019-----0,043

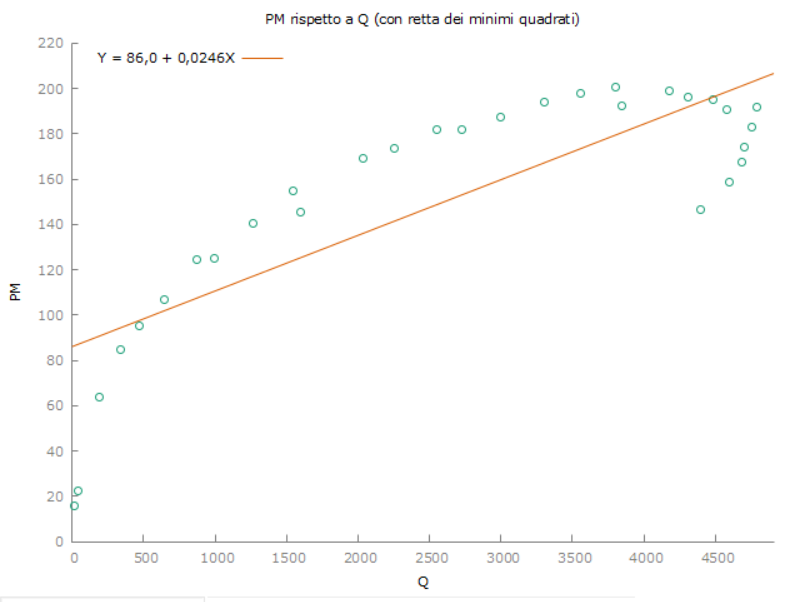
ESAME 3

Il file dati **ols\_240913.gdt** contiene le informazioni relative alla produzione di scarpe.

Rispondete alle seguenti domande:

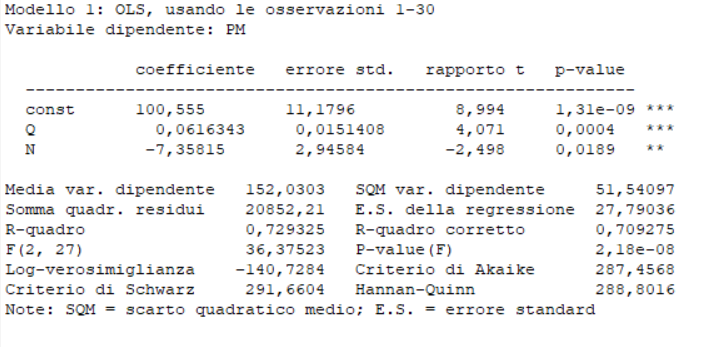
1. Si stimi la curva della produttività media.
2. Si verifichi la presenza di eteroschedasticità e si indichi, eventualmente, la soluzione per questo problema. Cosa cambia nelle stime quando si corregge per l’eventuale presenza di eteroschedasticità?
3. Si calcoli in base alla curva stimata (**cioè in base ai coefficienti stimati**) il numero di lavoratori in corrispondenza del quale la produttività media è massimizzata.

1.a stimare la curva della produttività media

Immagine che contiene grafico

Descrizione generata automaticamenteAGGIUNGI > DEFINISCI NUOVA VARIABILE > PM= Q/N

MODELLO > OLS > var dip PM; regressori Q e N

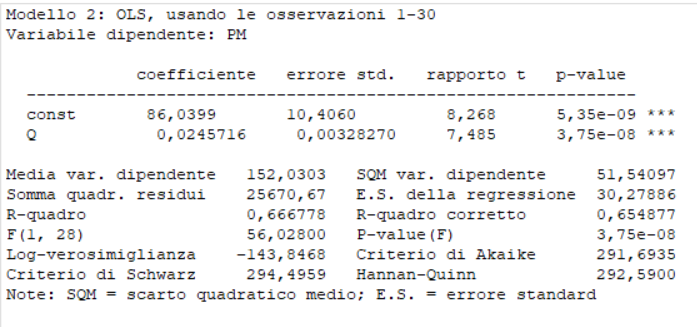
**p-value:** molto piccolo, nella distribuzione siamo molto lontani dalla media e rifiutiamo H0, la cui H è che i coefficienti siano uguali a 0. Ovviamente noi li vogliamo lontani da 0 perche dobbiamo fare delle stime con quei coefficienti.

**Statistica F:** valore molto grande, molto distanti dalla media, rifiutiamo l'H0 che entrambi i coeff siano uguali a 0

PM significativa con \*\*\*

PROVIAMO ALTRI MODELLI CON SOLO 1 VARIABILE CONSIDERATA

Immagine che contiene testo, schermata, Carattere, numero

Descrizione generata automaticamente

R^2 e R^2corretto scendono, criteri di informazioni salgono: male. Anche i criteri informativi peggiorano. Meglio il modello 1 in cui questi sono inferiori.

CONTINUIAMO CON L’ANALI DEL MODELLO

Alternativa meno evidente: modellare in modo non-lineare PM e N, oppure PM e Q. Questo si può dedurre dai grafici visti sopra. Aggiungiamo due variabili al quadrato e le usiamo:

Immagine che contiene testo, schermata, Carattere, numero

Descrizione generata automaticamenteSELEC N e Q > AGGIUNGI > QUADRATI DELLE VARIABILI SELEZIONATE

Immagine che contiene testo, schermata, Carattere, numero

Descrizione generata automaticamenteR^2 e R^2corretto molto alti, Criteri di informazioni molto bassi. Bene.

Stessa cosa con Q e Q^2, solo con risultati peggiori rispetto al Modello 6. QUINDI SCEGLIAMO IL MODELLO 6.

Immagine che contiene testo, schermata, Carattere, menu

Descrizione generata automaticamente1.b verificare presenza di eteroschedasticità

Prendo il modello scelto > TEST > LMTEST- ETEROSCHEDASTICITA > TEST DI WHITE

p-value 0.045 minore del 5%, rifiuto H0. Qui gli errori sono eteroschedastici.

Nel test di eteroschedasticità l’H0 dice che non è presente eteroschedasticità, quindi gli errori sono eteroschedastici.

Calcoliamo il modello spuntando ERRORI STANDARD ROBUSTI per migliorare il test.

Immagine che contiene testo, schermata, Carattere, numero

Descrizione generata automaticamente

Coefficienti uguali, come ci dovremmo aspettare, significatività sempre molto alta (cioè p value molto bassi).

1.c N LAVORATORI CON PRODUTTIVITA MASSIMA

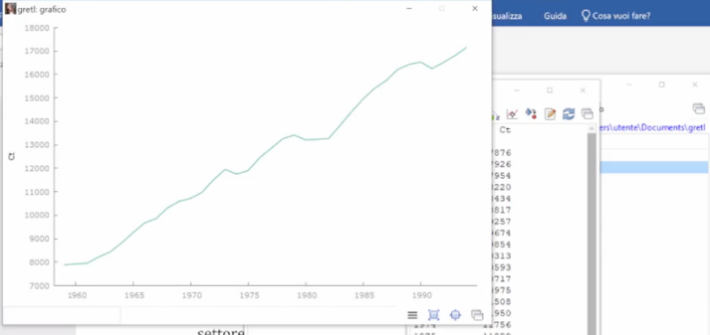
Fare ragionamento con i grafici sopra (attenzione al dato vero e quello predetto), oppure prendere la relazione stimata PM= 2,38+19,68\*N-0,49\*N^2 e imporre la parte destra =0 (sappiamo che è un massimo dal grafico), cioè

19,68-0,49\*2\*N=0🡪 N\_massimo= 20,08.

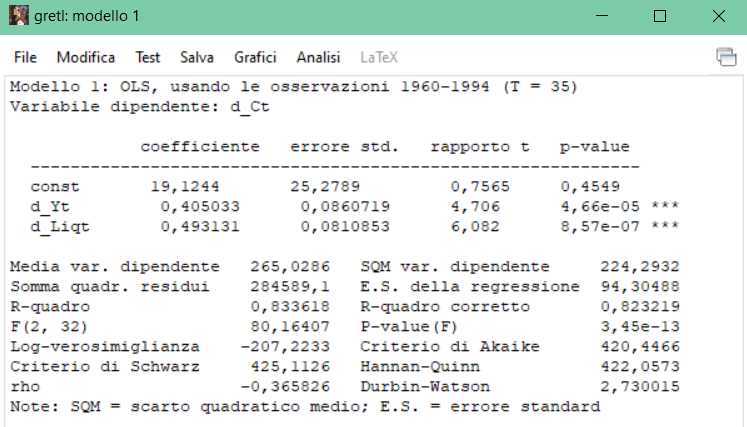
1. Il file dati **ols\_270514b.gdt** contiene le informazioni riguardanti le serie storiche dei consumi (Ct), del reddito disponibile (Yt) e di una misura di accesso al credito (Liqt). Tutte le variabili sono a prezzi costanti. Rispondete alle seguenti domande (I risultati delle stime **vanno riportati nella allegata tabella** in formato Word):
2. Si stimi il miglior modello che lega i consumi alle altre variabili.
3. Si verifichi la presenza di eteroschedasticità.
4. Si calcoli in base **ai coefficienti stimati** l’elasticità dei consumi al reddito disponibile.

1.a

In questo caso bisogna vedere com’è la crescita delle variabili cliccando almeno su una e si fa il grafico



Una modalità per eliminare il problema della crescita delle variabili (il trend delle variabili) è fare la differenza tra le variabili stesse, questo per renderle più stazionarie e ottenere una varianza costante e una media costante.



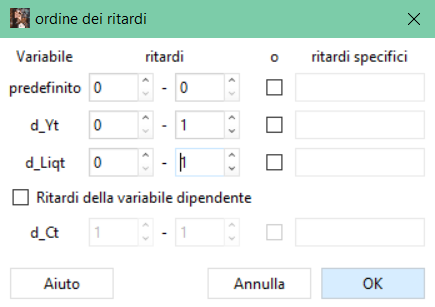
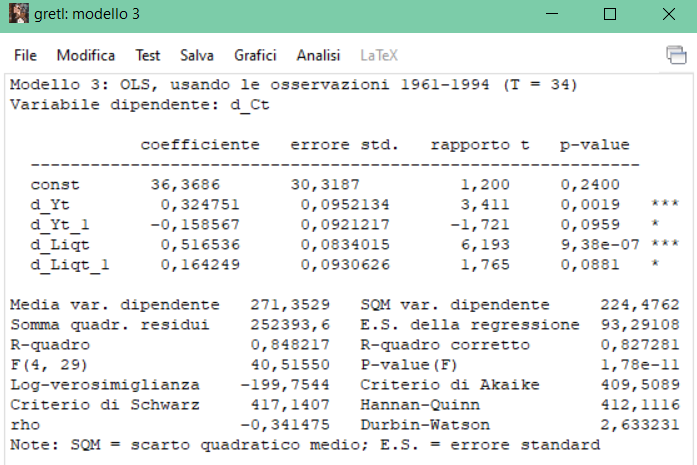
- AGGIUNGI > differenze delle variabili SEL.

- MODELLO > OLS con le differenze

Variabili significative, p value piccolo, accetto H0.

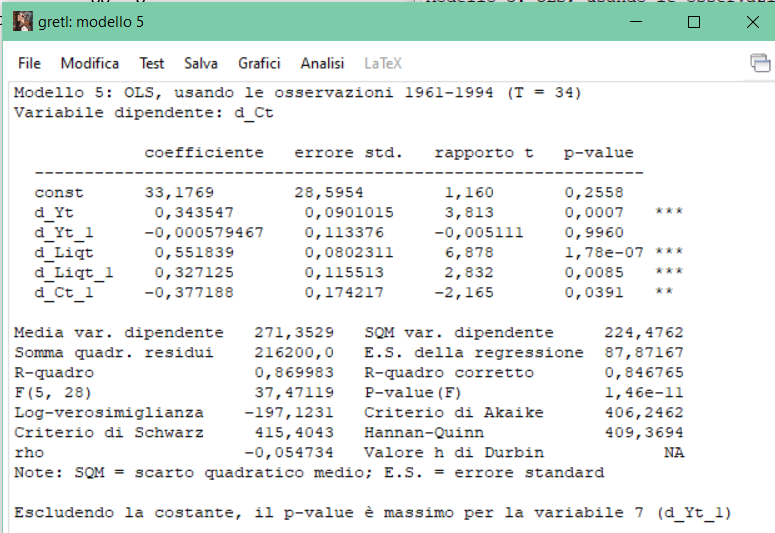
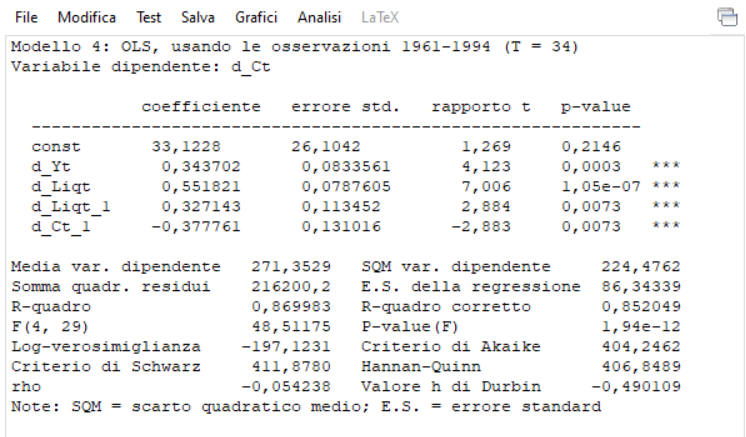
Cerco un altro modello che sia migliore.

MODIFICA > MODIFICA MODELLO > AGGIUNGIAMO 1 RITARDO



>>>>>>>>>>>>>>>>>>

Le variabili aggiunte sono meno statisticamente significative in quanto hanno un solo asterisco, sono al 10% statisticamente significative.

Andiamo avanti aggiungendo ritardi anche alla variabile dipendente.

A questo punto bisogna analizzare i modelli fin qui ottenuti:

* R-quadro
* Schwarz
* Durbin Watson
* Hannan Quinn

Dati i criteri di informazione analizzati, il modello migliore come abbiamo visto è il modello 4. Il modello 4 è quello con R^2 corretto e Akaike, Schwarz e Hannan Quinn minori, con rho piccolo e h in valore assoluto molto sotto a 2.

1.b **TEST DI ETEROSCHEDASTICITA’**

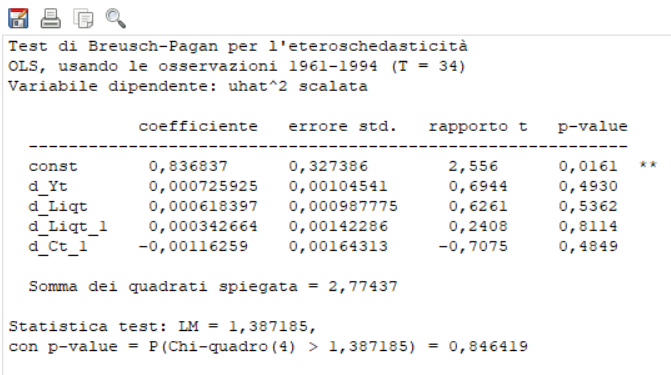
dal modello scelto facciamo TEST🡪 LMTEST – AUTOCORRELAZIONE > > TEST DI BREUSCH- GODFREY

ORDINE DI RITARDI 2

Da LMF, non rifiuto l’ipotesi nulla di assenza di autocorrelazioneImmagine che contiene testo, schermata, Carattere, menu

Descrizione generata automaticamente

dal modello scelto facciamo TEST🡪 LMTEST eteroschedasticità e applichiamo uno dei test > TEST DI BREUSCH- PAGAN

Andiamo a vedere il p-value che è 0,84 ciò vuol dire che all’84%

Quindi, non rifiuto l’ipotesi nulla di assenza di eteroschedasticità.

1.c ELASTICITA DEI CONSUMI AL REDDITO DISPONIBILE

VISUALIZZA > STATISTICHE DESCRITTIVE

Immagine che contiene testo, schermata, Carattere, ricevuta

Descrizione generata automaticamente

Per elasticità: nonostante i coefficienti stimati siano riferiti a (delta\_var), possiamo usarli anche per le variabili in livelli. Quindi:

Elasticità\_(Ct al variare di Yt)= (Coeff. Stimato del nostro modello migliore)\*(media\_y/media\_c)

(Ricorda, Elasticità di due variabili= variazione percentuale di una rispetto alla variazione percentuale dell’altra).