

## Επαναληπτικές Ασκήσεις-Λύσεις

### Άσκηση 1:

1. Αν το ποσοστό μεγέθυνσης του ΑΕΠ είναι μηδέν (μηδενική ανάπτυξη) και ο πληθωρισμός είναι και αυτός μηδέν (σταθερές τιμές), τότε το κυβερνών κόμμα στις εκλογές θα λάβει περίπου 52%, και συνεπώς θα εκλέξει τον νέο πρόεδρο. Η παράμετρος αυτή είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%. Αν το ποσοστό μεγέθυνσης του ΑΕΠ αυξηθεί κατά 1 ποσοστιαία μονάδα (π.μ.) το προηγούμενο τρίμηνο τότε το κυβερνών κόμμα θα λάβει 0.64% περισσότερο στις προεδρικές εκλογές. Η παράμετρος είναι στατιστικά σημαντική στο 1%, οπότε υπάρχει μια στατιστικά θετική σχέση μεταξύ οικονομικής ανάπτυξης και εκλογικού αποτελέσματος για το κυβερνών κόμμα. Αν ο μέσος πληθωρισμός αυξηθεί κατά 1 π.μ. τότε το κυβερνών θα λάβει 0.17% λιγότερο στις εκλογές. Η παράμετρος όμως δεν είναι στατιστικά σημαντική ούτε στο 10% επίπεδο σημαντικότητας, υποδεικνύοντας ότι ο πληθωρισμός δεν καθορίζει τελικά το εκλογικό αποτέλεσμα. Ο συντελεστής προσδιορισμού είναι 35%, οπότε το 35% της διακύμανσης του ποσοστού των ψήφων του κυβερνώντος κόμματος στις εκλογές ερμηνεύεται από το υπόδειγμα ενώ το υπόλοιπο 65% παραμένει ανεξηγήτο.

2.  $\widehat{VOTE} = 52.15 + 0.64 \times (-4) - 0.17 \times 4 = 48.91\%$ . Άρα ο κυβερνητικός υποψήφιος θα λάβει 48.91%, δηλαδή αναμένεται ότι θα χάσει τις εκλογές.

3. Θέλουμε να πραγματοποιήσουμε τον εξής έλεγχο:  $H_0: \beta_2 = 0.5$ ,  $H_1: \beta_2 \neq 0.5$ . Η στατιστική του ελέγχου στο δείγμα είναι

$$t = \frac{0.64 - 0.5}{0.16} = 0.875$$

Η κριτικές τιμές του ελέγχου με 1% επίπεδο σημαντικότητας είναι  $t_{(0.995,30)} = 2.750$  και  $t_{(0.005,30)} = -2.750$ . Παρατηρούμε ότι  $-2.750 < 0.875 < 2.750$  οπότε δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση στο 1% επίπεδο σημαντικότητας. Αυτό σημαίνει ότι τα δεδομένα του δείγματος συνάδουν με μια τιμή για την παράμετρο  $\beta_2$  ίση με 0.5.

4. Παρατηρούμε ότι τα p-values του ελέγχου (είτε χρησιμοποιήσουμε την  $F$  στατιστική είτε την  $\chi^2$  στατιστική) είναι μεγαλύτερα από 10%, οπότε καταλήγουμε στο συμπέρασμα ότι δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση της ομοσκεδαστικότητας.
5. Αν ο πρόεδρος δεν είναι εκ νέου υποψήφιος ( $PERSON = 0$ ) τότε το κυβερνών κόμμα αναμένεται να λάβει στις εκλογές 51% όταν το ποσοστό μεγέθυνσης του ΑΕΠ είναι μηδέν (μηδενική ανάπτυξη) και ο πληθωρισμός είναι και αυτός μηδέν (σταθερές τιμές). Συνεπώς θα εκλέξει τον νέο πρόεδρο. Σημειώστε ότι η παράμετρος αυτή είναι στατιστικά σημαντική στο 1%. Αν ο πρόεδρος είναι εκ νέου υποψήφιος το κυβερνών κόμμα θα λάβει 2.54% περισσότερο, αλλά όπως παρατηρούμε από τα αποτελέσματα του πίνακα η παράμετρος αυτή δεν είναι στατιστικά σημαντική. Συνεπώς η εκ νέου υποψηφιότητα του απερχόμενου προέδρου από μόνης της δεν επηρεάζει σημαντικά τα εκλογικά ποσοστά του κυβερνώντος κόμματος. Αν το ποσοστό μεγέθυνσης του ΑΕΠ αυξηθεί κατά 1 π.μ. το προηγούμενο τρίμηνο τότε το κυβερνών κόμμα θα λάβει 0.59%

περισσότερο στις προεδρικές εκλογές. Η παράμετρος είναι στατιστικά σημαντική στο 1%, οπότε υπάρχει μια στατιστικά θετική σχέση μεταξύ οικονομικής ανάπτυξης και εκλογικού αποτελέσματος για το κυβερνών κόμμα. Αν ο μέσος πληθωρισμός αυξηθεί κατά 1 π.μ. τότε το κυβερνών θα λάβει 0.28% λιγότερο στις εκλογές. Η παράμετρος όμως δεν είναι στατιστικά σημαντική ούτε στο 10% επίπεδο σημαντικότητας, υποδεικνύοντας ότι ο πληθωρισμός δεν καθορίζει τελικά το εκλογικό αποτέλεσμα. Ο συντελεστής προσδιορισμού είναι 39%, οπότε το 39% της διακύμανσης του ποσοστού των ψήφων του κυβερνώντος κόμματος στις εκλογές ερμηνεύεται από το υπόδειγμα ενώ το υπόλοιπο 61% παραμένει ανεξηγήτο.

Το δεύτερο υπόδειγμα έχει μεγαλύτερο  $\text{adjusted-}R^2$  και μικρότερο AIC οπότε φαίνεται ότι το δεύτερο υπόδειγμα έχει καλύτερη προσαρμογή στα δεδομένα. Από την άλλη πρέπει να σημειώσουμε ότι το BIC είναι μεγαλύτερο στο δεύτερο υπόδειγμα. Αυτό οφείλεται πρώτον στο ότι εισάγουμε μια νέα μεταβλητή η οποία δεν είναι στατιστικά σημαντική, και δεύτερον ότι το κριτήριο BIC είναι πιο αυστηρό (σε σχέση με τα υπόλοιπα) στην απώλεια βαθμών ελευθερίας.

## Άσκηση 2:

1. Σε ένα τέτοιο υπόδειγμα είναι δύσκολο να ερμηνεύσουμε τον σταθερό όρο (πώς θα μπορούσαμε να υποστηρίξουμε ότι η ηλικία του μεγαλύτερου μέλους της οικογένειας είναι 0!). Αν η συνολική κατανάλωση του νοικοκυριού αυξηθεί (μειωθεί) κατά 1% τότε το ποσοστό κατανάλωσης του νοικοκυριού σε αλκοολούχα ποτά θα αυξηθεί (μειωθεί) κατά  $0.032/100 = 0.00032$  π.μ. Η παράμετρος, αν και μικρή σε μέγεθος, είναι στατιστικά σημαντική στο 1%, οπότε υπάρχει μια στατιστικά θετική σχέση μεταξύ της συνολικής κατανάλωσης του νοικοκυριού και του ποσοστού κατανάλωσης σε αλκοολούχα ποτά. Αν η ηλικία του μεγαλύτερου μέλους της οικογένειας αυξηθεί (μειωθεί) κατά 1 έτος τότε το ποσοστό κατανάλωσης του νοικοκυριού σε αλκοολούχα ποτά θα μειωθεί (αυξηθεί) κατά 0.0022 π.μ. Η παράμετρος, αν και μικρή σε μέγεθος, είναι στατιστικά σημαντική στο 1%. Έτσι, όσο μεγαλύτερη είναι η ηλικία του μεγαλύτερου μέλους της οικογένειας (που αποτελεί έναν μάλλον σημαντικό καταναλωτή αλκοολούχων ποτών, σε σχέση με τα μικρότερα μέλη) τόσο το ποσοστό κατανάλωσης μειώνεται. Τέλος, αν αυξηθεί ο αριθμός ανηλίκων στην οικογένεια κατά 1 παιδί τότε το ποσοστό κατανάλωσης του νοικοκυριού σε αλκοολούχα ποτά αναμένεται να μειωθεί (αυξηθεί) κατά 0.014 π.μ. Η παράμετρος είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 5%. Έτσι όσα περισσότερα ανήλικα μέλη έχει μια οικογένεια, τόσο μειώνεται το ποσοστό κατανάλωσης αλκοολούχων ποτών. Αυτό πιθανόν να οφείλεται στο ότι η οικογένεια κατευθύνει τις καταναλωτικές δαπάνες σε προϊόντα και υπηρεσίες που καλύπτουν τις ανάγκες και προτιμήσεις των ανήλικων μελών, και σίγουρα αυτά δεν περιέχουν τα αλκοολούχα ποτά! Ο συντελεστής προσδιορισμού είναι 8%, οπότε το 8% της διακύμανσης του ποσοστού κατανάλωσης του νοικοκυριού σε αλκοολούχα ποτά ερμηνεύεται από το υπόδειγμα, ενώ το υπόλοιπο 92% παραμένει ανεξηγήτο.

2. Το 95% διάστημα εμπιστοσύνης δίνεται ως  $(-0.0022 - 1.96 \times 0.0004, -0.0022 + 1.96 \times 0.0004) = (-0.003, -0.0014)$  λαμβάνοντας υπόψη ότι  $t_{(0.025, 496)} = -1.96$  και  $t_{(0.975, 496)} = 1.96$ . Άρα με επίπεδο εμπιστοσύνης 95% η πληθυσμιακή παράμετρος  $\beta_2$  λαμβάνει τιμές μεταξύ του -0.003 και του -0.0014, άρα μια αύξηση της ηλικίας του μεγαλύτερου μέλους της οικογένειας κατά 1 έτος αναμένεται να μειώσει το ποσοστό κατανάλωσης σε αλκοολούχα ποτά μεταξύ -0.003 και -0.0014 π.μ. Το διάστημα αυτό είναι σχετικά μικρό, οπότε η συγκεκριμένη παράμετρος είχε εκτιμηθεί με αξιοπιστία.
3. Μπορούμε να γράψουμε την εξαρτημένη μεταβλητή ως:

$$WALC = \frac{ALCEXP}{TOTEXP}$$

όπου  $ALCEXP$  η κατανάλωση του νοικοκυριού σε αλκοολούχα ποτά και  $TOTEXP$  η συνολική κατανάλωση. Η παράμετρος  $\beta_2$  ορίζεται ως:

$$\begin{aligned} \beta_2 &= \frac{\partial WALC}{\partial \log(TOTEXP)} = \frac{\partial \left( \frac{ALCEXP}{TOTEXP} \right)}{\frac{\partial TOTEXP}{TOTEXP}} \\ &= TOTEXP \left( \frac{\frac{\partial ALCEXP}{\partial TOTEXP} TOTEXP - ALCEXP}{TOTEXP^2} \right) \\ &= \frac{\partial ALCEXP}{\partial TOTEXP} - \frac{ALCEXP}{TOTEXP} \end{aligned}$$

Αν τώρα υποθέσουμε ότι το  $\beta_2 < 0$  τότε:

$$\begin{aligned} \frac{\partial ALCEXP}{\partial TOTEXP} - \frac{ALCEXP}{TOTEXP} &< 0 \Leftrightarrow \frac{\partial ALCEXP}{\partial TOTEXP} < \frac{ALCEXP}{TOTEXP} \\ &\Leftrightarrow \frac{\partial ALCEXP / ALCEXP}{\partial TOTEXP / TOTEXP} < 1 \Leftrightarrow \varepsilon < 1 \end{aligned}$$

Άρα όταν το  $\beta_2 < 0$  τότε η ελαστικότητα της κατανάλωσης του νοικοκυριού σε αλκοολούχα ποτά (σε σχέση με την συνολική κατανάλωση) είναι μικρότερη από 1. Άρα όταν το νοικοκυριό μειώνει (αυξάνει) την συνολική του κατανάλωση κατά 1%, η κατανάλωση σε αλκοολούχα ποτά μειώνεται (αυξάνεται) κατά ένα μικρότερο ποσοστό. Αυτό σημαίνει ότι τα αλκοολούχα ποτά είναι είδη πρώτης ανάγκης. Αντίστροφα όταν το  $\beta_2 > 0$  τότε η ελαστικότητα είναι μεγαλύτερη του 1, και τα αλκοολούχα ποτά θεωρούνται είδη πολυτελείας.

Μπορούμε βέβαια να εξηγήσουμε τα παραπάνω με έναν λιγότερο αυστηρά μαθηματικό τρόπο βασιζόμενοι όμως και πάλι τον ορισμό της παραμέτρου  $\beta_2$ . Ας υποθέσουμε ότι το νοικοκυριό υφίσταται μια μείωση της συνολικής κατανάλωσης. Αν το  $\beta_2 > 0$  τότε το ποσοστό της κατανάλωσης σε αλκοολούχα ποτά θα μειωθεί. Για να μειωθεί όμως το ποσοστό θα πρέπει η κατανάλωση σε αλκοολούχα ποτά να μειωθεί περισσότερο σε σχέση με την συνολική κατανάλωση (που είναι ήδη μειωμένη). Αυτό σημαίνει ότι τα αλκοολούχα ποτά είναι είδη πολυτελείας. Αν τώρα το  $\beta_2 < 0$  τότε το ποσοστό της κατανάλωσης σε αλκοολούχα ποτά θα έπρεπε να αυξηθεί. Έτσι ενώ μειώνεται η συνολική κατανάλωση, τα αλκοολούχα ποτά καταλαμβάνουν ένα μεγαλύτερο μέρος (ποσοστό) αυτής. Αυτό σημαίνει ότι η κατανάλωση τους δεν μειώνεται με

αντίστοιχο ρυθμό, οπότε μπορούμε να θεωρήσουμε ότι είναι είδος πρώτης ανάγκης.

4. Η στατιστική του ελέγχου είναι  $t = \frac{0.032}{0.0081} = 3.95$ . Η κριτική τιμή σε επίπεδο σημαντικότητας 5% είναι  $t_{(0.95,496)} = 1.645$ . Παρατηρούμε ότι  $3.95 > 1.645$  άρα η μηδενική υπόθεση απορρίπτεται σε 5% επίπεδο σημαντικότητας. Δεχόμαστε την εναλλακτική υπόθεση  $\beta_2 > 0$  και συμπεραίνουμε ότι τα αλκοολούχα ποτά είναι είδη πολυτελείας.

### Άσκηση 3:

1. Αν η ποσοστιαία μεταβολή του κατά κεφαλήν εισοδήματος είναι 0 (δηλαδή το εισόδημα δεν μεταβλήθηκε καθόλου το προηγούμενο τρίμηνο) τότε αναμένεται μια αύξηση της κατανάλωσης κατά 0.95 π.μ. Η παράμετρος είναι στατιστικά σημαντική στο 1%, υποδεικνύοντας ότι ακόμα και σε περιόδους (τρίμηνα) που το εισόδημα δεν άλλαξε σημαντικά η κατανάλωση συνέχισε να αυξάνεται, πιθανόν διότι οι καταναλωτές μπορούν να αυξήσουν την κατανάλωση χρησιμοποιώντας άλλες πηγές (πχ. αποταμιεύσεις). Μια αύξηση (μείωση) της ποσοστιαίας μεταβολής του εισοδήματος κατά 1 π.μ. θα αυξήσει (μειώσει) το ποσοστό μεταβολής της κατανάλωσης κατά 0.45 π.μ. Η παράμετρος είναι στατιστικά σημαντική στο 1% επίπεδο σημαντικότητας υποδεικνύοντας μια στατιστικά σημαντική σχέση μεταξύ μεγέθυνσης (ή επιβράδυνσης) του εισοδήματος και της κατανάλωσης. Ο συντελεστής προσδιορισμού είναι 30%, οπότε το 30% της διακύμανσης της ποσοστιαίας μεταβολής της κατά κεφαλήν κατανάλωσης ερμηνεύεται από το υπόδειγμα, ενώ το υπόλοιπο 70% παραμένει ανεξηγήτο.
2. Θέλουμε να ελέγξουμε την υπόθεση  $H_0: \beta_2 \leq 0.5$ ,  $H_1: \beta_2 > 0.5$ . Η στατιστική του ελέγχου ισούται με:

$$t = \frac{0.45 - 0.5}{0.05} = -1$$

Η κριτική τιμή του ελέγχου ισούται με  $t_{(0.95,298)} = 1.645$ . Παρατηρούμε ότι  $-1 < 1.645$  οπότε δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση σε 5% επίπεδο σημαντικότητας. Οπότε καταλήγουμε στο συμπέρασμα ότι με βάση το δείγμα δεν φαίνεται η αύξηση του κατά κεφαλήν εισοδήματος κατά 2 π.μ. να οδηγεί στην αύξηση της κατά κεφαλήν κατανάλωσης περισσότερο από 1 π.μ.

3. Ο έλεγχος Lagrange multiplier είναι ένας έλεγχος αυτοσυσχέτισης. Στον πίνακα 2 παρατηρούμε ότι τα p-values της  $F$  στατιστικής και της  $\chi^2$  στατιστικής είναι μικρότερα από 1% οπότε απορρίπτουμε την μηδενική υπόθεση μη ύπαρξης αυτοσυσχέτισης. Καταλήγουμε στο συμπέρασμα ότι ο διαταρακτικός όρος παρουσιάζει αυτοσυσχέτιση. Με βάση αυτό, οι εκτιμητές ελάχιστων τετραγώνων είναι αμερόληπτοι και συνεπείς αλλά δεν είναι πλέον αποτελεσματικοί. Επίσης αν η αυτοσυσχέτιση προκαλείται από ανεξάρτητες μεταβλητές που έχουμε παραλείψει και οι οποίες μπορεί να σχετίζονται με το ποσοστό μεγέθυνσης του κατά κεφαλήν εισοδήματος τότε το υπόδειγμα παράγει εκτιμητές οι οποίοι είναι μεροληπτικοί και ασυνεπείς λόγω της ενδογένειας.

4. Στο δεύτερο υπόδειγμα μια αύξηση (μείωση) της ποσοστιαίας μεταβολής του εισοδήματος κατά 1 π.μ. θα αυξήσει (μειώσει) το ποσοστό μεταβολής της κατανάλωσης κατά 0.35 π.μ. Η παράμετρος είναι στατιστικά σημαντική στο 1% επίπεδο σημαντικότητας υποδεικνύοντας μια στατιστικά σημαντική σχέση μεταξύ μεγέθυνσης (ή επιβράδυνσης) του εισοδήματος και της κατανάλωσης. Όμως σε σχέση με το πρώτο υπόδειγμα παρατηρούμε μια μείωση στην εκτίμηση αυτής της παραμέτρου της τάξεως των 0.10 π.μ. Η διαφορά αυτή μπορεί να σημαίνει ότι η εκτίμηση του αρχικού υποδείγματος είναι προς τα πάνω μεροληπτική (upward biased). Αυτή η μεροληψία οφείλεται στην απουσία από το πρώτο υπόδειγμα σημαντικών μεταβλητών (omitted variable bias) η οποία προκαλεί ενδογένεια. Έτσι η προς τα πάνω μεροληψία μπορεί να δημιουργείται από την απουσία της μεταβλητής  $INCGWTH_{t-1}$ . Η μεταβλητή αυτή έχει μια θετική σχέση με την εξαρτημένη μεταβλητή (όπως συνάδει από την εκτίμηση της παραμέτρου  $\beta_5$ ) και λογικά μια θετική συσχέτιση με την  $INCGWTH_t$ . Αυτά τα δύο από κοινού εξηγούν την προς τα πάνω μεροληψία της παραμέτρου  $\beta_2$  που παρατηρούμε στο πρώτο υπόδειγμα.
5. Θα επιλέγαμε το δεύτερο υπόδειγμα για την ερμηνεία της σχέσης μεταξύ της μεταβολής της κατανάλωσης και του εισοδήματος. Πρώτον, το δεύτερο υπόδειγμα έχει καλύτερη προσαρμοστικότητα όπως συνάδεται από το υψηλότερο  $adjusted-R^2$  και τα χαμηλότερα AIC και BIC. Δεύτερον, τα αποτελέσματα του Πίνακα 4 υποδεικνύουν ότι στο δεύτερο υπόδειγμα η μηδενική υπόθεση τη μη-αυτοσυσχέτισης δεν μπορεί να απορριφθεί σε όλα τα συμβατικά επίπεδα σημαντικότητας. Έτσι, φαίνεται ότι η συμπερίληψη των υστερήσεων της εξαρτημένης και ανεξάρτητης μεταβλητής στο υπόδειγμα αφαιρεί την αυτοσυσχέτιση που παρατηρήθηκε στον διαταρακτικό όρο του πρώτου υποδείγματος, ή με άλλα λόγια αυτή η αυτοσυσχέτιση αντανάκλασσε σε μεγάλο βαθμό την αυτοσυσχέτιση της εξαρτημένης μεταβλητής και την δυναμική σχέση που έχει με την ανεξάρτητη μεταβλητή. Τρίτον, όπως αναφέρθηκε και παραπάνω, στο πρώτο υπόδειγμα οι εκτιμήσεις μπορεί να είναι μεροληπτικές.
6.  $CONG\widehat{WTH}_t = 0.31 + 0.013 \times 1 + 0.21 \times 1.3 + 0.35 \times 0.6 + 0.23 \times 0.9 = 1.013$ . Άρα προβλέπεται ότι το πρώτο τρίμηνο του 2010 η μεταβολή της κατά κεφαλήν κατανάλωσης θα είναι 1.013 π.μ.

#### Άσκηση 4:

1. Ο σταθερός όρος σε ένα τέτοιο υπόδειγμα δεν έχει κάποια οικονομική ερμηνεία, παρά μόνο αν τον σκεφτούμε σαν ένα προνοιακό επίδομα που λαμβάνει κάποιος με 0 έτη στην εκπαίδευση, 0 χρόνια προϋπηρεσία και 0 ώρες εργασίας. Σε αυτή τη περίπτωση το υπόδειγμα υποδεικνύει ότι  $E(\ln(WAGE)) = 0.97$  ή ισοδύναμα ότι  $E(WAGE) = \exp(0.97) \exp(0.49^2/2) = \$2.97$ . Αν αυξηθούν (μειωθούν) τα έτη στην εκπαίδευση κατά 1 χρόνο τότε ο ημερήσιος μισθός αναμένεται να αυξηθεί (μειωθεί) κατά 9.28%. Η παράμετρος είναι στατιστικά σημαντική στο 1%. Άρα υπάρχει μια θετική και στατιστική σχέση μεταξύ ετών στην εκπαίδευση και μισθού. Αν αυξηθούν (μειωθούν) τα έτη με εργασιακή εμπειρία κατά 1 χρόνο τότε ο ημερήσιος μισθός αναμένεται να αυξηθεί

(μειωθεί) κατά 0.58%. Η παράμετρος είναι στατιστικά σημαντική στο 1%. Άρα υπάρχει μια θετική και στατιστική σχέση μεταξύ ετών εργασιακής εμπειρίας και μισθού. Παρόλα αυτά παρατηρούμε ότι αν ένα άτομο έχει να διαλέξει μεταξύ 1 έτους παραπάνω στην εκπαίδευση (παρατείνοντας την παραμονή του στις εκπαιδευτικές δομές) ή 1 έτους παραπάνω εργασιακής εμπειρίας (εγκαταλείποντας γρηγορότερα τις εκπαιδευτικές δομές) τότε η πρώτη επιλογή φαίνεται πιο συμφέρουσα γιατί θα αυξήσει το μισθό του κατά 9.28% το οποίο είναι πολλαπλάσια μεγαλύτερο σε σχέση με το 0.58%. Τέλος τα αποτελέσματα του Πίνακα 1 υποδεικνύουν ότι αν αυξηθούν (μειωθούν) οι ώρες εβδομαδιαίας εργασίας κατά 1 ώρα τότε ο μισθός αναμένεται να αυξηθεί (μειωθεί) κατά 1.07%. Η παράμετρος είναι στατιστικά σημαντική στο 1%. Άρα υπάρχει μια θετική και στατιστική σχέση μεταξύ των ωρών εργασίας και του μισθού. Ο συντελεστής προσδιορισμού είναι 25% οπότε το 25% της διακύμανσης της εξαρτημένης μεταβλητής ( $\ln(WAGE)$ ) ερμηνεύεται από το υπόδειγμα, ενώ το υπόλοιπο 75% παραμένει ανεξηγήτο.

2. Ο Πίνακας 2 παρουσιάζει τον έλεγχο White για ετεροσκεδαστικότητα. Παρατηρούμε ότι τα p-values του ελέγχου (είτε χρησιμοποιήσουμε την  $F$  στατιστική είτε την  $\chi^2$  στατιστική) είναι μικρότερα από 1%, οπότε καταλήγουμε στο συμπέρασμα ότι θα απορρίψουμε την μηδενική υπόθεση της ομοσκεδαστικότητας, καταλήγοντας στο συμπέρασμα ότι ο διαταρακτικός όρος παρουσιάζει ετεροσκεδαστικότητα. Σε μια τέτοια περίπτωση είναι καλύτερα να εκτιμήσουμε το υπόδειγμα χρησιμοποιώντας robust standard errors (βλ. Πίνακα 3), διορθώνοντας έτσι τα κλασσικά τυπικά σφάλματα για την ύπαρξη ετεροσκεδαστικότητας.
3. Για την κατασκευή του διαστήματος εμπιστοσύνης χρησιμοποιούμε τα αποτελέσματα του Πίνακα 3. Οι κριτικές τιμές όπως δίνονται από την τυποποιημένη κανονική κατανομή (θυμηθείτε ότι το δείγμα είναι πολύ μεγάλο) είναι  $\pm 1.96$ . Οπότε ένα 95% διάστημα εμπιστοσύνης για την παράμετρο  $\beta_2$  είναι το εξής  $(0.092 - 1.96 \times 0.003, 0.092 + 1.96 \times 0.003) = (0.086, 0.098)$ . Άρα με επίπεδο εμπιστοσύνης 95% η πληθυσμιακή παράμετρος  $\beta_2$  λαμβάνει τιμές μεταξύ του 0.086 και του 0.094, άρα μια αύξηση των ετών εκπαίδευσης κατά 1 έτος αναμένεται να αυξήσει τον ωριαίο μισθό μεταξύ 8.6% και 9.8%. Το διάστημα αυτό είναι σχετικά μικρό, οπότε η συγκεκριμένη παράμετρος είχε εκτιμηθεί με αξιοπιστία.
4. Το υπόδειγμα προβλέπει ότι ο μέσος λογαριθμοποιημένος μισθός για έναν άνδρα δίνεται ως:

$$\begin{aligned} E(\ln(WAGE_M)) \\ &= 1.24 + 0.086 \times EDUC + 0.006 \times EXPR \\ &\quad + 0.008 \times HRSWK \end{aligned}$$

Αντίστοιχα ο μέσος λογαριθμοποιημένος μισθός για μια γυναίκα δίνεται ως:

$$\begin{aligned} E(\ln(WAGE_W)) \\ &= 1.24 - 0.52 + 0.086 \times EDUC + 0.023 \times EDUC \\ &\quad + 0.006 \times EXPR + 0.008 \times HRSWK \end{aligned}$$

Άρα η διαφορά μεταξύ των δύο είναι:

$$E(\ln(WAGE_M)) - E(\ln(WAGE_W)) = 0.52 - 0.023 \times EDUC$$

Έτσι το υπόδειγμα υποδεικνύει ότι διαφορά στον μέσο μισθό εξαρτάται από τα έτη στην εκπαίδευση. Αν για παράδειγμα τα έτη στην εκπαίδευση είναι 0, τότε η διαφορά είναι 52%, δηλαδή οι άνδρες με 0 έτη στην εκπαίδευση αμείβονται κατά 52% περισσότερο σε σχέση με τις γυναίκες. Όσο όμως τα έτη στην εκπαίδευση αυξάνονται η απόκλιση μεταξύ αντρών και γυναικών μειώνεται. Για παράδειγμα, αν και οι δύο έχουν 16 χρόνια στην εκπαίδευση τότε η διαφορά είναι ίση με  $0.52 - 0.023 \times 16 = 0.15$  ή 15%.

Η σχέση μισθού και εκπαίδευσης για τους άνδρες είναι 8.6%, ενώ για τις γυναίκες είναι  $8.6 + 2.3 = 10.9\%$ . Έτσι η αύξηση (μείωση) των ετών εκπαίδευσης κατά 1 έτος στους άνδρες αναμένεται να αυξήσει (μειώσει) τον μισθό κατά 8.6% ενώ για τις γυναίκες το ίδιο ποσοστό είναι 10.9%.

5. Για τον άνδρα έχουμε ότι:

$$E(WAGE_M) = \exp(1.24 + 0.086 \times 16 + 0.006 \times 5 + 0.008 \times 35) \exp(0.48^2/2) = \$19.01$$

Για την γυναίκα έχουμε ότι:

$$E(WAGE_W) = \exp(1.24 - 0.52 + 0.109 \times 16 + 0.006 \times 5 + 0.008 \times 35) \exp(0.48^2/2) = \$15.56$$

6. Η μηδενική υπόθεση είναι  $H_0: \beta_2 = \beta_4 = 0$ , και η εναλλακτική  $H_1: \beta_2 \neq 0$  ή  $\beta_4 \neq 0$ . Για να πραγματοποιήσουμε τον από κοινού έλεγχο υποθέσεων πρέπει να γνωρίζουμε το τετραγωνικό άθροισμα καταλοίπων του restricted και του unrestricted υποδείγματος. Η εκτίμηση του unrestricted υποδείγματος παρουσιάζεται στον Πίνακα 4, οπότε  $SSE_U = 1140.73$ . Η εκτίμηση του restricted υποδείγματος (δηλαδή όταν υποθέσουμε ότι  $\beta_2 = \beta_4 = 0$ ) δίνεται στον Πίνακα 3, οπότε  $SSE_R = 1190.22$ . Η στατιστική του ελέγχου ισούται με:

$$F = \frac{(1190.22 - 1140.73)/2}{1140.73/(4838 - 6)} = 104.8$$

Η κριτική τιμή του ελέγχου είναι  $F_{(0.95, 4832, 2)} = 3$ . Παρατηρούμε ότι  $104.8 > 3$  οπότε η μηδενική υπόθεση απορρίπτεται, άρα πράγματι η μεταβλητή *FEMALE* επηρεάζει σημαντικά το υπόδειγμα, ή με άλλα λόγια η σχέση μεταξύ των ανεξάρτητων μεταβλητών και της εξαρτημένης είναι διαφορετική για γυναίκες και άνδρες.