

## Rešenja drugog kontrolnog testa

1. Testira se  $H_0: v(\varepsilon_i) = \sigma^2$ ,  $H_1: v(\varepsilon_i) = \sigma_i^2$  za  $\forall i, i = 1, \dots, n$   
Testira se statistička značajnost ocene parametra uz promenljivu  $X_i$  primenom t testa u pomoćnoj regresiji, odnosno rezultatima iz tabele 2.

$$t = -3,49$$

Na osnovu Glejzerovog testa možemo zaključiti da se odbacuje nulta hipoteza o konstantnoj varijansi, odnosno da postoji heteroskedastičnost.

2. Nulta i alternativna hipoteza su definisane isto kao u prvom zadatku.  $WH = n * R^2 = 74 * 0,0988 = 7,31$  što je veće od kritične vrednosti  $\chi_2^2 = 5,991$ . Odbacujemo nultu hipotezu o homoskedastičnosti i zaključujemo da postoji heteroskedastičnost.
3. Korigovana ocena varijanse ocene b je  $s_{kb}^2 = \frac{\sum x_i^2 e_i^2}{(\sum x_i^2)^2} = \frac{19.191.078.912}{5.970.492} = 3.214$ . Korigovana standardna greška ocene b je  $s_{kb} = 57$ . Obratite pažnju da kada postoji heteroskedastičnost,  $s_b^2$  potcenjuje stvarnu varijansu ocene b, stoga je korigovana standardna greška ocene  $s_{kb} = 57$  veća od  $s_b = 53$  u polaznom modelu (Tabela 1).
4. Nulta hipoteza je  $H_0: \rho_1 = \rho_2 = 0$ . Na osnovu pomoćne regresije vrednost Godfri-Brojšove test statistike iznosi  $GB = T * R^2 = 248 * 0,8992 = 223$ . GB test statistika je veća od kritične vrednosti  $\chi_2^2 = 5,991$ , stoga odbacujemo  $H_0$ , i zaključujemo da postoji autokorelacija.
5. Darbin-Votsonova test statistika je definisana na sledeći način:

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} = \frac{235.916}{1.683.913} = 0,1401$$

Pošto je  $d < 2$ , sprovodimo test pozitivne autokorelacije. Testiramo hipotezu da autokorelacija prvog reda ne postoji, protiv alternativne hipoteze o prisustvu pozitivne autokorelacije.

$$H_0: \rho = 0, H_1: \rho > 0$$

Pošto su  $d_d = 1,65$  i  $d_g = 1,69$  (za uzorak velikog obima  $T > 100$ ), vidimo da je  $d < d_d < d_g$ , što znači da odbacujemo nultu hipotezu i zaključujemo da u modelu postoji pozitivna autokorelacija.

Obratite pažnju kako je definisana d statistika, odnosno da druge dve date sume se ne koriste za izračunavanje, tj.  $\sum_{t=3}^T (e_t - e_{t-2})^2$  i  $\sum_{t=2}^T (e_t - e_{t-1})$  Vam nisu potrebne.

6. Ocenjeni autokorelacioni koeficijent iznosi  $\hat{\rho} = 1 - 0,5 * d = 0,93$ . U cilju otklanjanja autokorelacije ocenjuje se model  $Y_t^* = \beta_0^* + \beta X_t^* + v_t$ , pri čemu je  $Y_t^* = Y_t - \hat{\rho} Y_{t-1}$ ,  $X_t^* = X_t - \hat{\rho} X_{t-1}$ ,  $\beta_0^* = \beta(1 - \hat{\rho})$ .
7.  $d = 1,74$ , postupak testiranja je kao u zadatku 5. Izračunata statistika je manja od 2, stoga se testira prisustvo pozitivne autokorelacije. Pošto je d statistika veća od  $d_g = 1,69$ , zaključujemo da u transformisanom modelu ne postoji autokorelacija prvog reda.

8. U modelu bez kubikaže, ocena broja pređenih milja će iznositi  $b_1^* = b_1 + b_2 \frac{\sum_{i=1}^n x_{i1}x_{i2}}{\sum_{i=1}^n x_{i1}^2}$ .

Na osnovu informacija datih u tabelama 6 i 7  $b_1^* = -121 + 11 * \frac{-27.370}{2.444} = -244$

Pristrasnost je razlika između  $b_1^*$  (ocena u modelu bez kubikaže) i  $b_1$  (ocena u modelu sa kubikažom), odnosno  $-244 + 121 = -123$ , ili  $11 * \frac{-27.370}{2.444} = -123$ .

9. Bolji je model predstavljen u tabeli 9, jer u tabeli 8 postoji irelevantna objašnjavajuća promenljiva.

$$t_1 = \frac{1,82-2}{0,850} = -0,212, t_2 = \frac{2,04-2}{0,377} = 0,106$$

Na osnovu t testa možemo zaključiti da u oba modela (tabela 8 i tabela 9), ocenjeni koeficijent za težinu ( $X_{i1}$ ) se ne razlikuje statistički značajno od 2 (kritična vrednost za t test je 2 za uzorak obima 60, za uzorak obima 120 je 1,98). Možemo da primetimo da standardna greška ocene za promenljivu težina u tabeli 9 iznosi 0,377 i manja je u odnosu na standardnu grešku ocene za promenljivu težina u tabeli 8 koja iznosi 0,850. Takođe, možemo da primetimo da kubikaža nema statistički značajan uticaj na cenu, kao i da je korigovani  $R^2$  veći u modelu bez kubikaže (0,2802) nego u modelu sa kubikažom (0,2710).

Uključivanjem irelevantne objašnjavajuće promenljive (u ovom primeru kubikaže) dobija se nepristrasna ocena druge objašnjavajuće promenljive (u ovom primeru težine), ali neefikasna jer je ocena varijanse druge objašnjavajuće promenljive (u ovom primeru težine) veća u pogrešno specifikovanom modelu.