

Analiza reoloških i kemijskih osobina ploda masline u ovisnosti o sorti i zrelosti ploda analizom varijance ponovljenih mjerenja

Markovinović, Gabriijela

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:516179>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-11-14**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Gabrijela Markovinović

**ANALIZA REOLOŠKIH I KEMIJSKIH
OSOBINA PLODA MASLINE U
OVISNOSTI O SORTI I ZRELOSTI
PLODA ANALIZOM VARIJANCE
PONOVLJENIH MJERENJA**

Diplomski rad

Voditelj rada:
prof. dr. sc. Anamarija Jazbec

Zagreb, veljača 2022.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Sadržaj

Sadržaj	iii
Uvod	1
1 Analiza varijance	2
1.1 Uvod	2
1.2 Jednofaktorska ANOVA	3
1.3 Tukey-ev <i>post hoc</i> test	5
1.4 <i>LS-means</i>	5
1.5 Dvofaktorska ANOVA ponovljenih mjerenja	6
1.6 Uvjet sferičnosti	9
2 Analiza reoloških i kemijskih osobina ploda masline	11
2.1 Opis podataka i deskriptivna statistika	11
2.2 Analiza varijance ponovljenih mjerenja	13
2.2.1 Elastičnost ploda	13
2.2.2 Tvrdća ploda	23
2.2.3 Suha tvar ploda	32
2.3 Zaključak	42
3 Dodatak	43
Bibliografija	47

Uvod

Maslinovo ulje se zbog blagotvornog djelovanja na ljudsko zdravlje sve više upotrebljava u prehrani. Povećava se potražnja za visoko kvalitetnim maslinovim uljima, a time i potreba za unaprijeđenjem proizvodnje i što boljom analizom plodova masline kako bi se odredio optimalan period berbe i uočile kvalitete pojedinih sorti masline.

U ovom radu bavit ćemo se analizom reoloških i kemijskih osobina ploda masline u ovisnosti o sorti i zrelosti ploda analizom varijance ponovljenih mjerenja. Analizirat ćemo elastičnost i tvrdoću kao reološka svojstva te suhu tvar kao kemijsko svojstvo ploda. Podaci koje ćemo analizirati prikupljeni su za projekt „Analiza teksturnog profila kao čimbenik preradbenih značajki i vrednovanja bioraznolikosti autohtonih sorti maslina” Instituta za Jadranske kulture i melioraciju krša u Splitu u sklopu kojega je provedeno istraživanje na odabranim autohtonim i introduciranim sortama maslina.

U prvom dijelu rada uvest ćemo osnovne statističke pojmove i dati teorijsku podlogu za bolje razumijevanje analize varijance ponovljenih mjerenja, a u nastavku ćemo interpretirati rezultate dobivene analizom podataka putem programskog jezika SAS.



Slika 1: Stablo masline (privatna arhiva)

Poglavlje 1

Analiza varijance

1.1 Uvod

Pretpostavimo da nam je dano k uzoraka za koje trebamo odrediti jesu li uzeti iz iste populacije. Kada bismo imali samo dva uzorka, promatrali bismo razliku njihovih očekivanja. Međutim, ukoliko bismo u slučaju većeg broja uzoraka promatrali razliku svaka dva očekivanja međusobno, to nam ne bi u potpunosti pomoglo u rješavanju problema. U opisanom slučaju upotrijebit ćemo statističku metodu analize varijance (ANOVA). Ideja metode je da se problem razlike više uzoraka svede na analizu varijabilnosti unutar uzoraka (neobjašnjena varijabilnosti) i varijabilnosti između uzoraka (objašnjena varijabilnost). Ukoliko je varijabilnost između uzoraka statistički značajno veća od varijabilnosti unutar uzoraka, zaključujemo da se radi o uzorcima iz različitih populacija.

Analizu varijance možemo koristiti ukoliko je zadovoljeno nekoliko pretpostavki:

- nezavisnost uzoraka
- homogenost varijance - populacije iz kojih su uzeti uzorci moraju imati približno jednake varijance
- normalna distribucija populacije iz koje su uzeti uzorci

ANOVA nam daje informaciju samo o tome postoji li statistički značajna razlika među uzorcima, ali ne odgovara na pitanje koji su uzorci statistički značajno različiti. Zbog toga u slučaju da postoji statistički značajna razlika među uzorcima, provodimo neki od *post hoc* testova (mi ćemo provoditi Tukey-ev test). [4]

Nas će posebno zanimati analiza varijance ponovljenih mjerenja jer se podaci koje ćemo analizirati sastoje od tri izmjere prikupljene u različitim vremenskim točkama na istim subjektima. Ovdje će očito postojati korelacija među različitim izmjerama nad istim subjektom, a pretpostavka nezavisnosti bi trebala biti zadovoljena unutar jednog ponavljanja i između subjekata.

1.2 Jednofaktorska ANOVA

Najprije ćemo pobliže objasniti model analize varijance u kojem imamo jednu zavisnu varijablu i jednu nezavisnu varijablu, odnosno faktor.

Pretpostavimo da nam je dano k ($k \geq 3$) nizova podataka

$$\begin{cases} y_{11}, & y_{12}, & y_{13}, & \dots, & y_{1n_1} \\ y_{21}, & y_{22}, & y_{23}, & \dots, & y_{2n_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{k1}, & y_{k2}, & y_{k3}, & \dots, & y_{kn_k} \end{cases} \quad (1.1)$$

pri čemu je i -ti niz dobiven mjerenjem slučajne varijable $Y_i \sim N(\mu_i, \sigma^2)$ te da su Y_1, \dots, Y_k nezavisne slučajne varijable. U tom slučaju, y_{ij} ($i = 1, \dots, k, j = 1, \dots, n_i$) može se interpretirati kao vrijednost slučajne varijable

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \quad (1.2)$$

pri čemu su $\varepsilon_{ij} \sim N(0, \sigma^2)$ nezavisne slučajne varijable.

Testiramo nul-hipotezu

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \quad (1.3)$$

naprema alternativnoj hipotezi da u (1.3) postoje i i j takvi da je $\mu_i \neq \mu_j$.

Stavimo

$$n = \sum_{i=1}^k n_i \quad (1.4)$$

$$\mu = \frac{1}{n} \sum_{i=1}^k n_i \mu_i \quad (1.5)$$

$$\delta_i = \mu_i - \mu, \quad i = 1, \dots, k. \quad (1.6)$$

Uobičajeno je da se veličina μ zove opća srednja vrijednost, a δ_i efekt i -te razine djelujućeg faktora.

Sada jednadžbu (1.2) zapisujemo kao

$$Y_{ij} = \mu + \delta_i + \varepsilon_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i, \quad (1.7)$$

a nul-hipoteza iz (1.3) može se zapisati kao

$$H_0 : \delta_1 = \delta_2 = \dots = \delta_k = 0. \quad (1.8)$$

Uvodimo sljedeće oznake

$$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}, \quad i = 1, \dots, k, \quad (1.9)$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij} = \frac{1}{n} \sum_{i=1}^k n_i \bar{Y}_i, \quad (1.10)$$

$$SST = \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})^2, \quad (1.11)$$

$$SSE = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2, \quad (1.12)$$

$$SS = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2 = SST + SSE, \quad (1.13)$$

pri čemu je \bar{Y}_i aritmetička sredina i -tog niza podataka (i -te grupe), SST je suma kvadrata zbog tretmana, SSE suma kvadrata pogrešaka, a SS suma kvadrata svih odstupanja od sredine.

Testna statistika glasi

$$F = \frac{MST}{MSE} = \frac{SST/(k-1)}{SSE/(n-k)} \sim F(k-1, n-k). \quad (1.14)$$

[4] [5] [6]

Kako bismo preglednije zapisali rezultate dobivene analizom varijance, uobičajeno je koristiti ANOVA tablicu. Opći oblik ANOVA tablice prikazan je u tablici 1.1.

Tablica 1.1: Model jednofaktorske ANOVA-e

Izvor varijabilnosti	Stupnjevi slobode	Suma kvadrata	Varijanca	F vrijednost
Između uzoraka	$k - 1$	SST	MST	$F = \frac{MST}{MSE}$
Unutar uzoraka	$n - k$	SSE	MSE	
Ukupno	$n - 1$	SS		

1.3 Tukey-ev *post hoc* test

Pretpostavimo da nam je dano k nizova podataka definiranih s (1.1) i pretpostavimo najprije da vrijedi $n_1 = n_2 = \dots = n_k$, odnosno da imamo balansirani dizajn. Neka je w raspon tih podataka ($max - min$) te neka je s^2 procjenitelj od σ^2 s ν stupnjeva slobode i nezavisan je od Y_i za $i = 1, \dots, k$. Studentizirani raspon definiran je kao

$$q_\alpha(k, \nu) = \frac{w}{s}, \quad (1.15)$$

a njegova distribucija je tabelirana.

Rubovi Tukey-evog intervala pouzdanosti za usporedbu aritmetičkih sredina grupa po parovima su

$$\bar{y}_i - \bar{y}_j \pm q_\alpha(k, n - k) \sqrt{\frac{s^2}{n}}, \quad i, j = 1, \dots, k, i \neq j. \quad (1.16)$$

Metoda koja se koristi u slučaju nebalansiranog dizajna poznata je kao Tukey-Kramer metoda. Ukoliko pretpostavimo da su n_1 i n_2 veličine grupa koje uspoređujemo, rubovi intervala pouzdanosti računaju se po formuli

$$\bar{y}_i - \bar{y}_j \pm q_\alpha(k, n - k) \sqrt{\frac{s^2(\frac{1}{n_1} + \frac{1}{n_2})}{2}}, \quad i, j = 1, \dots, k, i \neq j. \quad (1.17)$$

[1] [4]

1.4 *LS-means*

Tukey-Kramer metoda može se koristiti za usporedbu aritmetičkih sredina grupa po parovima, ali i za usporedbu *least square* aritmetičkih sredina (*LS-means*). *LS-means* su procjene marginalnih aritmetičkih sredina za balansiranu populaciju.

Razliku između aritmetičke sredine i *LS-means*-a pokazat ćemo na primjeru dvofaktorskog modela s interakcijom. Pretpostavimo da su obavljena mjerenja mase ploda masline za dvije sorte maslina u dva maslinika. Podaci koje ćemo koristiti dani su u tablici 1.2.

Tablica 1.2: Masa ploda masline

	sorta _a	sorta _b
maslinik ₁	a_1, a_2	b_1, b_2, b_3, b_4
maslinik ₂	a_3, a_4, a_5	b_5, b_6

Želimo izračunati aritmetičku sredinu i *LS-means* za masu ploda za sorte *a* i *b*. Aritmetičke sredine po sortama računamo po formulama:

$$\bar{a} = \frac{\sum_{i=1}^5 a_i}{5}, \quad \bar{b} = \frac{\sum_{i=1}^6 b_i}{6}. \quad (1.18)$$

LS-means po sortama računa se tako da prvo izračunamo aritmetičku sredinu za prvi i drugi maslinik, a zatim aritmetičku sredinu dobivenih rezultata, odnosno

$$a_{LS-means} = \frac{\frac{\sum_{i=1}^2 a_i}{2} + \frac{\sum_{j=3}^5 a_j}{3}}{2}, \quad b_{LS-means} = \frac{\frac{\sum_{i=1}^4 b_i}{4} + \frac{\sum_{j=5}^6 b_j}{2}}{2}. \quad (1.19)$$

Iz formula (1.18) i (1.19) vidimo da bi vrijednosti aritmetičkih sredina i *LS-means*-a bile jednake u slučaju balansirano dizajna, odnosno kada bismo u svakom masliniku imali jednak broj izmjera. U slučaju nebalansiranog dizajna, *LS-means* uzima u obzir i ostale efekte (u našem primjeru je to efekt maslinika). [1], [4]

1.5 Dvofaktorska ANOVA ponovljenih mjerenja

Budući da ćemo prilikom analize osobina ploda masline imati dva faktora, sortu kao nezavisan faktor i rok berbe koji označava ponavljanje, proučit ćemo model dvofaktorske analize varijance ponovljenih mjerenja.

Pretpostavimo da su provedena mjerenja u *t* vremenskih točaka za *s* grupa subjekata.

Tablica 1.3: Opažanja na subjektima unutar grupe *h*

	Subjekt	b_1	...	b_j	...	b_t
a_h	1	y_{h11}	...	y_{h1j}	...	y_{h1t}
	⋮	⋮		⋮		⋮
	<i>i</i>	y_{hi1}	...	y_{hij}	...	y_{hit}
	⋮	⋮		⋮		⋮
	n_h	y_{hn_1}	...	y_{hn_j}	...	y_{hn_t}

Neka je n_h broj subjekata u grupi *h* i $n = \sum_{h=1}^s n_h$. Neka y_{hij} predstavlja mjerenje u trenutku *j* na *i*-tom subjektu koji pripada grupi *h* za $h = 1, \dots, s$, $i = 1, \dots, n_h$ i $j = 1, \dots, t$. U tablici 1.3 dan je prikaz mjerenja na subjektima koji pripadaju *h*-toj grupi.

Prema [2], model možemo zapisati kao

$$Y_{hij} = \mu + \alpha_h + \beta_j + (\alpha\beta)_{hj} + \pi_{i(h)} + \varepsilon_{hij}, \quad (1.20)$$

pri čemu μ označava ukupnu aritmetičku sredinu, α_h efekt h -te grupe, β_j efekt j -tog ponavljanja, a $(\alpha\beta)_{hj}$ efekt interakcije h -te grupe i j -tog ponavljanja. Pritom pretpostavimo da vrijedi $\sum_{h=1}^s \alpha_h = 0$ i $\sum_{j=1}^t \beta_j = 0$. Interakciju efekata možemo interpretirati kao promatranje ponašaju li se grupe jednako tijekom vremena. Parametri $\pi_{i(h)}$ označavaju slučajni efekt za i -ti subjekt u h -toj grupi i pretpostavljamo da su to nezavisne normalno distribuirane varijable s očekivanjem 0 i varijancom σ_π^2 . Na kraju, pretpostavljamo da su $\varepsilon_{hij} \sim N(0, \sigma_\varepsilon^2)$ nezavisne slučajne varijable.

U svrhu lakšeg definiranja suma kvadrata, uvodimo oznake (1.21) - (1.25) kojima su redom označene ukupna aritmetička sredina, aritmetička sredina grupe h , aritmetička sredina j -tog ponavljanja, aritmetička sredina grupe h za j -to ponavljanje te aritmetička sredina za i -ti subjekt u grupi h .

$$\bar{Y}_{...} = \frac{\sum_{h=1}^s \sum_{i=1}^{n_h} \sum_{j=1}^t Y_{hij}}{nt} \quad (1.21)$$

$$\bar{Y}_{h..} = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^t Y_{hij}}{n_h t} \quad (1.22)$$

$$\bar{Y}_{..j} = \frac{\sum_{h=1}^s \sum_{i=1}^{n_h} Y_{hij}}{n} \quad (1.23)$$

$$\bar{Y}_{h.j} = \frac{\sum_{i=1}^{n_h} Y_{hij}}{n_h} \quad (1.24)$$

$$\bar{Y}_{hi.} = \frac{\sum_{j=1}^t Y_{hij}}{t} \quad (1.25)$$

Ukupnu sumu kvadrata svih odstupanja od sredine (SS) u ovom slučaju možemo podijeliti na sumu kvadrata između subjekata ($SS_{između}$) i sumu kvadrata unutar subjekata (SS_{unutar}). Prvi izvor varijabilnosti proizlazi iz činjenice da su promatrani subjekti različiti, a drugi se definira tako da se za svaki subjekt izračuna kvadratna udaljenost od njegove aritmetičke sredine, a suma svih odstupanja za sve subjekte je preostala varijabilnost. Sada sume kvadrata između i unutar subjekata možemo rastaviti kao

$$SS_{između} = SSG + SSE_G, \quad (1.26)$$

$$SS_{unutar} = SSP + SSGP + SSE_P, \quad (1.27)$$

pri čemu su korištene sume kvadrata definirane s

$$SSG = \sum_{h=1}^s \sum_{i=1}^{n_h} \sum_{j=1}^t (\bar{Y}_{h..} - \bar{Y}_{...})^2 = t \sum_{h=1}^s n_h (\bar{Y}_{h..} - \bar{Y}_{...})^2, \quad (1.28)$$

$$SSE_G = \sum_{h=1}^s \sum_{i=1}^{n_h} \sum_{j=1}^t (\bar{Y}_{hi.} - \bar{Y}_{h..})^2 = t \sum_{h=1}^s \sum_{i=1}^{n_h} (\bar{Y}_{hi.} - \bar{Y}_{h..})^2, \quad (1.29)$$

$$SSP = \sum_{h=1}^s \sum_{i=1}^{n_h} \sum_{j=1}^t (\bar{Y}_{..j} - \bar{Y}_{...})^2 = n \sum_{j=1}^t (\bar{Y}_{..j} - \bar{Y}_{...})^2, \quad (1.30)$$

$$SSGP = \sum_{h=1}^s \sum_{i=1}^{n_h} \sum_{j=1}^t (\bar{Y}_{h.j} - \bar{Y}_{h..} - \bar{Y}_{..j} + \bar{Y}_{...})^2, \quad (1.31)$$

$$SSE_P = \sum_{h=1}^s \sum_{i=1}^{n_h} \sum_{j=1}^t (Y_{hij} - \bar{Y}_{h.j} - \bar{Y}_{hi.} + \bar{Y}_{h..})^2. \quad (1.32)$$

F-statistika pomoću koje testiramo postoji li statistički značajna razlika između grupa definirana je s

$$F_1 = \frac{MSG}{MSE_G} = \frac{SSG/(s-1)}{SSE_G/(n-s)} \sim F(s-1, n-s). \quad (1.33)$$

Testna statistika pomoću koje testiramo postoji li statistički značajna razlika između ponavljanja definirana je s

$$F_2 = \frac{MSP}{MSE_P} = \frac{SSP/(t-1)}{SSE_P/[(n-s)(t-1)]} \sim F(t-1, (n-s)(t-1)). \quad (1.34)$$

Testna statistika pomoću koje testiramo značajnost interakcije Grupa x Ponavljanje definirana je s

$$F_3 = \frac{MSGP}{MSE_P} = \frac{SSGP/[(s-1)(t-1)]}{SSE_P/[(n-s)(t-1)]} \sim F((s-1)(t-1), (n-s)(t-1)). \quad (1.35)$$

[2] [5]

Kao i kod jednofaktorske analize varijance, sve sume kvadrata, stupnjeve slobode, vrijednosti srednjekvadratnih odstupanja te testne statistike radi preglednosti zapisujemo u ANOVA tablicu (Tablica 1.4).

Tablica 1.4: Model dvofaktorske ANOVA-e ponovljenih mjerenja

Izvor varijabilnosti	Stupnjevi slobode	Suma kvadrata	Varijanca	F vrijednost
Između subjekata				
Grupa	$s - 1$	SSG	MSG	$F = \frac{MSG}{MSE_G}$
Greška _a	$n - s$	SSE_G	MSE_G	
Unutar subjekata				
Ponavljanje	$t - 1$	SSP	MSP	$F = \frac{MSP}{MSE_P}$
Grupa x Ponavljanje	$(s - 1)(t - 1)$	$SSGP$	$MSGP$	$F = \frac{MSGP}{MSE_P}$
Greška _b	$(n - s)(t - 1)$	SSE_P	MSE_P	

1.6 Uvjet sferičnosti

Prilikom provođenja analize varijance ponovljenih mjerenja, dodatan uvjet koji treba provjeriti jest uvjet sferičnosti. Taj uvjet možemo poistovjetiti s ranije spomenutim uvjetom homogenosti varijance kod analize varijance bez ponovljenih mjerenja. Prema [2], uvjet sferičnosti možemo zapisati na dva načina:

1. $Var(Y_{ij} - Y_{i'j'})$ je konstantna za sve j, j' pri čemu je Y_{ij} j -to ponovljeno mjerenje na i -tom subjektu.
2. $\varepsilon = 1$, pri čemu je

$$\varepsilon = \frac{t^2(\bar{\sigma}_{ii} - \bar{\sigma}_{..})^2}{(t - 1)(S - 2t \sum \bar{\sigma}_i^2 + t^2 \bar{\sigma}_{..}^2)}. \quad (1.36)$$

i $\bar{\sigma}_{ii}$ označava aritmetičku sredinu dijagonalnih elemenata kovarijacijske matrice modela Σ , $\bar{\sigma}_{..}$ označava aritmetičku sredinu svih elemenata matrice Σ , $\bar{\sigma}_i$ označava aritmetičku sredinu elemenata i -tog retka matrice Σ , S je suma kvadrata svih elemenata matrice Σ , a t je broj ponovljenih mjerenja.

Prvi uvjet poznat je kao Huynh-Feldtov uvjet, a epsilon koji je dan formulom (1.36) naziva se Box-ov epsilon.

Ukoliko je uvjet sferičnosti zadovoljen, prilikom testiranja postoji li statistički značajna razlika između ponavljanja i testiranja značajnosti interakcije, koristimo testne statistike (1.34) i (1.35).

Ukoliko uvjet sferičnosti nije zadovoljen, koristimo prilagodbe stupnjeva slobode F statistike. Prema [7], može se pokazati da za epsilon iz (1.36) vrijedi

$$\frac{1}{t-1} \leq \varepsilon \leq 1. \quad (1.37)$$

Greenhouse i Geisser ([7]) preporučuju provođenje niza koraka koje navodimo u nastavku. Najprije uspoređujemo realizaciju testne statistike s kritičnom vrijednosti iz F distribucije s brojem stupnjeva slobode bez prilagodbe. Dakle, u prvom koraku za testnu statistiku F_2 iz (1.34) uzimamo da je broj stupnjeva slobode $t-1$ i $(n-s)(t-1)$, a za testnu statistiku F_3 iz (1.35) $(s-1)(t-1)$ i $(n-s)(t-1)$. Ukoliko je realizacija testne statistike manja od izračunate kritične vrijednosti, tada ne odbacujemo nultu hipotezu i zaustavljamo postupak. U slučaju da se realizacija testne statistike nalazi u kritičnom području, realizaciju testne statistike uspoređujemo s kritičnom vrijednosti iz F distribucije s brojem stupnjeva slobode pomnoženima faktorom $\frac{1}{t-1}$. Ovdje ćemo za F_2 koristiti distribuciju $F(1, n-s)$, a za F_3 distribuciju $F(s-1, n-s)$. Ukoliko se realizacija testne statistike nalazi u kritičnom području, odbacujemo nultu hipotezu na nivou značajnosti α i zaustavljamo postupak. Međutim, ukoliko ne možemo odbaciti nultu hipotezu, potrebno je izračunati procjenu Box-ovog epsilon, $\hat{\varepsilon}$, koristeći uzoračku kovarijacijsku matricu i prilagoditi stupnjeve slobode F statistike pomoću izračunate aproksimacije parametra epsilon koja se naziva Greenhouse-Geisserov epsilon. Zatim se ponovno provodi test koristeći izračunatu prilagodbu stupnjeva slobode F statistike.

Hyunh i Feldt ([3]) predlažu sljedeću procjenu Box-ovog epsilon:

$$\tilde{\varepsilon} = \min \left\{ 1, \frac{n(t-1)\hat{\varepsilon} - 2}{(t-1)(n-1-(t-1)\hat{\varepsilon})} \right\}, \quad (1.38)$$

pri čemu je $\hat{\varepsilon}$ ranije definirani Greenhouse-Geisserov epsilon. Procjenu $\tilde{\varepsilon}$ nazivamo Hyunh-Feldtov epsilon. Naime, prema [3], $\hat{\varepsilon}$ je manje pristrana procjena za Box-ov epsilon kada je $\varepsilon \leq 0.5$, a $\tilde{\varepsilon}$ je manje pristrana procjena u slučaju da je $\varepsilon \geq 0.75$. Međutim, vrijednost Box-ovog epsilon u praksi nije poznata.

Procedura *glm* koju ćemo koristiti za provedbu analize varijance u SAS-u, daje nam i Greenhouse-Geisserovu i Hyunh-Feldtovu prilagodbu stupnjeva slobode, a mi ćemo koristiti Greenhouse-Geisserovu prilagodbu.

Poglavlje 2

Analiza reoloških i kemijskih osobina ploda masline

2.1 Opis podataka i deskriptivna statistika



Slika 2.1: Plod sorte oblica (privatna arhiva)

Podaci koje ćemo analizirati prikupljeni su u vremenskom razdoblju od 18. rujna 2019. do 29. listopada 2019. u masliniku OPG-a Jagoda Alfirević u Kaštel Sućurcu. Promatrat ćemo četiri sorte: oblica, leccino, lastovka i coratina te tri roka berbe koji su određeni prema obojenosti ploda masline tijekom dozrijevanja - u prvom roku je plod nezreli, u drugom optimalno zreli, a u trećem prezreli. U svakom roku berbe s određenog broja stabala maslina iz maslinika ubran je određeni broj plodova te su za ubrane plodove utvrđene vrijednosti tvrdoće, elastičnosti i suhe tvari.

U tablicama 2.2, 2.3 i 2.4 dana je deskriptivna statistika. Možemo primijetiti da je prikupljeno zamjetno manje podataka o suhoj tvari u odnosu na podatke o elastičnosti i tvrdoći ploda masline. To je opravdano činjenicom da je svojstvo suhe tvari manje varijabilno pa je za svako promatrano stablo suha tvar određena samo za jedan plod masline, dok su elastičnost i tvrdoća određeni za veći broj plodova po stablu.

Tablica 2.2: Deskriptivna statistika za elastičnost ploda u tri roka berbe (Ispis iz SAS-a)

sorta	Variable	N	Mean	Std Dev	Minimum	Maximum	Median
coratina	elasticnost_1	88	1.0162090	0.0541711	0.9940080	1.2814203	0.9996157
	elasticnost_2	86	1.0228881	0.1367514	0.5191729	1.9137258	1.0002976
	elasticnost_3	84	0.9999067	0.1258224	0.0478569	1.2342944	1.0008185
lastovka	elasticnost_1	75	1.0508236	0.1192568	0.9809375	1.6172751	1.0015734
	elasticnost_2	75	1.0597504	0.1169851	0.9858317	1.4073726	1.0043895
	elasticnost_3	73	1.0799195	0.1540788	0.9572162	1.6208645	1.0016916
leccino	elasticnost_1	101	1.0053177	0.0379811	0.9376428	1.3460640	0.9995463
	elasticnost_2	99	0.9977205	0.0720291	0.7042453	1.5409969	0.9999916
	elasticnost_3	99	1.0107472	0.0607307	0.9936272	1.5564461	0.9998505
oblica	elasticnost_1	89	1.0035081	0.0768446	0.5415543	1.3564858	1.0004528
	elasticnost_2	86	1.0035952	0.0800817	0.3996950	1.2547603	1.0012842
	elasticnost_3	86	0.9977828	0.0505179	0.5459950	1.0578469	1.0005333

Tablica 2.3: Deskriptivna statistika za tvrdoću ploda u tri roka berbe (Ispis iz SAS-a)

sorta	Variable	N	Mean	Std Dev	Minimum	Maximum	Median
coratina	tvrdoca_1	88	65.6504512	47.7819008	11.5053989	145.4072285	49.6002425
	tvrdoca_2	86	68.1843354	44.2112003	8.8224570	176.7239774	69.6862061
	tvrdoca_3	84	69.0505098	48.8819290	11.3577207	165.3187490	55.4848672
lastovka	tvrdoca_1	75	59.7065829	17.0023706	21.5561950	98.8595786	59.8327368
	tvrdoca_2	75	61.1754289	20.4303839	23.8746124	142.4839963	61.0974293
	tvrdoca_3	73	59.0328364	21.8628122	18.6037949	159.5404554	59.6378032
leccino	tvrdoca_1	101	83.8263781	60.1000470	7.9913759	230.5408903	88.1596647
	tvrdoca_2	99	86.1629974	55.8000807	10.1017603	207.2940846	92.3287086
	tvrdoca_3	99	86.1152622	59.0763579	8.3655878	197.5632822	95.0524260
oblica	tvrdoca_1	89	60.3456440	36.3477971	16.5652934	131.9740489	41.9087503
	tvrdoca_2	86	62.4977324	37.7419788	15.6343342	149.8671263	42.3737228
	tvrdoca_3	86	60.9936186	36.2047152	14.4524027	164.4885101	48.8288401

Tablica 2.4: Deskriptivna statistika za suhu tvar ploda u tri roka berbe (Ispis iz SAS-a)

sorta	Variable	N	Mean	Std Dev	Minimum	Maximum	Median
coratina	suha_tvar_1	4	63.2037469	4.9034655	59.4807122	70.4175258	61.4583748
	suha_tvar_2	4	44.0195410	0.9142426	42.7646454	44.8914432	44.2110377
	suha_tvar_3	3	51.8872452	1.7388447	50.7658834	53.8903171	51.0055351
lastovka	suha_tvar_1	4	66.8521535	0.8363803	66.1488512	67.8571429	66.7013100
	suha_tvar_2	4	62.4262146	0.7150093	61.5772727	63.3022571	62.4126644
	suha_tvar_3	4	64.0958252	0.9776320	63.1787109	65.3539382	63.9253258
leccino	suha_tvar_1	6	67.0802691	1.2626409	65.5794206	69.0112705	66.9851385
	suha_tvar_2	4	65.8786132	0.6123150	65.3532875	66.7468720	65.7071467
	suha_tvar_3	3	62.3653260	2.9055085	59.6343780	65.4185268	62.0430733
oblica	suha_tvar_1	6	60.0850910	1.0956605	58.2012467	61.0657119	60.3827752
	suha_tvar_2	4	59.0124555	2.0913425	56.3798220	61.4593781	59.1053109
	suha_tvar_3	4	64.7097871	2.4067226	61.6392318	67.0488889	65.0755139

2.2 Analiza varijance ponovljenih mjerenja

Provodimo analizu varijance ponovljenih mjerenja koristeći proceduru *glm* u programskom jeziku SAS za elastičnost, tvrdoću i suhu tvar u ovisnosti o sorti i zrelosti ploda. Za svaku zavisnu varijablu najprije ćemo interpretirati rezultate jednofaktorske analize varijance u svakom od tri roka berbe, zatim rezultate Tukey-evog *post hoc* testa i testa sferičnosti, a na kraju i protumačiti rezultate analize varijance ponovljenih mjerenja. Prilikom interpretacije rezultata provedenih testova, za razinu značajnosti koristit ćemo vrijednost od 5%.

Prilikom interpretacije ispisa iz SAS-a za analizu varijance ponovljenih mjerenja, potrebno je posebno obratiti pozornost na tablicu s rezultatima testa sferičnosti (engl. *Sphericity Tests*), tablicu između subjekata (engl. *Between Subjects Effects*) te na tablicu unutar subjekata (engl. *Within Subject Effects*). U tablici s rezultatima testa sferičnosti dani su rezultati testa za transformirane podatke i za ortogonalne komponente. Prema [1], potrebno je analizirati rezultate testa za ortogonalne komponente. Ukoliko je p-vrijednost manja od odabrane razine značajnosti, uvjet sferičnosti nije zadovoljen. U tom slučaju pri interpretaciji rezultata iz tablice unutar subjekata gledamo Greenhouse-Geisserovu prilagodbu p-vrijednosti (u tablici unutar subjekata označena s G-G) koja je dobivena prilagodbom stupnjeva slobode koristeći Greenhouse-Geisserov epsilon. U tablici između subjekata testira se ima li razlike među sortama za promatranu zavisnu varijablu. Test se provodi tako da se uprosječe tri izmjere zavisne varijable za tri roka berbe te se testira postoji li statistički značajna razlika u vrijednosti zavisne varijable među četiri sorte. U tablici unutar subjekata testira se postoji li statistički značajna razlika za zavisnu varijablu između rokova berbe na način da se uprosječe vrijednosti za svaki rok berbe i testira se postoji li statistički značajna razlika između tih prosjeka. Nadalje, testira se ponaša li se zavisna varijabla približno jednako po sortama za različite rokove berbe. [4]

2.2.1 Elastičnost ploda

Promotrimo najprije rezultate jednofaktorskih analiza varijance za elastičnost ploda po rokovima berbe u ovisnosti o sorti i rezultate Tukey-evog *post hoc* testa za svaki od tri roka berbe.

Tablica 2.5: Rezultati jednofaktorske analize varijance za elastičnost ploda u prvom roku berbe (Ispis iz SAS-a)

Dependent Variable: elasticnost_1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.11232832	0.03744277	6.48	0.0003
Error	338	1.95302000	0.00577817		
Corrected Total	341	2.06534831			

R-Square	Coeff Var	Root MSE	elasticnost_1 Mean
0.054387	7.466961	0.076014	1.018008

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	0.11232832	0.03744277	6.48	0.0003

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	0.11232832	0.03744277	6.48	0.0003

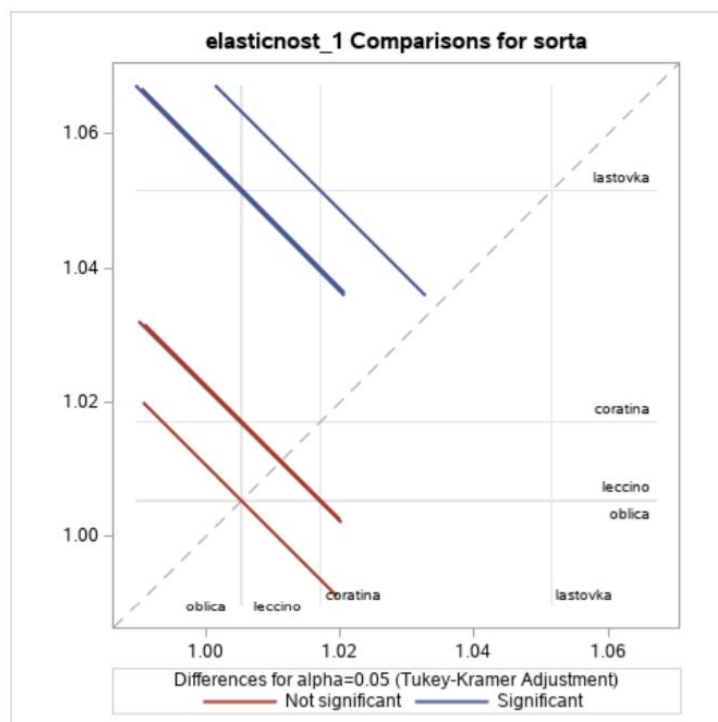
U tablici 2.5 prikazan je ispis rezultata jednofaktorske analize varijance u prvom roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance: koeficijenta determinacije (engl. *R-square*), koeficijenta varijacije (oznaka u tablici: engl. *Coeff Var*) te aritmetičke sredine. Također, dane su i sume kvadrata tipa I i III (engl. *Type I SS*, *Type III SS*) čiji je iznos u ovom slučaju jednak jer se radi o jednofaktorskoj analizi varijance. Iz rezultata u tablici 2.5 vidimo da postoji statistički značajna razlika među sortama za elastičnost ploda u prvom roku berbe (p-vrijednost je 0.0003). Koeficijent determinacije iznosi 0.054387 pa zaključujemo da je mali dio varijabilnosti zavisne varijable (oko 5%) objašnjen ovim modelom.

U tablici 2.6 nalaze se rezultati Tukey-evog *post hoc* testa za prvi rok berbe. Vidimo da statistički značajnu razliku među sortama čini lastovka obzirom na preostale tri promatrane sorte.

Tablica 2.6: Rezultati Tukey-evog *post hoc* testa za elastičnost ploda u prvom roku berbe (Ispis iz SAS-a)

sorta	elasticnost_1 LSMEAN	LSMEAN Number
coratina	1.01694236	1
lastovka	1.05155187	2
leccino	1.00544396	3
oblica	1.00503832	4

Least Squares Means for effect sorta Pr > t for H0: LSMean(i)=LSMean(j) Dependent Variable: elasticnost_1				
i/j	1	2	3	4
1		0.0242	0.7380	0.7373
2	0.0242		0.0006	0.0008
3	0.7380	0.0006		1.0000
4	0.7373	0.0008	1.0000	



Slika 2.2: Grafički prikaz rezultata Tukey-evog *post hoc* testa za elastičnost ploda u prvom roku berbe (Ispis iz SAS-a)

Tablica 2.7: Rezultati jednofaktorske analize varijance za elastičnost ploda u drugom roku berbe (Ispis iz SAS-a)

Dependent Variable: elasticnost_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.19705533	0.06568511	6.08	0.0005
Error	338	3.64884198	0.01079539		
Corrected Total	341	3.84589732			

R-Square	Coeff Var	Root MSE	elasticnost_2 Mean
0.051238	10.19570	0.103901	1.019066

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	0.19705533	0.06568511	6.08	0.0005

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	0.19705533	0.06568511	6.08	0.0005

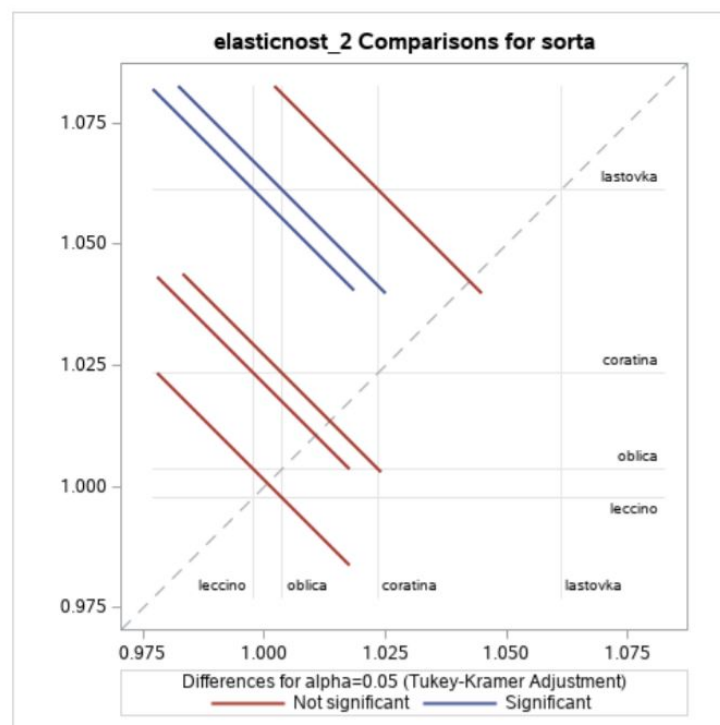
U tablici 2.7 prikazan je ispis rezultata jednofaktorske analize varijance u drugom roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Iz rezultata u tablici 2.7 vidimo da postoji statistički značajna razlika među sortama za elastičnost ploda u drugom roku berbe (p-vrijednost je 0.0005). Primijetimo da je iznos F vrijednosti približno jednak u prvom i drugom roku berbe (u prvom roku berbe iznosi 6.48, a u drugom 6.08). Iznos koeficijenta determinacije je ponovno oko 5% kao i u prvom roku berbe pa zaključujemo da je mali dio varijabilnosti zavisne varijable objašnjen ovim modelom.

U tablici 2.8 nalaze se rezultati Tukey-evog *post hoc* testa za drugi rok berbe. Vidimo da statistički značajnu razliku među sortama čini lastovka obzirom na oblicu i leccino.

Tablica 2.8: Rezultati Tukey-evog *post hoc* testa za elastičnost ploda u drugom roku berbe (Ispis iz SAS-a)

sorta	elasticnost_2 LSMEAN	LSMEAN Number
coratina	1.02341983	1
lastovka	1.06122873	2
leccino	0.99772047	3
oblica	1.00359521	4

Least Squares Means for effect sorta Pr > t for H0: LSMean(i)=LSMean(j)				
Dependent Variable: elasticnost_2				
i/j	1	2	3	4
1		0.1061	0.3426	0.5994
2	0.1061		0.0005	0.0031
3	0.3426	0.0005		0.9808
4	0.5994	0.0031	0.9808	



Slika 2.3: Grafički prikaz rezultata Tukey-evog *post hoc* testa za elastičnost ploda u drugom roku berbe (Ispis iz SAS-a)

Tablica 2.9: Rezultati jednofaktorske analize varijance za elastičnost ploda u trećem roku berbe (Ispis iz SAS-a)

Dependent Variable: elasticnost_3

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.34687748	0.11562583	10.85	<.0001
Error	338	3.60166568	0.01065582		
Corrected Total	341	3.94854316			

R-Square	Coeff Var	Root MSE	elasticnost_3 Mean
0.087849	10.12437	0.103227	1.019589

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	0.34687748	0.11562583	10.85	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	0.34687748	0.11562583	10.85	<.0001

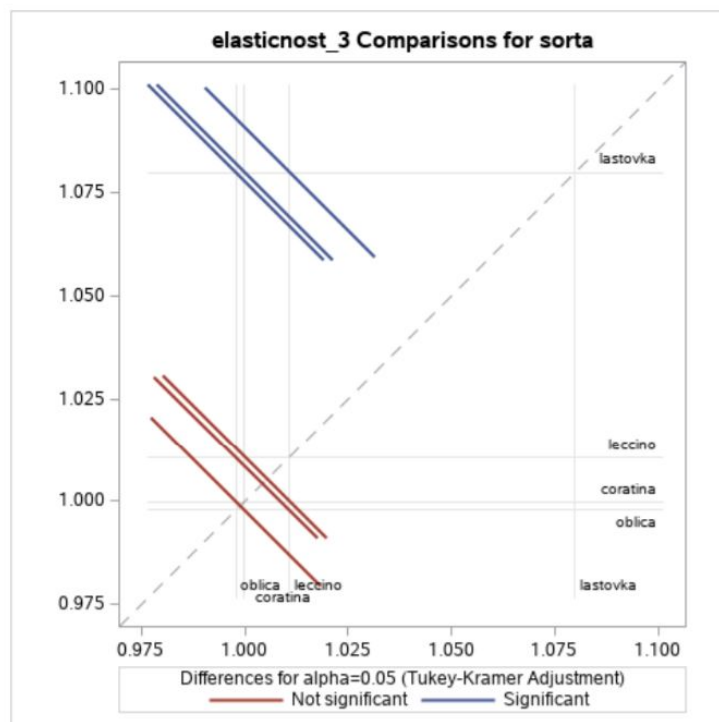
U tablici 2.9 prikazan je ispis rezultata jednofaktorske analize varijance u trećem roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Vidimo da postoji statistički značajna razlika među sortama za elastičnost ploda u trećem roku berbe (p-vrijednost je manja od 0.0001). U usporedbi s vrijednostima u prvom i drugom roku berbe, F vrijednost ima veći iznos (10.85) što znači da je u trećem roku berbe razlika među sortama veća. Koeficijent determinacije ima malo veću vrijednost nego u prethodna dva roka berbe (iznosi 0.087849), ali i dalje zaključujemo da je mali dio varijabilnosti zavisne varijable objašnjen ovim modelom.

U tablici 2.10 nalaze se rezultati Tukey-evog *post hoc* testa za drugi rok berbe. Vidimo da statistički značajnu razliku među sortama čini lastovka obzirom na preostale tri promatrane sorte.

Tablica 2.10: Rezultati Tukey-evog *post hoc* testa za elastičnost ploda u trećem roku berbe (Ispis iz SAS-a)

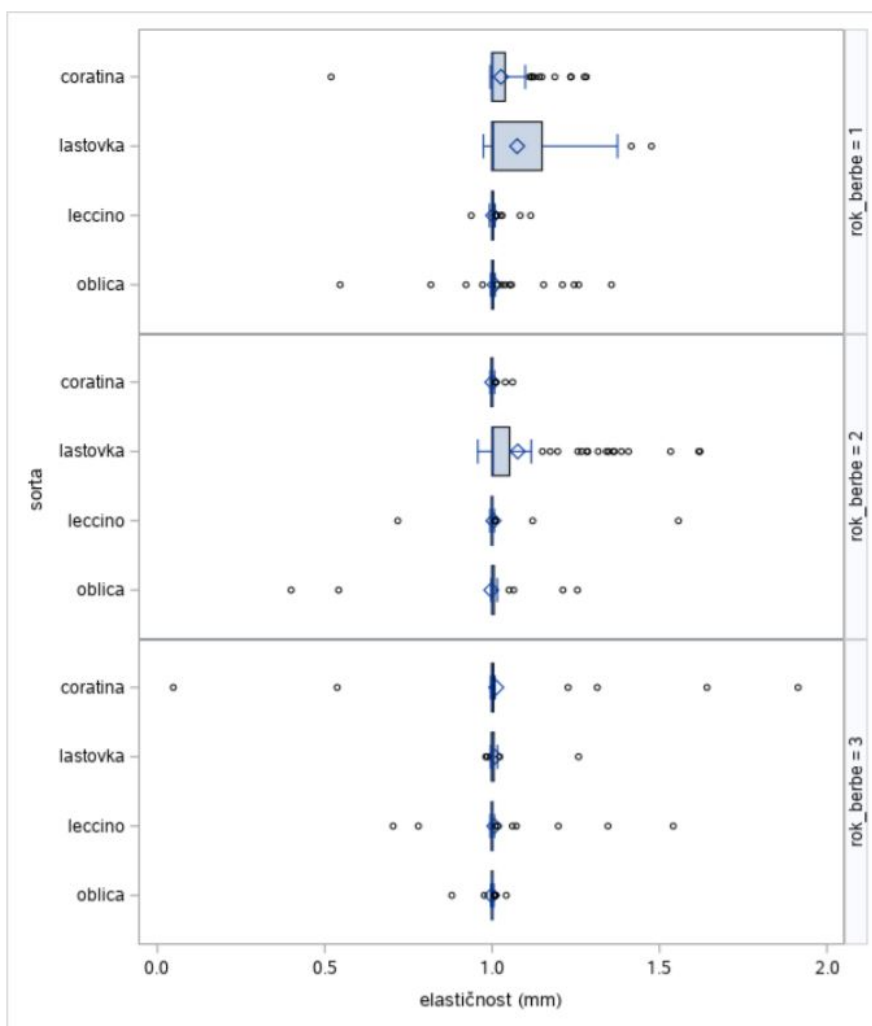
sorta	elasticnost_3 LSMEAN	LSMEAN Number
coratina	0.99990669	1
lastovka	1.07991947	2
leccino	1.01074717	3
oblica	0.99778278	4

Least Squares Means for effect sorta Pr > t for H0: LSMean(i)=LSMean(j)				
Dependent Variable: elasticnost_3				
i/j	1	2	3	4
1		<.0001	0.8939	0.9991
2	<.0001		0.0001	<.0001
3	0.8939	0.0001		0.8294
4	0.9991	<.0001	0.8294	



Slika 2.4: Grafički prikaz rezultata Tukey-evog *post hoc* testa za elastičnost ploda u trećem roku berbe (Ispis iz SAS-a)

Na slici 2.5 prikazan je *boxplot* za elastičnost ploda s podjelom po rokovima berbe i sortama. Najveću razliku u elastičnosti ploda za različite rokove berbe primjećujemo kod sorte lastovka. Za navedenu sortu se razlika između gornjeg i donjeg kvartila smanjuje od prvog do trećeg roka berbe iz čega zaključujemo da se sazrijevanjem ploda smanjuje varijabilnost elastičnosti ploda.



Slika 2.5: *Boxplot* prikaz elastičnosti ploda s podjelom po rokovima berbe i sortama (Ispis iz SAS-a). Sivi pravokutnik označava dio između donjeg (q_L) i gornjeg kvartila (q_U), plava okomita linija unutar pravokutnika medijan, a plavi kvadratić označava aritmetičku sredinu. *Outlieri* su označeni kružićima, a okomite plave linije van pravokutnika označavaju vrijednosti koje su udaljene najviše $1.5(q_U - q_L)$ od donjeg, odnosno gornjeg kvartila.

Tablica 2.11: Test sferičnosti za elastičnost ploda (Ispis iz SAS-a)

Sphericity Tests				
Variables	DF	Mauchly's Criterion	Chi-Square	Pr > ChiSq
Transformed Variates	2	0.5770771	185.27565	<.0001
Orthogonal Components	2	0.9171932	29.129328	<.0001

U tablici 2.11 dani su rezultati testa sferičnosti. Iz rezultata vidimo da odbacujemo nultu hipotezu da je uvjet sferičnosti zadovoljen (p-vrijednost je manja od 0.0001).

Tablica 2.12: Tablica rezultata ANOVA-e ponovljenih mjerenja između subjekata za elastičnost ploda (Ispis iz SAS-a)

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	0.59052936	0.19684312	20.82	<.0001
Error	338	3.19527683	0.00945348		

U tablici između subjekata testiramo postoji li statistički značajna razlika u elastičnosti ploda među sortama. Iz rezultata u tablici 2.12 vidimo da postoji statistički značajna razlika između sorti (p-vrijednost je manja od 0.0001). Dakle, uprosječimo li podatke po rokovima berbe za svaku sortu, vidimo da postoji statistički značajna razlika među sortama.

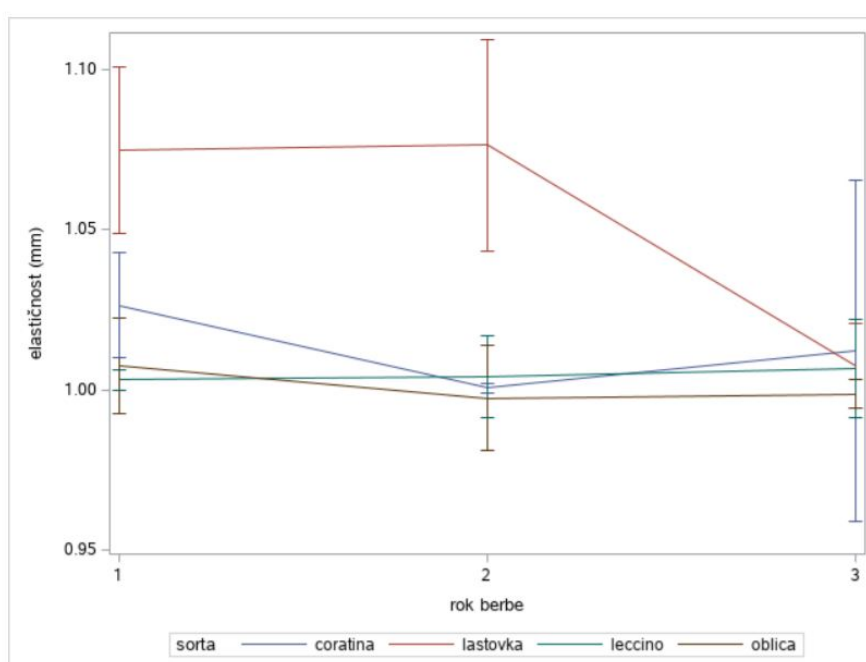
Tablica 2.13: Tablica rezultata ANOVA-e ponovljenih mjerenja unutar subjekata za elastičnost ploda (Ispis iz SAS-a)

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
rok_berbe	2	0.00100373	0.00050186	0.06	0.9451	0.9343	0.9350
rok_berbe*sorta	6	0.06573177	0.01095530	1.23	0.2874	0.2897	0.2896
Error(rok_berbe)	676	6.00825083	0.00888795				

Greenhouse-Geisser Epsilon	0.9235
Huynh-Feldt-Lecoutre Epsilon	0.9284

U tablici unutar subjekata testira se postoji li statistički značajna razlika za elastičnost ploda između rokova berbe. Nadalje, testira se ponaša li se elastičnost približno jednako po sortama za različite rokove berbe. Budući da uvjet sferičnosti nije zadovoljen, pri donošenju zaključaka gledamo Greenhouse-Geisserovu prilagodbu p-vrijednosti. Iz rezultata u tablici 2.13 vidimo da su obje prilagođene p-vrijednosti veće od 0.05 pa zaključujemo da ne postoji statistički značajna razlika između rokova berbe te da interakcija "rok_berbe*sorta" nije statistički značajna što znači da se elastičnost ploda ponaša približno jednako po rokovima berbe.



Slika 2.6: Grafički prikaz aritmetičkih sredina i 95% pouzdanih intervala za elastičnost ploda po sortama u sva tri roka berbe (Ispis iz SAS-a)

Na slici 2.6 prikazane su vrijednosti aritmetičkih sredina i pripadajućih 95% pouzdanih intervala za aritmetičku sredinu za elastičnost ploda obzirom na sortu i rok berbe. Iz grafičkog prikaza možemo iščitati da je vrijednost aritmetičke sredine elastičnosti ploda za sortu lastovka zamjetno veća u prva dva roka berbe u odnosu na preostale tri sorte, dok u trećem roku berbe zamjećujemo pad u vrijednosti elastičnosti ploda za sortu lastovka i nema zamjetne razlike među sortama. Vidimo da su vrijednosti elastičnosti ploda približno jednake po rokovima berbe, osim spomenutog pada u vrijednosti elastičnosti za sortu lastovka u trećem roku berbe. Međutim, taj pad nije statistički značajan obzirom na kretanje elastičnosti ploda ostalih sorti (Tablica 2.13).

2.2.2 Tvrdoća ploda

Promotrimo najprije rezultate jednofaktorskih analiza varijance za tvrdoću ploda po rokovima berbe u ovisnosti o sorti i rezultate Tukey-evog *post hoc* testa za svaki od tri roka berbe.

Tablica 2.14: Rezultati jednofaktorske analize varijance za tvrdoću ploda u prvom roku berbe (Ispis iz SAS-a)

Dependent Variable: tvrdoca_1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	36378.9689	12126.3230	6.05	0.0005
Error	338	677618.7849	2004.7893		
Corrected Total	341	713997.7538			

R-Square	Coeff Var	Root MSE	tvrdoca_1 Mean
0.050951	64.41097	44.77487	69.51436

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	36378.96889	12126.32296	6.05	0.0005

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	36378.96889	12126.32296	6.05	0.0005

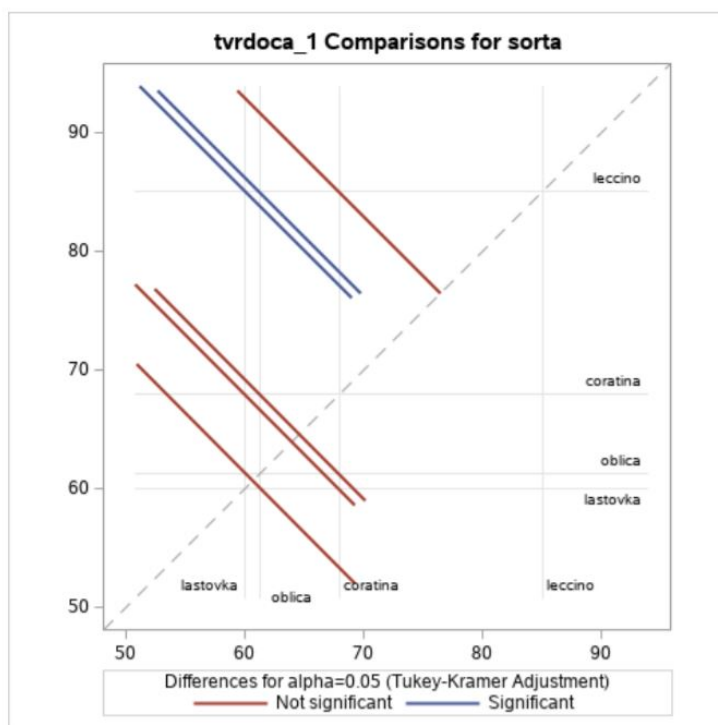
U tablici 2.14 prikazan je ispis rezultata jednofaktorske analize varijance u prvom roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Iz rezultata u tablici 2.14 vidimo da postoji statistički značajna razlika među sortama za tvrdoću ploda u prvom roku berbe (p-vrijednost je 0.0005). Koeficijent determinacije iznosi 0.050951 pa zaključujemo da je mali dio varijabilnosti zavisne varijable (oko 5%) objašnjen ovim modelom.

U tablici 2.15 nalaze se rezultati Tukey-evog *post hoc* testa za prvi rok berbe. Vidimo da statistički značajnu razliku među sortama čini leccino obzirom na lastovku i oblicu.

Tablica 2.15: Rezultati Tukey-evog *post hoc* testa za tvrdoću ploda u prvom roku berbe (Ispis iz SAS-a)

sorta	tvrdoca_1 LSMEAN	LSMEAN Number
coratina	67.9275391	1
lastovka	60.0552076	2
leccino	85.0096797	3
oblica	61.2559237	4

Least Squares Means for effect sorta Pr > t for H0: LSmean(i)=LSmean(j) Dependent Variable: tvrdoca_1				
i/j	1	2	3	4
1		0.6905	0.0513	0.7659
2	0.6905		0.0020	0.9983
3	0.0513	0.0020		0.0021
4	0.7659	0.9983	0.0021	



Slika 2.7: Grafički prikaz rezultata Tukey-evog *post hoc* testa za tvrdoću ploda u prvom roku berbe (Ispis iz SAS-a)

Tablica 2.16: Rezultati jednofaktorske analize varijance za tvrdoću ploda u drugom roku berbe (Ispis iz SAS-a)

Dependent Variable: tvrdoca_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	35716.4168	11905.4723	6.52	0.0003
Error	338	617511.8028	1826.9580		
Corrected Total	341	653228.2196			

R-Square	Coeff Var	Root MSE	tvrdoca_2 Mean
0.054677	60.34465	42.74293	70.83135

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	35716.41680	11905.47227	6.52	0.0003

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	35716.41680	11905.47227	6.52	0.0003

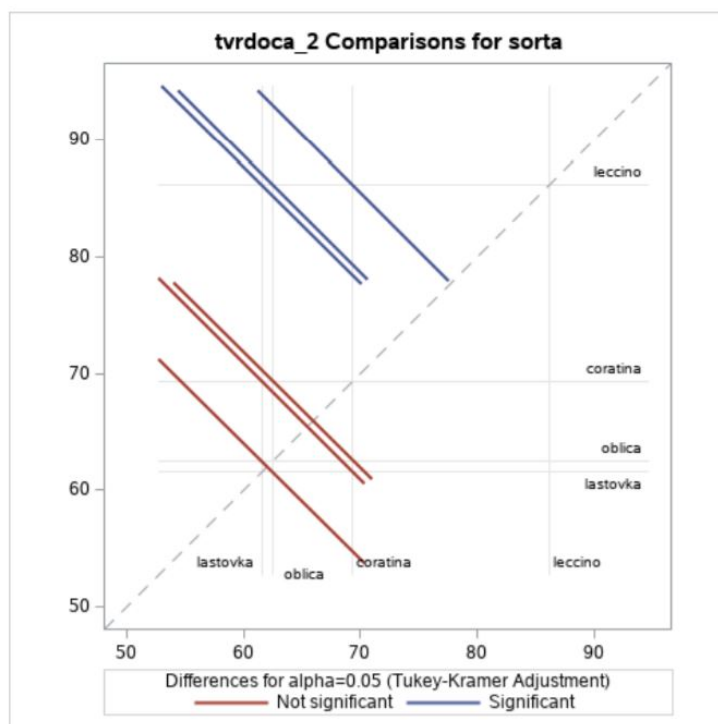
U tablici 2.16 prikazan je ispis rezultata jednofaktorske analize varijance u drugom roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Vidimo da postoji statistički značajna razlika među sortama za tvrdoću ploda u drugom roku berbe (p-vrijednost je 0.0003). Možemo zamijetiti da je iznos F vrijednosti približno jednak u prvom i drugom roku berbe (u prvom roku berbe iznosi 6.05, a u drugom 6.52). Iznos koeficijenta determinacije je ponovno oko 5% kao i u prvom roku berbe pa zaključujemo da je mali dio varijabilnosti zavisne varijable objašnjen ovim modelom.

U tablici 2.17 nalaze se rezultati Tukey-evog *post hoc* testa za drugi rok berbe. Vidimo da statistički značajnu razliku među sortama čini leccino obzirom na preostale tri sorte.

Tablica 2.17: Rezultati Tukey-evog *post hoc* testa za tvrdoću ploda u drugom roku berbe (Ispis iz SAS-a)

sorta	tvrdoća_2 LSMEAN	LSMEAN Number
coratina	69.3618418	1
lastovka	61.5477577	2
leccino	86.1629974	3
oblica	62.4977324	4

Least Squares Means for effect sorta Pr > t for H0: LSmean(i)=LSmean(j)				
Dependent Variable: tvrdoća_2				
i/j	1	2	3	4
1		0.6634	0.0418	0.7220
2	0.6634		0.0013	0.9990
3	0.0418	0.0013		0.0012
4	0.7220	0.9990	0.0012	



Slika 2.8: Grafički prikaz rezultata Tukey-evog *post hoc* testa za tvrdoću ploda u drugom roku berbe (Ispis iz SAS-a)

Tablica 2.18: Rezultati jednofaktorske analize varijance za tvrdoću ploda u trećem roku berbe (Ispis iz SAS-a)

Dependent Variable: tvrdoca_3

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	41532.0692	13844.0231	6.82	0.0002
Error	338	686176.5055	2030.1080		
Corrected Total	341	727708.5747			

R-Square	Coeff Var	Root MSE	tvrdoca_3 Mean
0.057072	64.52711	45.05672	69.82603

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	41532.06922	13844.02307	6.82	0.0002

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	41532.06922	13844.02307	6.82	0.0002

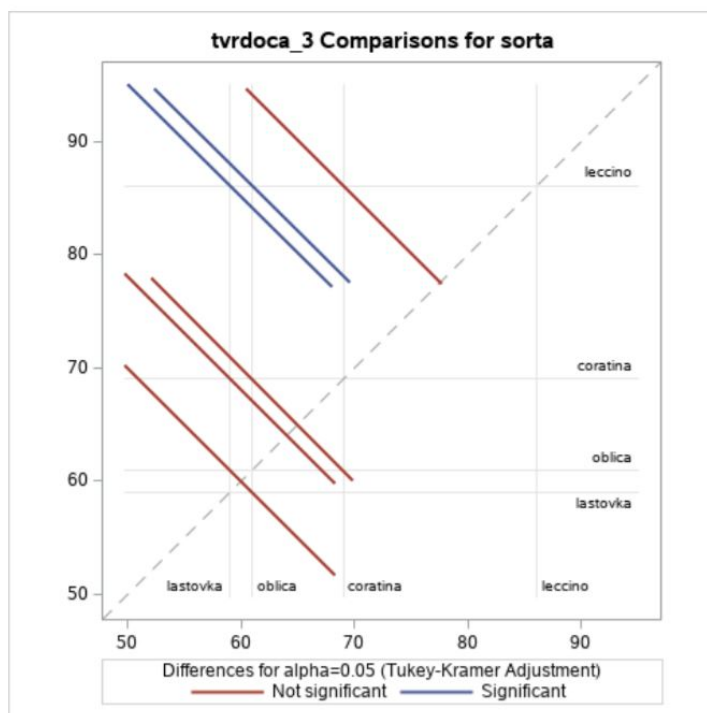
U tablici 2.18 prikazan je ispis rezultata jednofaktorske analize varijance u trećem roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Vidimo da postoji statistički značajna razlika među sortama za tvrdoću ploda u trećem roku berbe (p-vrijednost je 0.0002). Primijetimo da je iznos F vrijednosti približno jednak kao u prva dva roka berbe (iznosi 6.82) pa zaključujemo da se dozrijevanjem ploda ne povećava razlika između sorti. Iznos koeficijenta determinacije je ponovno oko 5% kao i u prethodna dva roka berbe pa zaključujemo da je mali dio varijabilnosti zavisne varijable objašnjen ovim modelom.

U tablici 2.19 nalaze se rezultati Tukey-evog *post hoc* testa za treći rok berbe. Vidimo da statistički značajnu razliku među sortama čini leccino obzirom na lastovku i oblicu.

Tablica 2.19: Rezultati Tukey-evog *post hoc* testa za tvrdoću ploda u trećem roku berbe (Ispis iz SAS-a)

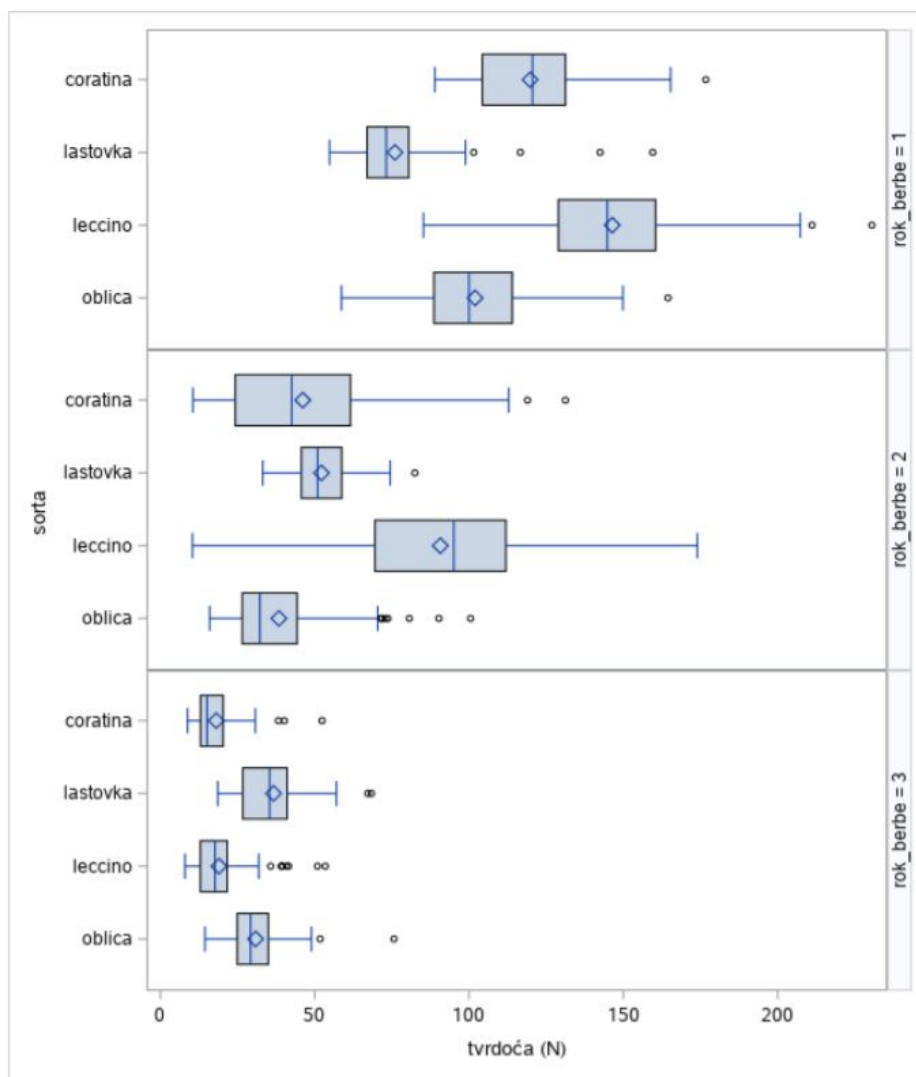
sorta	tvrdoca_3 LSMEAN	LSMEAN Number
coratina	69.0505098	1
lastovka	59.0328364	2
leccino	86.1152622	3
oblica	60.9936186	4

Least Squares Means for effect sorta Pr > t for H0: LSmean(i)=LSmean(j)				
Dependent Variable: tvrdoca_3				
ij	1	2	3	4
1		0.5068	0.0539	0.6489
2	0.5068		0.0007	0.9928
3	0.0539	0.0007		0.0010
4	0.6489	0.9928	0.0010	



Slika 2.9: Grafički prikaz rezultata Tukey-evog *post hoc* testa za tvrdoću ploda u trećem roku berbe (Ispis iz SAS-a)

Na slici 2.10 prikazan je *boxplot* za tvrdoću ploda s podjelom po rokovima berbe i sortama. Iz prikaza možemo zaključiti da se tvrdoća smanjuje sazrijevanjem ploda te da je varijabilnost najmanja za prezrele plodove.



Slika 2.10: *Boxplot* prikaz tvrdoće ploda s podjelom po rokovima berbe i sortama (Ispis iz SAS-a). Sivi pravokutnik označava dio između donjeg (q_L) i gornjeg kvartila (q_U), plava okomita linija unutar pravokutnika medijan, a plavi kvadratić označava aritmetičku sredinu. *Outlieri* su označeni kružićima, a okomite plave linije van pravokutnika označavaju vrijednosti koje su udaljene najviše $1.5(q_U - q_L)$ od donjeg, odnosno gornjeg kvartila.

Tablica 2.20: Test sferičnosti za tvrdoću ploda (Ispis iz SAS-a)

Sphericity Tests				
Variables	DF	Mauchly's Criterion	Chi-Square	Pr > ChiSq
Transformed Variates	2	0.7001197	120.14183	<.0001
Orthogonal Components	2	0.9758229	8.2477766	0.0162

U tablici 2.20 dani su rezultati testa sferičnosti. Iz rezultata vidimo da odbacujemo nultu hipotezu da je uvjet sferičnosti zadovoljen (p-vrijednost iznosi 0.0162).

Tablica 2.21: Tablica rezultata ANOVA-e ponovljenih mjerenja između subjekata za tvrdoću ploda (Ispis iz SAS-a)

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	113426.964	37808.988	7.55	<.0001
Error	338	1692834.174	5008.385		

U tablici između subjekata testiramo postoji li statistički značajna razlika u tvrdoći ploda među sortama. Iz rezultata u tablici 2.21 vidimo da postoji statistički značajna razlika između sorti (p-vrijednost je manja od 0.0001). Dakle, uprosječimo li podatke po rokovima berbe za svaku sortu, vidimo da postoji statistički značajna razlika među sortama.

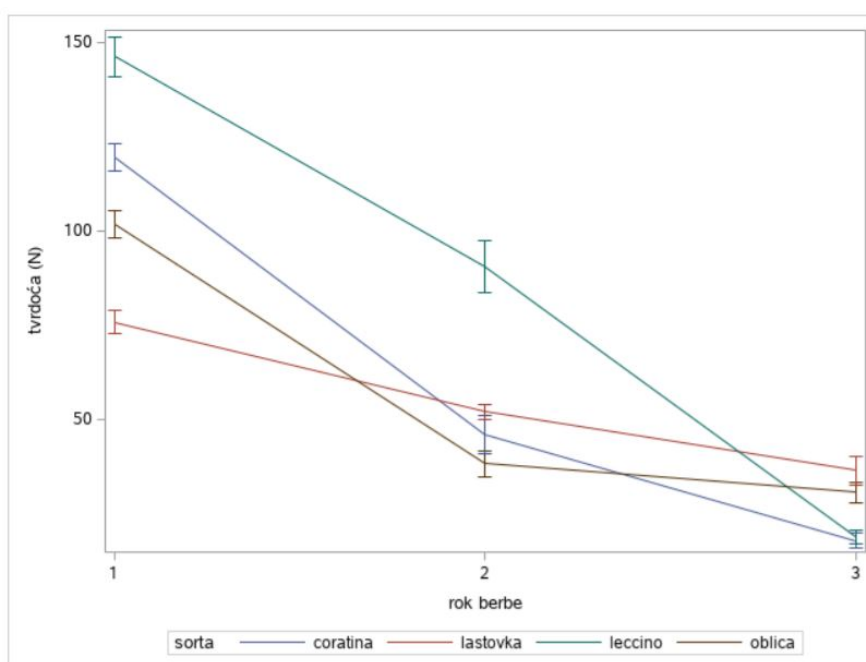
Tablica 2.22: Tablica rezultata ANOVA-e ponovljenih mjerenja unutar subjekata za tvrdoću ploda (Ispis iz SAS-a)

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
rok_berbe	2	340.7267	170.3634	0.40	0.6710	0.6661	0.6672
rok_berbe*sorta	6	200.4910	33.4152	0.08	0.9982	0.9979	0.9980
Error(rok_berbe)	676	288472.9193	426.7351				

Greenhouse-Geisser Epsilon	0.9764
Huynh-Feldt-Lecoutre Epsilon	0.9820

U tablici unutar subjekata testira se postoji li statistički značajna razlika za tvrdoću ploda između rokova berbe. Nadalje, testira se ponaša li se tvrdoća približno jednako po sortama za različite rokove berbe. Budući da uvjet sferičnosti nije zadovoljen, pri donošenju zaključaka gledamo Greenhouse-Geisserovu prilagodbu p-vrijednosti. Iz rezultata u tablici 2.22 vidimo da su obje prilagođene p-vrijednosti veće od 0.05 pa zaključujemo da ne postoji statistički značajna razlika između rokova berbe te da interakcija "rok_berbe*sorta" nije statistički značajna što znači da se tvrdoća ploda ponaša približno jednako po rokovima berbe.



Slika 2.11: Grafički prikaz aritmetičkih sredina i 95% pouzdanih intervala za tvrdoću ploda po sortama u sva tri roka berbe (Ispis iz SAS-a)

Na slici 2.11 prikazane su vrijednosti aritmetičkih sredina i pripadajućih 95% pouzdanih intervala za aritmetičku sredinu za tvrdoću ploda obzirom na sortu i rok berbe. Iz grafičkog prikaza vidimo da se vrijednosti tvrdoće za sve sorte smanjuju od prvog do trećeg roka berbe, a najveći pad u vrijednosti tvrdoće vidljiv je za sortu leccino. Međutim, iz rezultata iz tablice 2.22 zaključujemo da taj pad u vrijednosti tvrdoće tijekom rokova berbe nije statistički značajan, odnosno da ne postoji statistički značajna razlika između rokova berbe.

2.2.3 Suha tvar ploda

Promotrimo najprije rezultate jednofaktorskih analiza varijance za suhu tvar ploda po rokovima berbe u ovisnosti o sorti i rezultate Tukey-evog *post hoc* testa za svaki od tri roka berbe.

Tablica 2.23: Rezultati jednofaktorske analize varijance za suhu tvar ploda u prvom roku berbe (Ispis iz SAS-a)

Dependent Variable: suha_tvar_1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	148.4782867	49.4927622	6.61	0.0097
Error	10	74.9091907	7.4909191		
Corrected Total	13	223.3874774			

R-Square	Coeff Var	Root MSE	suha_tvar_1 Mean
0.664667	4.244214	2.736954	64.48671

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	148.4782867	49.4927622	6.61	0.0097

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	148.4782867	49.4927622	6.61	0.0097

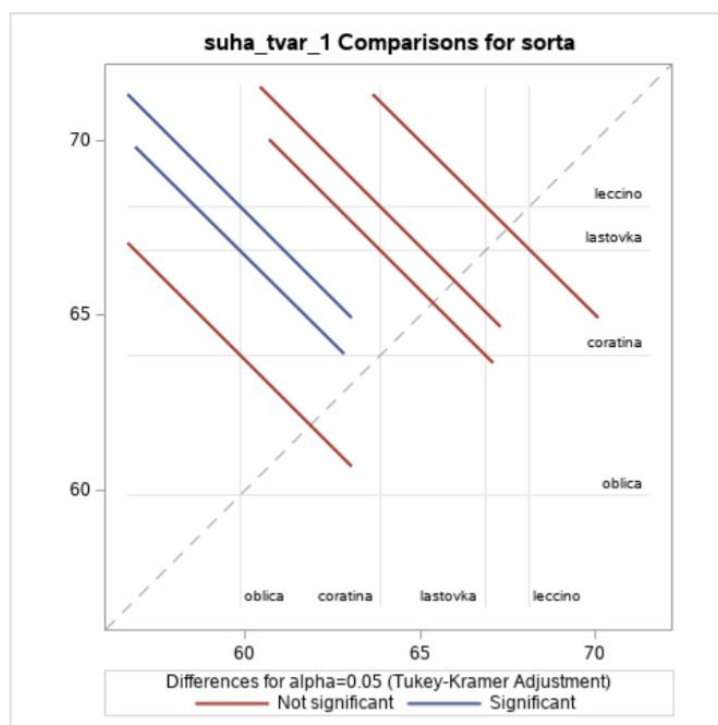
U tablici 2.23 prikazan je ispis rezultata jednofaktorske analize varijance u prvom roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Vidimo da postoji statistički značajna razlika među sortama za suhu tvar ploda u prvom roku berbe (p-vrijednost je 0.0097). Koeficijent determinacije iznosi 0.664667 pa zaključujemo da je oko 66% varijabilnosti zavisne varijable objašnjeno ovim modelom.

U tablici 2.24 nalaze se rezultati Tukey-evog *post hoc* testa za prvi rok berbe. Vidimo da statistički značajnu rezliku među sortama čini oblica obzirom na sorte lastovka i leccino.

Tablica 2.24: Rezultati Tukey-evog *post hoc* testa za suhu tvar ploda u prvom roku berbe (Ispis iz SAS-a)

sorta	suha_tvar_1 LSMEAN	LSMEAN Number
coratina	63.8736289	1
lastovka	66.8521535	2
leccino	68.1118787	3
oblica	59.8622142	4

Least Squares Means for effect sorta Pr > t for H0: LSmean(i)=LSmean(j) Dependent Variable: suha_tvar_1				
i/j	1	2	3	4
1		0.5131	0.2887	0.2801
2	0.5131		0.9289	0.0208
3	0.2887	0.9289		0.0123
4	0.2801	0.0208	0.0123	



Slika 2.12: Grafički prikaz rezultata Tukey-evog *post hoc* testa za suhu tvar ploda u prvom roku berbe (Ispis iz SAS-a)

Tablica 2.25: Rezultati jednofaktorske analize varijance za suhu tvar ploda u drugom roku berbe (Ispis iz SAS-a)

Dependent Variable: suha_tvar_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	870.7360840	290.2453613	161.84	<.0001
Error	10	17.9336360	1.7933636		
Corrected Total	13	888.6697200			

R-Square	Coeff Var	Root MSE	suha_tvar_2 Mean
0.979820	2.299445	1.339165	58.23863

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	870.7360840	290.2453613	161.84	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	870.7360840	290.2453613	161.84	<.0001

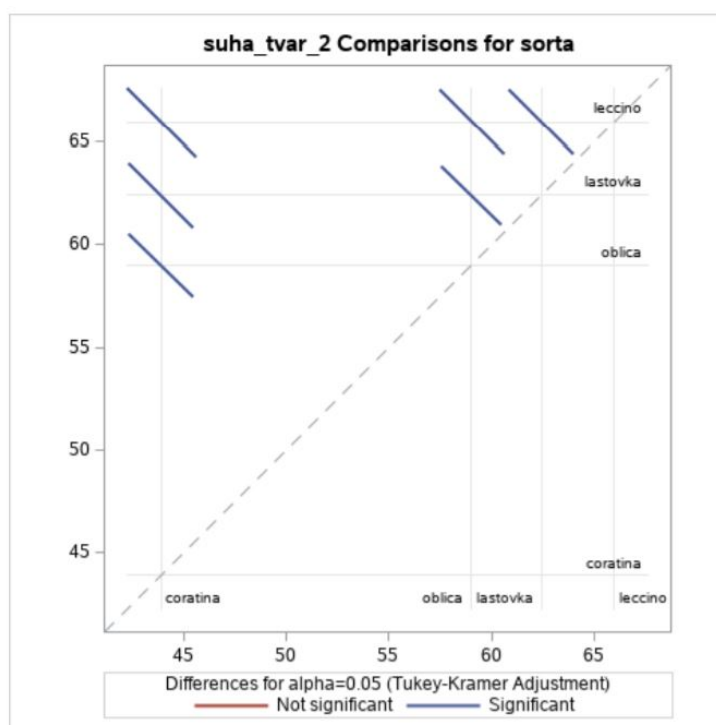
U tablici 2.25 prikazan je ispis rezultata jednofaktorske analize varijance u drugom roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Vidimo da postoji statistički značajna razlika među sortama za suhu tvar ploda u drugom roku berbe (p-vrijednost je manja od 0.0001). Primijetimo da je u usporedbi s prvim rokom berbe, iznos F vrijednosti u drugom roku berbe zamjetno veći (u prvom roku berbe F vrijednost je 6.61, a u drugom 161.84) što nam govori da je razlika među sortama veća u drugom roku berbe. Koeficijent determinacije iznosi 0.97982 pa zaključujemo da je gotovo 98% varijabilnosti zavisne varijable objašnjeno ovim modelom. Dakle, ovaj model jako dobro objašnjava varijabilnost zavisne varijable.

U tablici 2.26 nalaze se rezultati Tukey-evog *post hoc* testa za drugi rok berbe. Vidimo da je razlika između svakog para sorti statistički značajna.

Tablica 2.26: Rezultati Tukey-evog *post hoc* testa za suhu tvar ploda u drugom roku berbe (Ispis iz SAS-a)

sorta	suha_tvar_2 LSMEAN	LSMEAN Number
coratina	43.8814718	1
lastovka	62.4262146	2
leccino	65.9805812	3
oblica	59.0124555	4

Least Squares Means for effect sorta Pr > t for H0: LSmean(i)=LSmean(j)				
Dependent Variable: suha_tvar_2				
i/j	1	2	3	4
1		<.0001	<.0001	<.0001
2	<.0001		0.0258	0.0210
3	<.0001	0.0258		0.0002
4	<.0001	0.0210	0.0002	



Slika 2.13: Grafički prikaz rezultata Tukey-evog *post hoc* testa za suhu tvar ploda u drugom roku berbe (Ispis iz SAS-a)

Tablica 2.27: Rezultati jednofaktorske analize varijance za suhu tvar ploda u trećem roku berbe (Ispis iz SAS-a)

Dependent Variable: suha_tvar_3

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	346.9743142	115.6581047	26.79	<.0001
Error	10	43.1753546	4.3175355		
Corrected Total	13	390.1496689			

R-Square	Coeff Var	Root MSE	suha_tvar_3 Mean
0.889336	3.390539	2.077868	61.28430

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sorta	3	346.9743142	115.6581047	26.79	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	346.9743142	115.6581047	26.79	<.0001

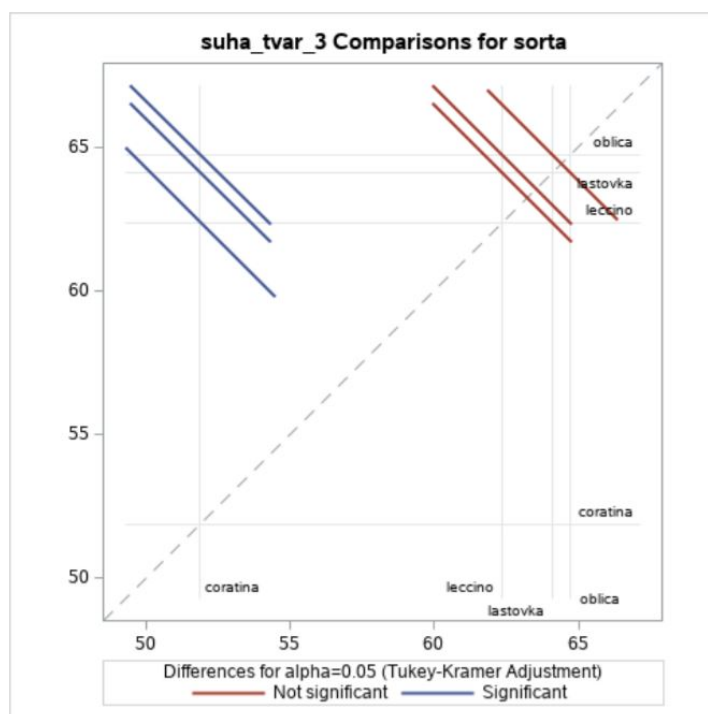
U tablici 2.27 prikazan je ispis rezultata jednofaktorske analize varijance u trećem roku berbe te vrijednosti ostalih dijagnostičkih informacija o modelu jednofaktorske analize varijance. Vidimo da postoji statistički značajna razlika među sortama za suhu tvar ploda u trećem roku berbe (p-vrijednost je manja od 0.0001). F vrijednost iznosi 26.79 što je manje u odnosu na vrijednost iz drugog roka berbe pa zaključujemo da je u trećem roku berbe manja razlika između sorti u usporedbi s drugim rokom berbe. Koeficijent determinacije iznosi 0.889336 pa zaključujemo da je gotovo 90% varijabilnosti zavisne varijable objašnjeno ovim modelom. Dakle, ovaj model jako dobro objašnjava varijabilnost zavisne varijable.

U tablici 2.28 nalaze se rezultati Tukey-evog *post hoc* testa za treći rok berbe. Vidimo da statistički značajnu razliku među sortama čini coratina obzirom na preostale tri promatrane sorte.

Tablica 2.28: Rezultati Tukey-evog *post hoc* testa za suhu tvar ploda u trećem roku berbe (Ispis iz SAS-a)

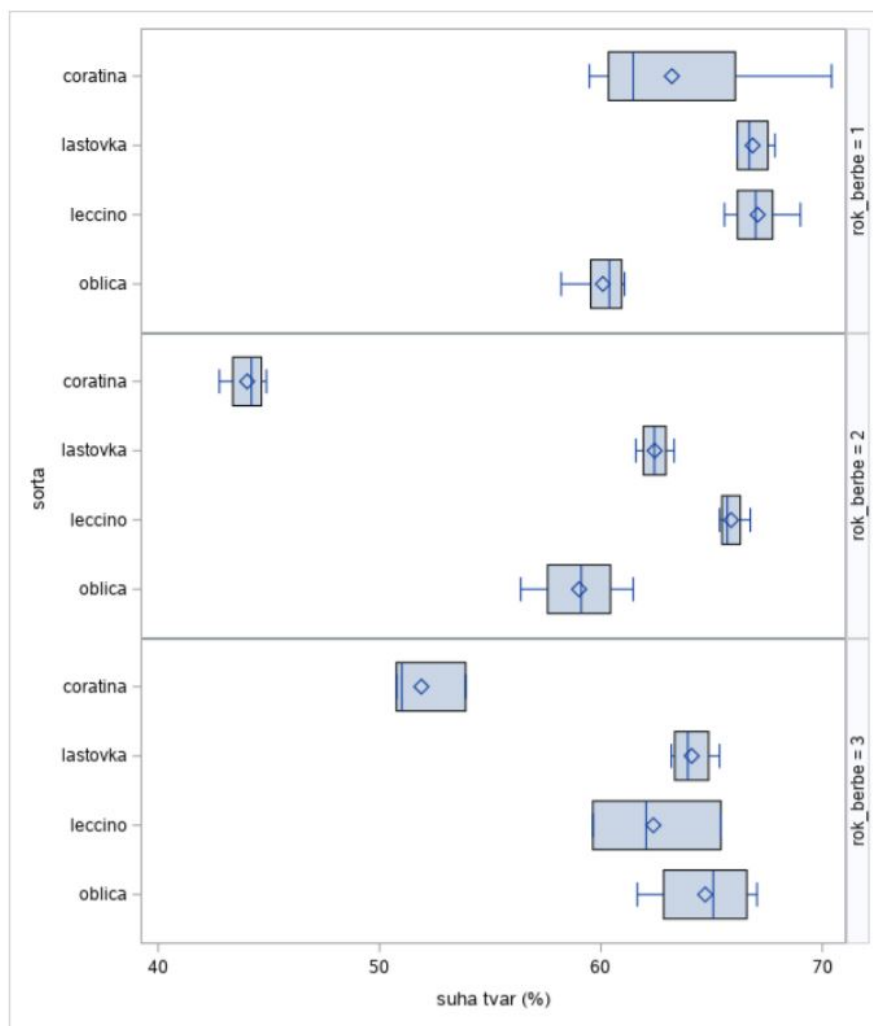
sorta	suha_tvar_3 LSMEAN	LSMEAN Number
coratina	51.8872452	1
lastovka	64.0958252	2
leccino	62.3653260	3
oblica	64.7097871	4

Least Squares Means for effect sorta Pr > t for H0: LSmean(i)=LSmean(j) Dependent Variable: suha_tvar_3				
i/j	1	2	3	4
1		<.0001	0.0005	<.0001
2	<.0001		0.7029	0.9741
3	0.0005	0.7029		0.4846
4	<.0001	0.9741	0.4846	



Slika 2.14: Grafički prikaz rezultata Tukey-evog *post hoc* testa za suhu tvar ploda u trećem roku berbe (Ispis iz SAS-a)

Na slici 2.15 prikazan je *boxplot* za suhu tvar ploda s podjelom po rokovima berbe i sortama. Možemo zamijetiti veće razlike u vrijednostima među sortama i po rokovima berbe. Primijetimo da je za sortu coratina u prvom roku berbe veći raspon vrijednosti, zatim se u drugom roku berbe raspon smanjuje kao i same vrijednosti suhe tvari ploda, a u trećem roku berbe vrijednosti su ponovno veće.



Slika 2.15: *Boxplot* prikaz suhe tvari ploda s podjelom po rokovima berbe i sortama (Ispis iz SAS-a). Sivi pravokutnik označava dio između donjeg (q_L) i gornjeg kvartila (q_U), plava okomita linija unutar pravokutnika medijan, a plavi kvadratić označava aritmetičku sredinu. *Outlieri* su označeni kružićima, a okomite plave linije van pravokutnika označavaju vrijednosti koje su udaljene najviše $1.5(q_U - q_L)$ od donjeg, odnosno gornjeg kvartila.

Tablica 2.29: Test sferičnosti za suhu tvar ploda (Ispis iz SAS-a)

Sphericity Tests				
Variables	DF	Mauchly's Criterion	Chi-Square	Pr > ChiSq
Transformed Variates	2	0.7692732	2.3607825	0.3072
Orthogonal Components	2	0.999302	0.0062843	0.9969

U tablici 2.29 dani su rezultati testa sferičnosti. Iz rezultata zaključujemo da je uvjet sferičnosti zadovoljen (p-vrijednost iznosi 0.9969).

Tablica 2.30: Tablica rezultata ANOVA-e ponovljenih mjerenja između subjekata za suhu tvar ploda (Ispis iz SAS-a)

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sorta	3	865.8888797	288.6296266	51.98	<.0001
Error	10	55.5240587	5.5524059		

U tablici između subjekata testiramo postoji li statistički značajna razlika u suhoj tvari ploda među sortama. Iz rezultata u tablici 2.12 vidimo da postoji statistički značajna razlika između sorti (p-vrijednost je manja od 0.0001). Dakle, uprosječimo li podatke po rokovima berbe za svaku sortu, vidimo da postoji statistički značajna razlika među sortama.

Tablica 2.31: Tablica rezultata ANOVA-e ponovljenih mjerenja unutar subjekata za suhu tvar ploda (Ispis iz SAS-a)

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
rok_berbe	2	323.8896941	161.9448470	40.24	<.0001	<.0001	<.0001
rok_berbe*sorta	6	500.2998052	83.3833009	20.72	<.0001	<.0001	<.0001
Error(rok_berbe)	20	80.4941227	4.0247061				

Greenhouse-Geisser Epsilon	0.9993
Huynh-Feldt-Lecoutre Epsilon	1.2488

U tablici unutar subjekata testira se postoji li statistički značajna razlika za suhu tvar ploda između rokova berbe. Nadalje, testira se ponaša li se suha tvar jednako po sortama za različite rokove berbe. Iz rezultata u tablici 2.31 vidimo da su obje p-vrijednosti manje od 0.0001 pa zaključujemo da postoji statistički značajna razlika između rokova berbe te da je interakcija "rok_berbe*sorta" statistički značajna. Budući da je interakcija statistički značajna, zaključak o tome razlikuje li se suha tvar ploda po sortama različit je u ovisnosti o tome koji rok berbe promatramo.

Tablica 2.32: Analiza varijance kontrastnih varijabli (Ispis iz SAS-a)

Repeated Measures Analysis of Variance
Analysis of Variance of Contrast Variables

rok_berbe_N represents the contrast between the nth level of rok_berbe and the 1st

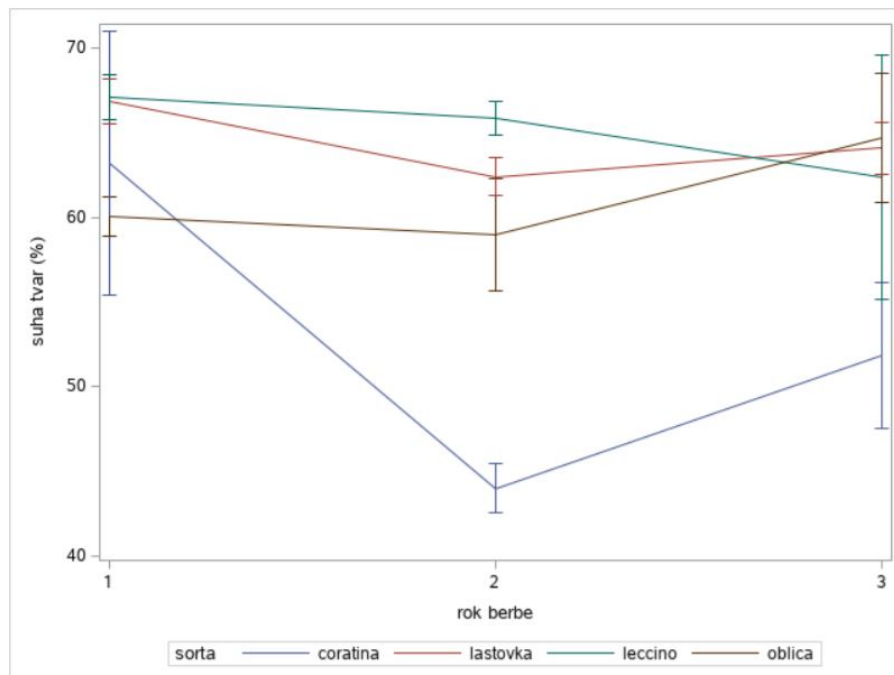
Contrast Variable: rok_berbe_2

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Mean	1	643.4687443	643.4687443	77.92	<.0001
sorta	3	747.3909444	249.1303148	30.17	<.0001
Error	10	82.5790790	8.2579079		

Contrast Variable: rok_berbe_3

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Mean	1	209.7107320	209.7107320	26.27	0.0004
sorta	3	510.8974256	170.2991419	21.34	0.0001
Error	10	79.8145256	7.9814526		

U tablici unutar subjekata dobili smo potvrđan odgovor na pitanje postoji li statistički značajna razlika za suhu tvar ploda među rokovima berbe. Sada ćemo usporediti suhu tvar ploda u drugom i trećem roku berbe s prvim rokom berbe. U tablici 2.32 testiramo postoji li statistički značajna razlika za suhu tvar ploda između prvog i drugog roka berbe te između prvog i trećeg roka berbe. Također, testiramo postoji li statistički značajna razlika između sorti u prvom i drugom, odnosno prvom i trećem roku berbe. Prema rezultatima u tablici 2.32 zaključujemo da postoji statistički značajna razlika za suhu tvar ploda između prvog i drugog roka berbe te da postoji statistički značajna razlika u suhoj tvari ploda među sortama u prvom i drugom roku berbe (p-vrijednosti su manje od 0.0001). Nadalje, isti zaključak vrijedi i prilikom usporedbe prvog i trećeg roka berbe (p-vrijednosti iznose 0.0004 i 0.0001).



Slika 2.16: Grafički prikaz aritmetičkih sredina i 95% pouzdanih intervala za suhu tvar ploda po sortama u sva tri roka berbe (Ispis iz SAS-a)

Na slici 2.16 prikazane su vrijednosti aritmetičkih sredina i pripadajućih 95% pouzdanih intervala za aritmetičku sredinu za suhu tvar ploda obzirom na sortu i rok berbe. Iz grafičkog prikaza možemo iščitati da se vrijednosti za sortu coratina u drugom i trećem roku berbe razlikuju od vrijednosti za preostale sorte te je za sortu coratina primjetna razlika u vrijednostima za različite rokove berbe.

2.3 Zaključak

Sumirajmo na kraju rezultate za sve tri zavisne varijable koje smo analizirali.

Što se tiče elastičnosti ploda, postoji statistički značajna razlika među sortama u svakom od tri roka berbe. U prvom i trećem roku berbe statistički značajnu razliku među sortama čini lastovka obzirom na preostale tri promatrane sorte, dok u drugom roku berbe statistički značajnu razliku među sortama čini lastovka obzirom na oblicu i leccino. Nadalje, postoji statistički značajna razlika u elastičnosti ploda među sortama, a ne postoji statistički značajna razlika u elastičnosti ploda između rokova berbe te interakcija "rok_berbe*sorta" nije statistički značajna.

Kod tvrdoće ploda također postoji statistički značajna razlika među sortama u svakom od tri roka berbe. U prvom i trećem roku berbe statistički značajnu razliku među sortama čini sorta leccino obzirom na lastovku i oblicu, dok u drugom roku berbe statistički značajnu razliku među sortama čini leccino obzirom na preostale tri promatrane sorte. Postoji statistički značajna razlika u tvrdoći ploda među sortama, a ne postoji statistički značajna razlika u tvrdoći ploda između rokova berbe te interakcija "rok_berbe*sorta" nije statistički značajna.

Posljednja analizirana osobina ploda masline bila je suha tvar ploda. Zaključili smo da postoji statistički značajna razlika među sortama u svakom od tri roka berbe. U prvom roku berbe statistički značajnu razliku među sortama čini oblica obzirom na lastovku i leccino. U drugom roku berbe razlika između svakog para sorti statistički je značajna, dok u trećem roku berbe statistički značajnu razliku među sortama čini coratina obzirom na preostale tri promatrane sorte. Nadalje, postoji statistički značajna razlika u suhoj tvari ploda među sortama te postoji statistički značajna razlika između rokova berbe i interakcija "rok_berbe*sorta" statistički je značajna.

Poglavlje 3

Dodatak

Kod u SAS-u

Podaci su na početku učitani u tablice podaci_st i podaci_tv_el koje se sastoje od stupaca sorta, rok_berbe, stablo (oznaka stabla s kojega je ubran plod) te oznake zavisne varijable (suha_tvar, elasticnost, tvrdoca). Nakon toga podaci su sortirani po sorti i oznaci stabla te restrukturirani kako bismo ih pripremili za *glm* proceduru.

```
/*restrukturiranje podataka*/
data podaci_el_glm(keep=sorta stablo elasticnost_1-elasticnost_3);
    array el(3) elasticnost_1-elasticnost_3;
    do Time=1 to 3;
        set podaci_tv_el;
        by sorta stablo;
        el(Time)=elasticnost;
        if last.stablo then return;
    end;
run;

data podaci_tv_glm(keep=sorta stablo tvrdoca_1-tvrdoca_3);
    array tv(3) tvrdoca_1-tvrdoca_3;
    do Time=1 to 3;
        set podaci_tv_el;
        by sorta stablo;
        tv(Time)=tvrdoca;
        if last.stablo then return;
    end;
run;
```

```
data podaci_st_glm(keep=sorta stablo suha_tvar_1-suha_tvar_3);
  array st(3) suha_tvar_1-suha_tvar_3;
  do Time=1 to 3;
    set podaci_st;
    by sorta stablo;
    st(Time)=suha_tvar;
    if last.stablo then return;
  end;
run;
```

```
/*deskriptivna statistika*/
```

```
proc means data=PODACI_TV_GLM chartype n mean std min max median
  vardef=df qmethod=os nonobs;
  var tvrdoca_1 tvrdoca_2 tvrdoca_3;
  class sorta;
run;
```

```
proc means data=PODACI_ST_GLM chartype n mean std min max median
  vardef=df qmethod=os nonobs;
  var suha_tvar_1 suha_tvar_2 suha_tvar_3;
  class sorta;
run;
```

```
proc means data=PODACI_EL_GLM chartype n mean std min max median
  vardef=df qmethod=os nonobs;
  var elasticnost_1 elasticnost_2 elasticnost_3;
  class sorta;
run;
```

```
/*grafički prikaz*/
```

```
proc sgplot data=podaci_st;
  vline rok_berbe / response=suha_tvar group=sorta stat=mean
    limitstat=CLM;
  yaxis label='suha tvar (%)';
  xaxis label='rok berbe';
run;
```



```
proc sgplot data=podaci_tv_el;
    vline rok_berbe / response=tvrdoca group=sorta stat=mean
        limitstat=CLM;
    yaxis label='tvrdoća (N)';
    xaxis label='rok berbe';
run;

proc sgplot data=podaci_tv_el;
    vline rok_berbe / response=elasticnost group=sorta stat=mean
        limitstat=CLM;
    yaxis label='elastičnost (mm)';
    xaxis label='rok berbe';
run;

/*boxplot-ovi*/
ods graphics / reset width=600px height=700px;

proc sgpanel data=PODACI_TV_EL;
    panelby rok_berbe / columns=1 layout=rowlattice;
    hbox elasticnost / category=sorta boxwidth=0.7;
    colaxis label='elastičnost (mm)';
run;

proc sgpanel data=PODACI_TV_EL;
    panelby rok_berbe / columns=1 layout=rowlattice;
    hbox tvrdoca / category=sorta boxwidth=0.7;
    colaxis label='tvrdoća (N)';
run;

proc sgpanel data=PODACI_ST;
    panelby rok_berbe / columns=1 layout=rowlattice;
    hbox suha_tvar / category=sorta boxwidth=0.7;
    colaxis label='suha tvar (%)';
run;
```

```
/*anova ponovljenih mjerenja*/
proc glm data=podaci_el_glm;
    class sorta;
    model elasticnost_1 - elasticnost_3=sorta;
    repeated rok_berbe 3 / printe;
    lsmeans sorta / pdiff adjust=tukey;
    means sorta;
run;

proc glm data=podaci_tv_glm;
    class sorta;
    model tvrdoca_1 - tvrdoca_3=sorta;
    repeated rok_berbe 3 / printe;
    lsmeans sorta / pdiff adjust=tukey;
    means sorta;
run;

proc glm data=podaci_st_glm;
    class sorta;
    model suha_tvar_1 - suha_tvar_3=sorta;
    repeated rok_berbe 3 contrast(1) / summary printm;
    lsmeans sorta / pdiff adjust=tukey;
    means sorta;
run;
```

Bibliografija

- [1] https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.4/statug/statug_glm_toc.htm.
- [2] Davis, C.S.: *Statistical methods for the analysis of repeated measurements*. Springer, New York, 2002.
- [3] H. Huynh, L.S. Feldt: *Estimation of the Box Correction for Degrees of Freedom from Sample Data in Randomized Block and Split-Plot Designs*. *Journal of Educational Statistics*, 1:69 – 82, 1976.
- [4] Jazbec, A.: *Odabrane statističke metode u biomedicini*. PMF-MO, nastavni materijali, 2020.
- [5] Molnar, N.: *Višefaktorska analiza varijance ponovljenih mjerenja*. diplomski rad, 2016.
- [6] Pažin, I.: *Analiza varijance ponovljenih mjerenja*. diplomski rad, 2014.
- [7] S.W. Greenhouse, S. Geisser: *On methods in the analysis of profile data*. *Psychometrika*, 24:95–112, 1959.

Sažetak

U ovom radu analizirali smo reološke i kemijske osobine ploda masline u ovisnosti o sorti i zrelosti ploda analizom varijance ponovljenih mjerenja s interakcijom između sorte i zrelosti ploda. Analizirali smo elastičnost i tvrdoću kao reološka svojstva te suhu tvar kao kemijsko svojstvo ploda. Promatrali smo četiri sorte: oblica, leccino, lastovka i coratina te tri roka berbe koji su određeni prema obojenosti ploda masline tijekom dozrijevanja. Podaci koje smo obrađivali prikupljeni su za projekt „Analiza teksturnog profila kao čimbenik preradbenih značajki i vrednovanja bioraznolikosti autohtonih sorti maslina” Instituta za Jadranske kulture i melioraciju krša u Splitu tijekom rujna i listopada 2019. godine u masliniku OPG-a Jagoda Alfirević u Kaštel Sućurcu.

Naglasak rada bio je na praktičnoj primjeni analize varijance ponovljenih mjerenja te na interpretaciji rezultata dobivenih korištenjem *glm* procedure u statističkom programu SAS. Prije praktične primjene, u prvom poglavlju objašnjeni su statistički pojmovi koje smo koristili u nastavku rada.

Za sve tri zavisne varijable (elastičnost, tvrdoća i suha tvar ploda) postoji statistički značajna razlika među sortama u svakom od tri roka berbe. Kod elastičnosti ploda, u prvom i trećem roku berbe statistički značajnu razliku među sortama čini lastovka obzirom na preostale tri promatrane sorte, dok u drugom roku berbe statistički značajnu razliku među sortama čini lastovka obzirom na oblicu i leccino. Što se tiče tvrdoće ploda, u prvom i trećem roku berbe statistički značajnu razliku među sortama čini sorta leccino obzirom na lastovku i oblicu, dok u drugom roku berbe statistički značajnu razliku među sortama čini leccino obzirom na preostale tri promatrane sorte. Kod suhe tvari ploda, u prvom roku berbe statistički značajnu razliku među sortama čini oblica obzirom na lastovku i leccino, u drugom roku berbe razlika između svakog para sorti statistički je značajna, dok u trećem roku berbe statistički značajnu razliku među sortama čini coratina obzirom na preostale tri promatrane sorte. Nadalje, postoji statistički značajna razlika u elastičnosti i tvrdoći ploda među sortama, a ne postoji statistički značajna razlika između rokova berbe te interakcija "rok_berbe*sorta" nije statistički značajna. Na kraju, zaključili smo da postoji statistički značajna razlika u suhoj tvari ploda među sortama te da postoji statistički značajna razlika između rokova berbe i interakcija "rok_berbe*sorta" statistički je značajna.

Summary

In this paper we have analysed the rheological and chemical properties of olive fruit depending on the cultivar and fruit maturity by using the repeated measures analysis of variance with the interaction between cultivar and fruit maturity. Springiness and hardness as rheological properties as well as dry matter as a chemical property of the fruit were examined. We observed four cultivars: 'Oblica', 'Leccino', 'Lastovka' and 'Coratina', and three harvest periods that were determined by the colour of the olive fruit during ripening. The data we processed were collected as a part of the project "Analysis of textural profile as a factor in processing characteristics and evaluation of biodiversity of autochthonous olive cultivars" conducted by the Institute for Adriatic Crops and Karst Reclamation in Split during September and October 2019 in the olive orchard Jagoda Alfirević in Kaštel Sućurac.

The emphasis of the paper was on the practical application of the repeated measures analysis of variance and the interpretation of the results obtained using the *glm* procedure in SAS. Before practical application, the first chapter explains the statistical concepts that we used in the following chapter.

For all three dependant variables (fruit springiness, hardness and dry matter) there is a statistically significant difference between cultivars in each of the three harvest periods. In terms of fruit springiness, in the first and third harvest period a statistically significant difference between cultivars is found at 'Lastovka' with regard to the remaining three observed cultivars, while in the second harvest period a statistically significant difference between cultivars is detected at 'Lastovka' with regard to 'Oblica' and 'Leccino'. As for fruit hardness, in the first and third harvest period a statistically significant difference between cultivars is detected at the cultivar 'Leccino' with regard to 'Lastovka' and 'Oblica', while in the second harvest period a statistically significant difference between cultivars is found at 'Leccino' with regard to the remaining three observed cultivars. In the case of fruit dry matter, in the first harvest period a statistically significant difference between cultivars is made by 'Oblica' with regard to 'Lastovka' and 'Leccino', in the second harvest period there is a statistically significant difference between every pair of cultivars, while in the third harvest period there is a statistically significant difference between 'Coratina' and remaining three observed cultivars. Furthermore, there is a statistically significant difference

in fruit springiness and hardness between cultivars, however there is no statistically significant difference between harvest periods and the interaction "harvest_period*cultivar" is not statistically significant. Finally, it was concluded that both difference in fruit dry matter between cultivars and a difference in harvest periods were found to be statistically significant, as well as the interaction "harvest_period*cultivar".

Životopis

Rođena sam 4. travnja 1997. godine u Zagrebu. Pohađala sam Osnovnu školu Vjenceslava Novaka u Zagrebu, a nakon toga II. gimnaziju u Zagrebu. Tijekom osnovne i srednje škole aktivno sam se bavila odbojkom. Nakon završetka srednje škole, 2016. godine upisala sam preddiplomski studij Matematike na Prirodoslovno-matematičkom fakultetu u Zagrebu. Preddiplomski studij završila sam 2019. godine te iste godine upisala diplomski studij Matematičke statistike na istom fakultetu.