Poglavje 1

POSTAVITEV IN TESTIRANJE HIPOTEZ

Testiranje hipotez je osrednja naloga pri vsaki obdelavi podatkov. Od postavitve hipotez je odvisen načrt preizkusa, torej moramo hipoteze postaviti še pred izpeljavo poskusa. Po izvedbi poskusa je včasih potrebno stvari celo popraviti ali prilagoditi, saj se kaj rado zgodi, da pri poskusu poteka kakšna reč drugače, kot smo predvideli.

Preizkus hipotez opravimo v treh korakih:

1. Preizkusimo, ali je model značilen.
2. Preizkusimo, kateri vplivi v modelu so značilni in kateri niso.
3. Preizkusimo, kateri nivoji pri značilnih vplivih se med seboj razlikujejo.

Nikoli ne preizkušamo razlik med nivojema dveh različnih vplivov, izogibamo se tudi kombiniranim razlikam. Hipoteze naredimo čimbolj enostavne, da jih je tudi enostavno razložiti.

1.1 Postavitev hipoteze

Pri študiju se poslužite prosojnic, kjer imate obrazložene hipoteze v skalarni obliki. Z matrično obliko se bomo seznanili samo toliko, da lahko povežete skalarni zapis hipotez z zapisom v programskem paketu SAS.

1.1.1 Ničelna in alternativna hipoteza

Hipoteza ima dve komponenti: ničelno hipotezo H_0 (1.1) in alternativno hipotezo H_1 (1.2). Ničelna hipoteza ima lahko dve obliki. Prva oblika pomeni, da se linearne kombinacije **K** (lokacijskih) parametrov $\boldsymbol{\beta}$ ne razlikujejo od vektorja **0**, v drugem primeru pa pričakujemo pri rezultatu linearnih kombinacij konstantno vrednost v vektorju **m**. Prvi primer je zelo običajen, saj najprej preverjamo ali so dobljeni rezultati od 0 različne.

$$H_0 : \mathbf{K}\boldsymbol{\beta} = \mathbf{0} \qquad [1.1]$$

$$H_0 : \mathbf{K}\boldsymbol{\beta} = \mathbf{m} \qquad [1.2]$$

Alternativna hipoteza (H_1 ali tudi H_a) lahko zavzema vse druge možnosti ali pa samo del. Zelo pomembno je, da alternativno hipotezo nazorno nakažemo. Hipotezi v naslednjih vrsticah vključujeta vse alternative ničelni hipotezi. Pri prvi hipotezi 1.3, ki je alternativa ničelni hipotezi 1.1, ovržemo ničelno hipotezo, če da katerakoli linearna kombinacija iz matrike **K** rezultat različen od 0. Druga hipoteza v 1.4 je alternativa ničelni hipotezi v enačbi 1.2. Alternativno hipotezo sprejmemo, če je vrednost najmanj ene linearne kombinacije iz matrike **K** različna od vrednosti v vektorju **m**. Ne moremo pa kombinirati niti ničelno hipotezo v 1.1 z alternativno hipotezo v 1.4 niti ničelno hipotezo v 1.2 z alternativno hipotezo v 1.3.

$$H_1 : \mathbf{K}\boldsymbol{\beta} \neq \mathbf{0} \qquad [1.3]$$

$$H_1 : \mathbf{K}\boldsymbol{\beta} \neq \mathbf{m} \quad [1.4]$$

Če npr. ničelna hipoteza pokriva možnost, da med pasmami ni razlik, oziroma bolj dosledno, da so razlike med pasmami enake nič, alternativna hipoteza predstavlja vse možnosti, ko med pasmami obstajajo razlike. Že ob eni sami od nič različni razliki bo ničelna hipoteza zavržena in sprejeta alternativna hipoteza. V primeru, da drži ničelna drži, nobena od razlik ni dokazano od nič različna. Vsako nadaljnje razglabljanje in iskanje razlik je neprimerno. Le v primeru, da so razvidni kakšni trendi, lahko predlagamo, da se poskus dopolni s potrebnimi meritvami ali pa ponovno zastavi s primernimi popravki (velikost vzorca, način vzorčenja itd.), da bi dobili potrditev ali zavrnitev nakazanega trenda.

Alternativna hipoteza pa lahko vključuje samo del alternativnih možnosti. Najpogostejši obliki sta v tem primeru hipotezi, ki vključujeta samo tiste možnosti, ko so ocene linearnih kombinacij večje od **0** (1.5), manjše od **0** (1.7), večje od konstant v vektorju **m** (1.6) ali manjše od konstant v vektorju **m** (1.8). Alternativni hipotezi v 1.5 in 1.7 lahko kombiniramo z ničelno hipotezo 1.1, ostali dve (1.6 in 1.8) pa z 1.2. Drugih možnosti ni.

$$H_1 : \mathbf{K}\boldsymbol{\beta} > \mathbf{0} \quad [1.5]$$

$$H_1 : \mathbf{K}\boldsymbol{\beta} > \mathbf{m} \quad [1.6]$$

$$H_1 : \mathbf{K}\boldsymbol{\beta} < \mathbf{0} \quad [1.7]$$

$$H_1 : \mathbf{K}\boldsymbol{\beta} < \mathbf{m} \quad [1.8]$$

Za ponazoritev moramo poiskati nov primer, nesmiselno bi bilo primer razlagati na primeru pasem. V selekcijkskem programu predvidevamo, da bo selekcijsko delo prineslo načrtovan genetski napredek. Čez leta lahko genetski napredek preverimo. Ker je bilo vloženega dela in kapitala mnogo, se ne moremo zadovoljiti z genetskim trendom, ki bi bil samo različen od nič. Negativni trendi, vrednosti manjše od nič, so še slabše, kot če bi genetskega napredka sploh ne gi bilo (genetski trend enak nič). Torej nas zanima le genetski napredek (trend) z ocenami, ki so večje od nič. Še bolj pogosto pa se odločamo v tem primeru za hipotezo, ki v vektorju **m** hrani načrtovane, planirane genetske spremembe.

Vlogi hipotez pa sta v praksi nekoliko drugačni kot v statistični teoriji. V praksi praviloma želimo dokazati, da obstajajo razlike, da obstajajo trendi ali povezave med spremenljivkami. Tako bi nam bilo skoraj bolj razumljivo, da poskušamo postaviti to kot ničelno, izhodiščno hipotezo. V statistiki pa vedno najprej ovržemo možnost, da razlik ni oziroma niso dovolj velike. Šele nato iščemo, kateri nivoji se med seboj razlikujejo. Če smo dobili odgovor, da nivoji niso različni od nič, je vsako nadaljne iskanje samo izguba časa. Nobena razlika ni značilna. Po domače bi rekli "ni dovolj pomembna" ali "ni dovolj prepričljiva". Paziti moramo, ker nam lahko napačno izbrani testi razliko pokažejo, čeprav so nam z njimi na krožniku postregli najboljši statistični paketi. Naloga statističnih paketov je, da uporabniku olajšajo delo tako, da jim ni potrebno poznati vseh številnih formul. Tudi mi bomo lahko po izpitu kakšno pozabili. Ne morejo pa pomagati pri izboru orodij, med njimi tudi pravilnih statističnih testov. Tako kot moramo na kmetiji vedeti, s katerimi stroji bomo pomolzli krave in s čim bomo orali njivo, moramo vedeti, katera so najprimernejša orodja za obdelavo podatkov, ki jih v živinoreji zbiramo. O izboru metod za obdelavo podatkov smo raypravljali v predhodnem poglavju.

1.1.2 Postavitev linearnih kombinacij

Hipoteze lahko predstavimo oziroma oblikujemo v matrični obliki. Z matrično obliko lahko nazorno ponazorimo posamezne hipoteze. Kot smo prikazali v skalarni obliki, so hipoteze pogosto enostavne. Takšne je tudi laže razložiti. Če pa je struktura podatkov nekoliko bolj zapletena (manjkajoči podatki, interakcije...), je lahko hipoteza tudi bolj sestavljena.

Našo hipotezo predstavimo v matriki linearnih kombinacij parametrov **K**. Če je hipoteza ocenljiva, potem bo produkt $\mathbf{K}\boldsymbol{\beta}$ vedno enak, ne glede na to katero izmed neskončno velikega števila možnih rešitev smo izbrali. Za matriko hipotez je pomembno, da ni v njej linearno odvisnih hipotez. Te dodatne hipoteze ne prinesejo novih spoznanj, ampak samo prikažejo rezultate v drugi luči.

PRIMER:

Vzemimo npr. primer mladic iz preizkusa v proizvodnih razmerah. Proučujemo le vpliva pasme (P_i) in farme $\left(F_j\right)$ s po tremi nivoji. Ocene parametrov za sistematske vplive so nanizane v vektorju $\widehat{\boldsymbol{\beta}}'$ (enačba 1.9). Pri tem ne smemo pozabiti na srednjo vrednost (μ).

$$\widehat{\boldsymbol{\beta}}' = \left[\begin{array}{ccccccc} \widehat{\mu} & \widehat{P}_1 & \widehat{P}_2 & \widehat{P}_3 & \widehat{F}_1 & \widehat{F}_2 & \widehat{F}_3 \end{array} \right] \qquad [1.9]$$

Zanimajo nas razlike med pasmami. Imamo tri možne razlike (prva-druga, prva-tretja in druga-tretja). Prvi dve razliki smo vnesli v prvi dve vrstici matrike **H**. Lahko pa bi nas zanimala tudi dvakratna razlika med drugo in tretjo pasmo, kar smo ponazorili v tretji vrstici. V matriki **H** je tretja vrstica dvakratna razlika med drugo in prvo vrstico: je linearna kombinacija prvih dveh. To v praksi pomeni, da je tretja razlika logični zaključek prvih dveh. V matriki, ki jo uporabljamo pri testiranju hipotez, uporabimo katerokoli kombinacijo samo linearno neodvisnih vrstic iz matrike **H**. Matriko katerekoli teh kombinacij bomo poimenovali **K**. Imeti mora polni rang v vrsticah, po stolpcih pa ni omejitve.

$$\mathbf{H} = \left[\begin{array}{ccccccc} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 \end{array} \right] \qquad [1.10]$$

Razlike med pasmami torej testiramo z naslednjimi hipotezami. Našli bi lahko še druge možnosti. Vrednosti, ki so različne od nič, so pogosto 1 in -1, tako kot v spodnjih dveh. Tako zapišemo npr. razliko med dvema nivjema znoraj vpliva.

$$\mathbf{K} = \left[\begin{array}{ccccccc} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \end{array} \right] \qquad [1.11]$$

$$\mathbf{K} = \left[\begin{array}{ccccccc} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{array} \right] \qquad [1.12]$$

Vajo bi lahko ponovili tudi za razlike med farmami. Z linearnimi kombinacijami iz 1.13 pa si ne moremo veliko pomagati. Poskušajmo prebrati prvo vrstico. Zanima nas razlika med prvo pasmo in drugo farmo. Takšna vrednost pa živinorejca bolj malo zanima. Kaj bi se iz razlike naučil? Ali bi kupil živali prve pasme, ali pa morda farmo 2? Vsekakor takšne dileme ne obstajajo. Odločamo se med pasmami ali med farmami. Konec koncev bi se lahko zgodilo, da bi želel kupiti farmo in živali. Še vedno pa bi farmo izbiral med farmami in bi te primerjave ločno opravil. Pasme (živali) pa bi izbiral med pasmami.

$$\mathbf{K} = \left[\begin{array}{ccccccc} 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{array} \right] \qquad [1.13]$$

Če bi bili pogoji med farmami zelo različni, bi pred nakupom živali hotel preveriti, ali s pasmami dosega različne proizvodne rezultate na posameznih farmah. V tem primeru pa bi želel preveriti tudi interakcije.

$$\widehat{\boldsymbol{\beta}}' = \left[\begin{array}{ccccccc} \widehat{\mu} & \widehat{P}_1 & \widehat{P}_2 & \widehat{P}_3 & \widehat{F}_1 & \widehat{F}_2 & \widehat{F}_3 \end{array} \right]$$

$$\widehat{\boldsymbol{\beta}}' = \left[\begin{array}{cccccccccccccccc} \widehat{\mu} & \widehat{P}_1 & \widehat{P}_2 & \widehat{P}_3 & \widehat{F}_1 & \widehat{F}_2 & \widehat{F}_3 & \widehat{PF}_{11} & \widehat{PF}_{12} & \widehat{PF}_{13} & \widehat{PF}_{21} & \widehat{PF}_{22} & \widehat{PF}_{23} & \widehat{PF}_{31} & \widehat{PF}_{32} & \widehat{PF}_{33} \end{array} \right]$$

Če imamo težavo s postavitvijo hipoteze, si lahko pomagamo na naslednji način.

1) Sestavite linearno kombinacijo (vrstico), ki predstavlja pričakovano vrednost pri določeni pasmi!

$$E(y_i) = 1\mu + 1P_i + 1/3\,(F_1 + F_2 + F_3) \qquad [1.14]$$

Sestavimo linearni kombinaciji za pričakovano vrednost pri pasmah 1 in 2. Pri tem upoštevamo srednjo vrednost, vpliv izbrane pasme in povprečen učinek farm. Ker so farme tri, vzamemo tretjino vsake farme.

$$k'_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \qquad [1.15]$$

$$k'_2 = \begin{bmatrix} 1 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \qquad [1.16]$$

2) Sestavite linearno kombinacijo (vrstico), ki predstavlja razliko pričakovanih vrednosti med izbranima pasmama i in i'.

Poiščimo razliko pasme 1 (1.15) in pasme 2 (1.16). Iz dobljenega rezultata 1.17 vidimo, da je razlika med pasmama očiščena drugih vplivov.

$$k'_{12} = k'_1 - k'_2 = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad [1.17]$$

1.2 Vsota kvadratov in stopinje prostosti

Metode najmanjših kvadratov, tehtanih najmanjših kvadratov in splošnih najmanjših kvadratov sprejemajo svoje zaključke na osnovi

- **vsote kvadratov**, ki ga pojasnijo posamezni vplivi,
- **vsote kvadratov za ostanek**, ki praviloma služi za primerjavo, in
- **stopinj prostosti**, to je, številu parametrov, ki smo jih porabili za opis posameznega vpliva.

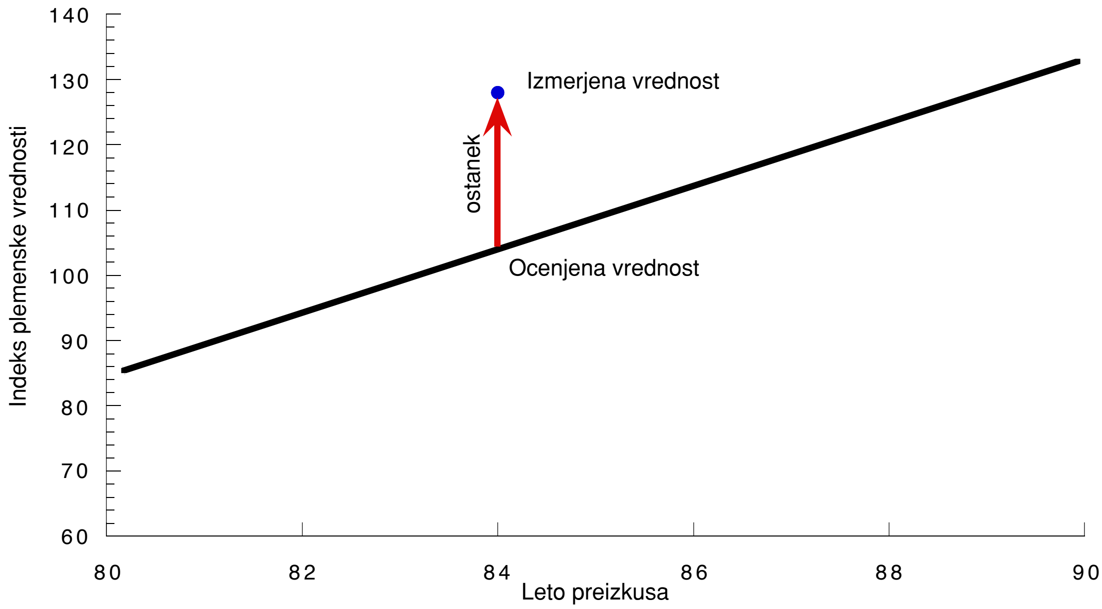
Pri biometriji moramo biti zelo natančni: ostanek (e) je razlika med resnično in ocenjeno vrednostjo. Ker pa resnične vrednosti ne poznamo, na njeni osnovi ne moremo narediti nobenih zaključkov. Preostane nam samo ena ali več meritev, s katerimi se poskušamo čimbolj približati dejanski vrednosti. Razlika med izmerjeno in ocenjeno vrednostjo je tako nadomestek dejanskega ostanka, je torej samo napoved za ostanek ($\hat{e}$). Brali boste lahko tudi o oceni ostanka, a ocena je povezana s sistematskimi vplivi, ostanek pa je naključna spremenljivka.

Vsote kvadratov si bomo ogledali kasneje, s stopinjami prostosti pa smo se spoznali že v poglavju o modelih.

1.3 Preveritev modela

Primer . Za ilustracijo primera ponovno obudimo primer enajstih merjenih mladic. V prvem delu bomo uporabili samo meritve za dnevni prirast (tabela 1.1). Poskusimo preveriti model! Zaradi lažjega razumevanja pa začnimo pri najbolj enostavnem modelu: v prvi model smo dali samo srednjo vrednost in ostanek. Ker bomo parametre ocenjevali po metodi najmanjših kvadratov, je kriterij za odločitve vsota kvadratov za ostanek. Seveda pa moramo najprej oceniti neznane parametre. V našem preprostem primeru je to samo srednja vrednost μ, ki znaša 550 g/dan.

$$y_i = \mu + e_i \qquad [1.18]$$



Slika 1.1: Napoved ostanka

Tabela 1.1: Izračun vsote kvadratov za ostanek pri modelu 1.18

Žival	Pasma	Mesec	Dnevni prirast (g/dan)	$\hat{\mu}$	$\hat{e}_i = y_i - E(y_i)$	$\hat{e}^2_{ij}$
1	SL	JAN	540	550	-10	100
2	SL	JAN	550	550	0	0
3	SL	FEB	550	550	0	0
4	SL	FEB	580	550	30	900
5	LW	JAN	520	550	-30	900
6	LW	FEB	500	550	-50	2500
7	LW	FEB	490	550	-60	3600
8	NL	JAN	560	550	10	100
9	NL	JAN	550	550	0	0
10	NL	FEB	600	550	50	2500
11	NL	FEB	610	550	60	3600
Skupaj						14200

Razvrstimo rezultate v tabelo 1.2. Vsoto kvadriranih meritev smo tako razdelili na del, ki ga pojasni srednja vrednost in ostanek. Vsoto kvadratov smo razdelili torej na dve neodvisni komponenti. Srednja vrednost je pojasnila skoraj vso variabilnost, za to pa smo porabili samo en parameter, eno stopinjo prostosti. V ostanku pa je ostalo še 10 stopinj prostosti. Ko ugotavljamo pomen parametrov, uporabimo srednji kvadrat. Ta pove, koliko vsote kvadratov v povprečju pojasni ena stopinja prostosti. Za primerjavo si praviloma izberemo srednji kvadrat za ostanek, le izjemona kaj drugega.

Tabela 1.2: Viri variabilnosti za dnevni prirast iz modela 1.18

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	2343.3	<0.0001
Ostanek	10	14200.00	1420.00		
Skupno	11	3341700.00			

•

Sedaj lahko obogatimo primer še s formulami. Uporabili bomo oznake izpeljane iz angleških izrazov.

a) Skupna vsota kvadratov (Total Sum of Square, TSS) je vsota kvadriranih opazovanj.

$$TSS = \sum y_i^2 \quad [1.19]$$

b) Korigirana skupna vsota kvadratov Iz skupne vsote kvadratov najprej odstranimo vsoto kvadratov, ki jo pojasnjuje ocena srednje vrednosti μ.

$$CTSS = \sum y_i^2 - \sum \widehat{\mu}^2 \quad [1.20]$$

c) Vsota kvadratov za model (model sum of square, MSS) je enaka vsoti kvadratov 1.21 za pričakovane vrednosti potem, ko smo odstranili vsoto kvadratov, ki jo pojasnjuje srednja vrednost. Z drugimi besedami MSS1.22 predstavlja razliko med korigirano vsoto kvadratov $CTSS$ in vsoto kvadratov za ostanek RSS.

$$MSS = \sum (E(y_i))^2 - \sum \widehat{\mu}^2 \quad [1.21]$$

$$MSS = CTSS - RSS \quad [1.22]$$

Izjema je model, ki vsebuje samo srednjo vrednost in ostanek. Tam ne izvrednotimo korigirane skupne vsote kvadratov, vsota kvadratov za model je kar vsota kvadratov, ki jo pojasnjuje srednja vrednost.

d) Vsota kvadratov za ostanek (residual sum of square, RSS) je vsota kvadriranih ostankov1.23.

$$RSS = \sum (y_i - E(y_i))^2 \quad [1.23]$$

e) Srednji kvadrat 1.24 dobimo tako, da vsoto kvadratov delimo s stopinjami prostosti. Znak x v enačbi zamenjamo s katerokoli vsoto kvadratov za model ali za posamezni vpliv.

$$MS_x = \frac{xSS}{d.f._x} \quad [1.24]$$

f) $F-$statistika je razmerje med dvema srednjima kvadratoma. V imenovalcu je tistai srednji kvadrat, s katerim primerjamo ostale. Kar praviloma je to srednji kvadrat za ostanek. F-statistika ima porazdelitev F, ko drži ničela hipoteza.

$$F = \frac{MS_x}{MS_e} \qquad [1.25]$$

g) $P-$vrednost je verjetnost, da vpliv (v našem primeru je to tudi celoten model ali pa smo srednja vrednost) ni pomemben. Pravzaprav bi morali reči, da je to verjetnost, da drži ničelna hipoteza. Da pa bi lažje razumeli, smo pač ubrali preprostejšo obliko. S testiranjem modela in posameznih vplivov presojamo, koliko variabilnosti smo pojasnili. Primerjavo praviloma delamo z variabilnostjo ostanka. Po domače bi lahko rekli, da del, ki ga pojasni posamezni vpliv, primerjamo z informacijami, ki so v ostanku še ostale.

h) Analiza variance za model Izračunane vrednosti uredimo v tabelo t1.3, kjer razčlenimo vsote kvadratov na posamezne komponente in določimo stopinje prostost ($d.f$). Pri preizkušanju modela v celoti imamo samo tri komponente: srednjo vrednost, ostali del modela in ostanek. Nato izvrednotimo srednje kvadrate, $F-$vrednosti in iz tabel odčitamo $P-$vrednosti. Praviloma nas ne zanima vrstica, ki je namenjena srednji vrednosti (prva vrstica v tabeli), test je usmerjen na model (druga vrstica v tabeli).

Tabela 1.3: Viri variabilnosti za dnevni prirast za model

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	$\sum \widehat{\mu}^2$	$\frac{\sum \widehat{\mu}^2}{1}$	$\frac{MS_\mu}{MS_e}$	
Model	$d.f.$	MSS	MS_M	$\frac{MS_M}{MS_e}$	
Ostanek	$n-1-d.f.$	RSS	MS_e		
Skupno	n	TSS			

Značilnost srednje vrednosti nas zanima, kadar obdelujemo razlike med pari. Tako bi poskus opravljali lahko na enojajčnih dvojčkih, sestrah/bratih in polsestrah/polbratih. Imamo dve pokusni skupini. Sorodnike uvrstimo v različni skupini in tako sestavimo pare. Skupini nista neodvisni: meritve povezuje genetski del variabilnosti. Da bi se motnji izognili, ne obdelamo meritev samih, ampak razlike med živalima v paru. Pri ničelni hipotezi je pričakovana vrednost (srednja vrednost) enaka nič.

Primer . Skupno vsoto kvadratov (TSS) bomo sedaj razdelili na tri komponente in sicer na tisto:

- kar pojasni srednja vrednost ($SS(\hat{\mu})$),
- kar pojasnijo ostali vplivi v modelu (MSS) in
- kar je ostalo (RSS).

Skupna vsota kvadratov 1.26 in vsota kvadratov 1.27, ki jo pojasni srednja vrednost μ, se nista spremenili v primerjavi z modelom 1.18 (glej tabelo 1.2). Korigirana vsota kvadratov $CTSS$ iz 1.28 je enaka kot vsota kvadratov za ostanek v enostavnem modelu 1.18.

$$TSS = 540^2 + 550^2 + 550^2 + 580^2 + ... + 600^2 + 610^2 = 3341700.00 \qquad [1.26]$$

$$SS(\hat{\mu}) = 550^2 + 550^2 + 550^2 + ... = 11 * 550^2 = 3327500.00 \qquad [1.27]$$

$$CTSS = TTS - SS(\hat{\mu}) = 3341700. - 3327500. = 14200.00 \qquad [1.28]$$

Uredimo v tabelo za analizo variance 1.4. Iz tabele lahko vidimo samo, da je srednja vrednost zelo različna od 0. Seveda to za dnevni prirast pri rastočih živalih tudi pričakujemo. Pri odraslih živalih, zlasti samicah v času laktacije, pa lahko imamo tudi negativne dnevne priraste. Ker živalim primanjkuje hranilnih snovi v zaužiti krmi za prirejo mleka, koristijo telesne rezerve. V takih primerih so lahko rezultati tudi drugačni. Na splošno pa nas povprečja ne zanimajo, da bi zmanjšali numerične probleme, ki jih računalnikom povzročajo velike številke, se statistični paketi srednje vrednosti znebijo in opravijo analizo variance brez nje. Mi jo bomo v prikazih zaradi kompletnosti obdržali, rezultati pa zaradi tega niso nič boljši in nič slabši. So enaki.

Tabela 1.4: Viri variabilnosti za dnevni prirast iz modela 1.18

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	2343.31	<0.0001
Ostanek	10	14200.00	1420.00		
Skupno	11	3341700.00			

Pravzaprav v ostanku pri modelu 1.18 ni ostalo veliko stvari nepojasnjenih. Vseeno dodajmo modelu 1.18 vpliv pasme ($\mathbf{P}_i$). Tako dobimo še vedno preprost model 1.29.

$$y_{ij} = \mu + P_i + e_{ij} \tag{1.29}$$

Vpliv pasme predstavlja edini res pravi vpliv v modelu1.29. Tako smo iz vsote kvadratov za ostanek iz tabele 1.36 oziroma $CTSS$ iz 1.28 pojasnili še dodatno variabilnost, ki je ocenjena v enačbi 1.30. Slednja vsota predstavlja kvadrirane odklone srednjih vrednosti po pasmah od skupne srednje vrednosti za vsako meritev. Ker vemo, da imamo pri pasmi 1 štiri meritve, pri pasmi 2 tri in pri pasmi 3 zopet štiri meritve, smo izračun pač nekoliko poenostavili (enačba 1.30).

$$MSS = SS(P) = 4 * (555 - 550)^2 + 3 * (503.33 - 550)^2 + 4 * (580 - 550)^2 = 10233.33 \tag{1.30}$$

Za isto vsoto pa se je zmanjšala vsota kvadratov za ostanek 1.31.

$$RSS = \hat{e}_{ij}^2 = 3966.67 \tag{1.31}$$

Izračun posameznih vsot kvadratov smo ponazorili tudi v tabeli 1.5.

Tabela 1.5: Izračun vsote kvadratov za ostanek pri modelu 1.29

Žival	Pasma	Mesec	Dnevni prirast (g/dan)	$\hat{\mu} + \hat{P}_i$	$\hat{P^2}_i$	$\hat{e}_i = y_i - E(y_i)$	$\hat{e}_{ij}^2$
1	SL	JAN	540	555.00	25.00	-15.00	225.00
2	SL	JAN	550	555.00	25.00	-5.00	25.00
3	SL	FEB	550	555.00	25.00	-5.00	25.00
4	SL	FEB	580	555.00	25.00	25.00	625.00
5	LW	JAN	520	503.33	2177.77	16.67	277.89
6	LW	FEB	500	503.33	2177.77	-3.33	11.09
7	LW	FEB	490	503.33	2177.77	-13.33	177.69
8	NL	JAN	560	580.00	900.00	-20.00	400.00
9	NL	JAN	550	580.00	900.00	-30.00	900.00
10	NL	FEB	600	580.00	900.00	20.00	400.00
11	NL	FEB	610	580.00	900.00	30.00	900.00

Sedaj uredimo vsote kvadratov še v tabelo za analizo variance (1.6) in izvrednotimo srednje kvadrate, $F-$statistiko in določimo $P-$vrednost. Srednja vrednost je tudi v tem modelu pojasnila največji del variabilnosti, za kar smo porabili 1 stopinjo prostosti. V podatkih smo imeli 3 pasme, zato porabimo za vpliv pasme 2 stopinji prostosti, za ostanek nam je ostalo samo 8 stopinj prostosti. Kljub temu razmerje med srednjim kvadratom za vpliv pasme in srednjim kvadratom za ostanek pokaže, da je vpliv pasme pomemben. Tudi $P-$vrednost, ki jo preberemo iz tabel oziroma izračunamo, potrjuje naše sklepanje. Ker je vpliv pasme edini vpliv v modelu 1.29, veljajo isti zaključki tudi za celotni model. Kot smo že omenili, vsoto kvadratov, ki jo povzroča srednja vrednost, obravnavamo posebej. Pravzaprav se z njo praviloma niti ne ukvarjamo.

Tabela 1.6: Viri variabilnosti za dnevni prirast iz modela 1.29

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	6710.97	<0.0001
Pasma	2	10233.33	5116.67	10.32	0.0061
Ostanek	8	3966.67	495.83		
CTSS	10	14200.00			
Skupno	11	3341700.00			

Primer . Vzemimo še en enostaven model in vključimo vanj le vpliv meseca (M_i)1.32.

$$y_{ij} = \mu + M_i + e_{ij} \qquad [1.32]$$

Vsoto kvadratov za model (enačba 1.33) izračunamo podobno kot v zgornjem primeru (enačba 1.30). Pojasnjena vsota je precej manjša kot pri pasmi. Dobili smo jo tako, da smo kvadrirali odklone srednjih vrednosti po pasmah od skupne srednje vrednosti za vsako meritev. Pričakovana vrednost za januar je 544, za februar pa 555. Ker vemo, da imamo v januarju pet meritev, v februarju pa šest, smo izračun pač nekoliko poenostavili (enačba 1.33).

$$MSS = SS(M) = 5 * (544 - 550)^2 + 6 * (555 - 550)^2 = 330.00 \qquad [1.33]$$

Izvrednotiti moramo še vsoto kvadratov za ostanek (1.34). Dobimo jo lahko tako, da izračunamo ostanke, jih kvadriramo in kvadrate seštejemo.

$$RSS = \hat{e}_{ij}^2 = 13870 \qquad [1.34]$$

Lahko pa uberemo krajšo pot (enačba 1.35). Od korigirane skupne vsote kvadratov ($CTSS$) smo odšteli tisti del (MSS), ki ga pojasni model.

$$RSS = CTSS - MSS = 14200 - 330 = 13870 \qquad [1.35]$$

Uredimo izračune v tabelo za analizo variance 1.7. Vpliv meseca je nepomemben. Vsota kvadratov in srednji kvadrat sta majhna v primerjavi z ostankom. Ker je bil vpliv pasme pomemben, že sedaj vemo, da so zaključki iz modela z mesecem neuporabni. Model smo uporabili le zato, da bomo kasneje lažje razmišljali o dodajanju vplivov v modele in presojanju pomena dodatnih vplivov.

Primer . Dodajmo modelu z vplivom pasme (enačba 1.29) še vpliv meseca, kot prikazuje model 1.36.

$$y_{ijk} = \mu + P_i + M_j + e_{ijk} \qquad [1.36]$$

Tabela 1.7: Viri variabilnosti za dnevni prirast iz modela 1.29

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	2159.16	<0.0001
Mesec	1	330.00	330.00	0.21	0.6545
Ostanek	9	13870.00	1541.11		
CTSS	10	14200.00			
Skupno	11	3341700.00			

Skupna vsota kvadratov in vsota kvadratov za srednjo vrednost sta ostali nespremenjeni. Iz tega sledi, da je nespremenjena tudi korigirana vsota kvadratov $CTSS$. Vsota kvadratov za ostanek 1.37 je zmanjšana, kar je pričakovano: z novim vplivom pričakujemo, da bodo podatki bolje predstavljeni.

$$RSS = 2900.00 \qquad [1.37]$$

Vsota kvadratov za model 1.38 je tako povečana. Oba vpliva v modelu skupaj pojasnita pomemben delež variabilnosti. Srednji kvadrat za model je zmanjšan, ker smo za pojasnitev porabili večje število stopinj prostosti. Nekoliko se je zmanjšala tudi $F-$statistika, kar pa ni močno vplivalo na verjetnost P. To seveda ne velja za vse modele. V našem primeru imamo majhno število opazovanj, dokaj izenačene skupine, izbrali pa smo tudi meseca, ko so proizvodni rezultati bolj podobni. Dodatno je bilo pojasnjeno le nekaj malega vsote kvadratov. Dobra informacija o tem, koliko model doprinese, je vsota kvadratov, ki jo pojasni ena stopinja prostosti. V modelu 1.36 imamo tri stopinje prostosti, vsota kvadratov za model je nekoliko povečana, srednji kvadrat, vsota kvadratov na stopinjo prostosti pa je zmanjšana. Model je še vedno značilen, med dvema vplivoma v modelu je vsaj eden statistično značilen, pri vsaj enem bomo ovrgli ničelno in sprejeli altrnativno hipotezo. Naša naloga je, da sedaj ugotovimo, kateri vpiv je to. Lahko pa bi bila tudi oba.

$$MSS = CTSS - RSS = 14200.00 - 2900.00 = 11300.00 \qquad [1.38]$$

Tabela 1.8: Viri variabilnosti za dnevni prirast iz modela 1.36

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	8031.90	<0.0001
Model	3	11300.00	3766.67	9.09	0.0082
Ostanek	7	2900.00	414.29		
Skupno	11	3341700.00			

1.4 Preveritev vplivov

Nadaljujmo kar z istim primerom. Novi model (1.36) je skupno pojasnil večjo vsoto kvadratov, na porabljeno stopinjo prostosti smo pojasnili nekoliko manj kot pri prejšnjem modelu (1.29), a je kljub temu zadostovalo, da je model značilen. Nadalje nas zanima, koliko k modelu doprineseta posamezna vpliva.

1.4.1 Vsota kvadratov tipa I

Vsota kvadratov tipa I je izračunana iz razlike med polnim modelom in poenostavljenim modelom, kjer smo predpostavili, da je opazovani vpliv nepomemben in smo ga zato izpustili. V tem primeru smo vsoto kvadratov razdelili tako, da je vsota vseh posameznih vsot kvadratov natanko skupna vsota kvadratov. Imenujemo jih tudi sekvenčne vsote kvadratov.

Nastavimo tabelo za analizo variance pri modelu 1.36. Vsoto kvadratov za model (enačba 1.38) moramo razdeliti na vsoto, ki jo pojasni pasma, in vsoto, ki jo pojasni mesec. Vsoto kvadratov za pasmo smo že izračunali v enačbi 1.30. Razlika (1.41) med vsotama kvadratov za modela 1.36 in 1.29 je vsota kvadratov, ki jo pri tipu I pripišemo vplivu mesec. Mesec je v tem primeru vključen za pasmo.

$$SS(M) = 11300.00 - 10233.33 = 1066.67 \qquad [1.39]$$

Uredimo rezultate v tabelo za analizo variance 1.9. Razvidno je, da so med pasmami razlike, med meseci pa ne. Toda pa bodite pozorni! Verjetnost (P-vrednost) se je za mesec precej zmanjšala v primerjavi, ko v modelu ni bilo pasme (tabela 1.7). Tako je vpliv meseca skoraj postal značilen, kar pri pitanju prašičev običajno pričakujemo. Neznačilen je morda zato, ker imamo malo opazovanj ali pa se meseca januar in februar nista bistveno razlikovala v temperaturi ali drugih klimatskih dejavnikih. Praviloma sta to tipična zimska meseca.

Tabela 1.9: Viri variabilnosti za dnevni prirast iz modela 1.36 tip - I

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	8031.90	<0.0001
Pasma	2	10233.33	5116.67	12.35	0.0051
Mesec	1	1066.67	1066.67	2.57	0.1526
Ostanek	7	2900.00	414.29		
Skupno	11	3341700.00			

Sedaj pa uporabimo isti model, le vrstni red vplivov v modelu zamenjajmo.

$$y_{ijk} = \mu + M_i + P_j + e_{ijk} \qquad [1.40]$$

Vsoto kvadratov za ostanek se ne spremeni. Ker je mesec prvi vpliv, zanj velja vsota kvadratov, izračunana v enačbi 1.33.

$$SS(P) = 11300.00 - 330.00 = 10970.00 \qquad [1.41]$$

Uredimo rezultate še v tabelo za analizo variance 1.10. Vsoto kvadratov za model smo razdelili v prvem (tabela 1.9) in drugem (tabela 1.10) primeru različno. Zaključki sicer slučajno niso različni, vendar pa se lahko zgodi celo to. Ko smo vpliv dodali kot drugi vpliv, je pojasnil več variance kot takrat, ko smo ga napisali na prvo mesto. Primer pa nam vseeno jasno pokaže, da je pri tem načinu izbora vsot lahko dobimo različne zaključke. Če se držimo nenapisanega pravila, da navajamo vplive v modelih glede na značilnost (oziroma glede na srednje kvadrate), in predvsem pravilno interpretiramo, pa se neljubim zapletom lahko izognemo. Kljub vsemu bi se radi izognili različnim rezultatom, zato bomo poiskali boljšo rešitev. Vrstni red v modelu pač ne sme vplivati na zaključke.

Tabela 1.10: Viri variabilnosti za dnevni prirast iz modela 1.40 tip - I

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	8031.90	<0.0001
Mesec	1	330.00	330.00	0.80	0.4018
Pasma	2	10970.00	5485.00	13.24	0.0042
Ostanek	7	2900.00	414.29		
CTSS	10	14200.00			
Skupno	11	3341700.00			

Vsota kvadratov tipa I je izračunana vsakokrat avtomatsko. Izračunana vsota kvadratov je odvisna od vrstnega reda vplivov v modelu. Zanje tudi velja, da vsota predstavlja vsoto kvadratov za model brez

Tabela 1.11: Zmanjšanje vsote kvadratov v modelu s tremi vplivi

Vpliv	Tip I	Tip II	Tip III	Tip IV
A	$R(A)$	$R(A \mid B, C)$		
B	$R(B \mid A)$	$R(B \mid A, C)$		
C	$R(C \mid A, B)$	$R(C \mid A, B)$		

Tabela 1.12: Zmanjšanje vsote kvadratov v modelu z dvema vplivoma in interakcijo

Vpliv	Tip I	Tip II	Tip III	Tip IV
A	$R(A)$	$R(A \mid B)$		
B	$R(B \mid A)$	$R(B \mid A)$		
$A * B$	$R(A * B \mid A, B)$	$R(A * B \mid A, B)$		

vsote kvadratov, ki jo pojasni srednja vrednost. V primeru neuravnoteženih podatkov vsot kvadratov tipa I ne smemo uporabljati, ker so odvisne od strukture podatkov.

Preizkusi tipa I so primerni za:

- uravnotežene ANOVA modele, če zagotovimo pravilni vrstni red vplivov (npr. interakcije za glavnimi vplivi ...)
- popolnoma hierarhične modele, če zagotovimo pravilni vrstni red vplivov (npr. vgnezdeni za nadrejenimi ...)
- regresijske modele s polinomi, če zagotovimo pravilni vrstni red vplivov (npr. višje stopnje sledijo nižjim ...).

Zmanjšanje vsote kvadratov za ostanek

Predno nadaljujemo se bomo dogovorili še za poseben zapis, s katerim bomo opisali zmanjšanje (redukcija) vsote kvadratov za ostanek.

$R(P)$ - zmanjšanje vsote kvadratov zaradi vpliva P

$R(P|\mu)$- zmanjšanje vsote kvadratov za ostanek, ko modelu s srednjo vrednostjo dodamo še vpliv P

$R(P|\mu, M)$- zmanjšanje vsote kvadratov za ostanek, ko modelu s srednjo vrednostjo in vplivom M dodamo še vpliv P

Pri modelu s tremi vplivi A, B in C razdelimo vsoto kvadratov na načina prikazana v tabeli 1.11. Vsote kvadratov pri tipu I dobimo tako, da sekvenčno dodajamo vplive. Vrstni red dodajanja vplivov je pomemben. Pri tipu II pa izvrednotimo, koliko pridobimo, če ostalim vplivom v modelu dodamo še vpliv, za katerega računamo vsoto kvadratov.

1.4.2 Vsota kvadratov tipa II

Vsota kvadratov pri tipu II ni odvisna od vrstnega reda vplivov v modelu. Hipoteze naj bi bile pravilne za večino setov podatkov, primerov, če lahko zagotovimo, da ni v modelu interakcij ali vgnezdenih vplivov. Vsota kvadratov za interakcijo in dodatni vpliv je pravilna, nepravilna je vsota kvadratov za vpliva, med katerima nastopa interakcija. Če je interakcija neznačilna, bo test za glavni vpliv tudi sprejemljiv.

Pričakovano se je spremenila razporeditev vsote kvadratov med vplivoma pasma in mesec ter ostankom. Novi vpliv mesec je pojasnil dobro četrtino ostanka iz enostavnejšega modela 1.29. Nekoliko večja je bila tudi vsota kvadratov za pasmo. Ta prerazporeditev je posledica nekoliko spremenjenih rešitev za vpliv pasme, ko vključimo dodatno še vpliv meseca.

$$SS(P) = 10970.00 \qquad [1.42]$$

$$SS\,(M) = 1066.67 \qquad [1.43]$$

Vsota kvadratov za model naj bi bila tudi vsota kvadratov vseh vplivov v modelu. V našem primeru imamo vpliv pasme in vpliv meseca. Če vsoti seštejemo 1.44, pa dobimo večjo vsoto kvadratov kot pri 1.38. Vsote kvadratov niso neodvisne. Tako smo razliko 12036.67 – 11300.00, kar znese 736.67, šteli dvakrat: enkrat pri pasmi in enkrat pri mesecu. Oba vpliva smo obravnavali s pretvezo, da drugega ni v modelu. Tako smo prišli do nelogičnega rezultata, da skupek vplivov pojasni več variabilnosti kot model. Na ta način pojasnjujemo neko dodatno variabilnost, ki je sploh ni.

$$MSS = SS(P)+SS\,(M)=10970.00+1066.67=12036.67 \qquad [1.44]$$

Vsota kvadratov za model, ko smo odstranili vsoto kvadratov za srednjo vrednost, pri tipu I znaša 11300.00. Pri tipu II je vsota kvadratov za model večja in sicer znaša 12036.67.

Vsekakor razliko 736.67 ne smemo kar izbrisati, potem bi bil seštevek premajhen. Ena od možnosti je prikazana v tabeli 1.13, da razliko 736.67 upoštevamo pri vplivu pasme, pri vplivu meseca pa ne. Vpliv pasme je značilen, pomemben, kar smo dokazali že s preprostejšim modelom 1.29, v modelu 1.36 z dodatnim vplivom se je vpliv pasme še bolj potrdil. To sicer ne smemo posplošiti na vse primere. Vsota kvadratov za vpliv meseca je v tem primeru sorazmeroma majhna (330.00). Ko imamo v modelu že pasmo, z mesecem ne pridobimo veliko.

Tabela 1.13: Viri variabilnosti za dnevni prirast iz modela 1.36 tip - II

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00	3327500.00	8031.81	<0.0001
Pasma	2	10970.00	5485.00	13.24	0.0042
Mesec	1	1066.67	330.00	0.797	0.1526
Ostanek	7	2900.00	414.29		
CTSS	10	14200.00			
Skupno	11	3341700.00			

$$y_{ijk} = \mu + P_i + M_j + PM_{ij} + e_{ijk} \qquad [1.45]$$

Pri modelu z interakcijami za glavna vpliva P in M ne moremo poiskati vsote kvadratov, ki bi model očistila tudi interakcije PM. Interakcijo lahko vključimo šele, ko sta v modelu oba glavna vpliva. Kadar je interakcija značilna, preizkus glavnih vplivov s pomočjo vsote kvadratov tipa II ni primeren.

Tabela 1.14: Viri variabilnosti za dnevni prirast iz modela 1.45 tip - II

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00			
Mesec	1	1066.67	1066.67	8.21	0.0352
Pasma	2	10970.00	5485.00	42.19	0.0007
Mesec*pasma	2	2250.00	1125.00	8.65	0.0238
Ostanek	5	650.00	130.00		
CTSS	10	14200.00			
Skupno	11	3341700.00			

Tip II vsote kvadratov so primerne:

- za uravnotežene primere (drugače odvisni od strukture podatkov)
- za modele samo z glavnimi vplivi

- za čiste regresijske modele
- za vpliv, ki ni vključen v drugi vpliv
- uporaben tudi za popolnoma hierarhične modele

1.4.3 Vsota kvadratov tipa III in IV

Vsote kvadratov za ta dva tipa so vse izračunane z medodo splošnih linearnih hipotez. Uporabnik mora poznati ocenljive funkcije ali pa si jih izpisati, da prepozna hipoteze, ki so bile preverjene. Potreboval jih bo pri interpretaciji rezultatov.

Vsota kvadratov tipa III 1.14 za posamezni vpliv je neodvisna od vrstnega reda. Predstavlja vsoto, ki je dodatno pojasnjena, če je vpliv v modelu, oziroma je izpuščen. Pri tem smo popustli pri dejstvu, da se vsote kvadratov seštejejo do skupne vsote kvadratov. Izračun vsote kvadratov temelji na hipotezi, ki jo želimo preveriti. O hipotezah in ocenljivosti se bomo pogovarjali kasneje. Ker izračun vsot kvadratov za posamezne vplive ni enostaven, bomo verjeli statističnim paketom. Iste preizkuse živinorejci poznajo iz “Harvey-evega programa”. Tip III lahko praktično vedno uporabljamo. Vsote tipa I ali II pa uporabljamo v živinoreji samo takrat, ko so vrednosti v tabelah enake kot pri tipu III ter pri popolnem hierarhičnem modelu. Učimo se jih bolj zaradi razumevanja. Včasih moramo poznati enostavnejši primer, da razumemo malo bolj zapletene.

Pomembna predpostavka pri tipu III je, da so vse celice – vsi podrazredi– zasedeni. Polna celica ima najmanj eno opazovanje. Praviloma to ni zadostno za dober poskus, a to je že druga zgodba. Če vemo, da je ena celica slabo zasedena, pa tistih nekaj podatkov pustimo v obelavi, ker bo izpeljava hipotez in s tem interpretacija lažja. Vedeti pa moramo, da bodo vse primerjave s slabo zasedeno celico oziroma skupino nezanesljive. Če so pri interakciji manjkajoče celice, izberemo vsote kvadratov tipa IV, ker so lahko boljše. Še vedno velja, da poskus ni bil najbolj posrečeno zasnovan. Zgodi pa se lahko, da smo šele na koncu poskusa ugotovili, da je interakcija pomembna. Takrat pa celic ne moremo več popolniti in iz poskusa poskušamo izvleči, kar se da.

Pri modelu z vplivom meseca in pasme 1.36 je delitev vsote kvadratov pri tipu III (tabela 1.15) enaka kot pri tipu II (tabela 1.13). Vpliv meseca ni značilen. Če dodamo vpliv pasme, dodatni vpliv pojasni vsoto kvadratov v znesku 10970.00 in za to porabi dve stopinji prostosti. Dodani del variabilnosti tudi v tem primeru pojasni pomemben delež variabilnosti. Vpliv pasme je značilen. Za preizkus vplivov lahko uporabimo vsoto kvadratov po tipu II ali III.

Tabela 1.15: Viri variabilnosti za dnevni prirast iz modela 1.36 tip - III

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00			
Mesec	1	1066.67	1066.67	2.57	0.1526
Pasma	2	10970.00	5485.00	13.24	0.0042
Ostanek	7	2900.00	414.29		
CTSS	10	14200.00			
Skupno	11	3341700.00			

Pri modelu z vključeno interakcijo (1.45) se vsote kvadratov med tipoma II (1.14) in III (1.16) razlikujeta. Pri interakciji je vsota kvadratov enaka, pri glavnih vplivih pa je pri tipu II precenjena. Pri podobnih modelih uporabljamo pri preizkušanju vplivov vsote kvadratov, izračunane po tipu III.

V uporabljenih modelih so vsote kvadratov pri tipu IV enake kot pri tipu III, ker nimamo praznih celic.

Tabela 1.16: Viri variabilnosti za dnevni prirast iz modela 1.45 tip - III

Vir variabilnosti	d.f.	Vsota kvadratov	Srednji kvadrat	F-vrednost	P-vrednost
Srednja vrednost	1	3327500.00			
Mesec	1	558.57	558.57	4.45	0.0887
Pasma	2	8450.00	4225.00	32.50	0.0014
Mesec*pasma	2	2250.00	1125.00	8.65	0.0238
Ostanek	5	650.00	130.00		
CTSS	10	14200.00			
Skupno	11	3341700.00			