Dizajni i latinski kvadrati

Andrea Švob
(asvob@math.uniri.hr)

25. svibnja 2011.

Dizajni

Definicija

Konačna incidencijska struktura $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ *je* $t - (v, k, \lambda)$ **dizajn** *ako vrijedi sljedeće:*

1. $|\mathcal{P}| = v$,
2. *svaki element skupa* $\mathcal{B}$ *incidentan je s točno* k *elemenata skupa* $\mathcal{P}$,
3. *svakih* t *elemenata skupa* $\mathcal{P}$ *incidentno je s točno* λ *elemenata skupa* $\mathcal{B}$.

Definicija

$2 - (v, k, \lambda)$ *dizajn naziva se blok dizajn.*

Teorem

Neka je $\mathcal{D}$ $t-(v,k,\lambda)$ dizajn. Tada je $\mathcal{D}$ ujedno i $s-(v,k,\lambda_s)$ dizajn, za svaki s iz skupa $\{1,\ldots t-1\}$, gdje je

$$\lambda_s\binom{k-s}{t-s}=\lambda\binom{v-s}{t-s}.$$

Svaki $t-$ dizajn ujedno je i $s-$ dizajn za $s<t$.

Statistička teorija dizajna

- Dizajni su zanimljivi i statističarima,
- koriste ih u eksperimentima, posebno komparativnim eksperimentima,
- postoje različiti načini određivanja efikasnosti dizajna za eksperiment. (primjer 1-dizajn)

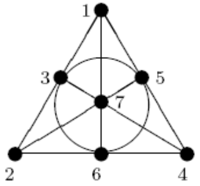
- Cilj eksperimenta: Testiranje 7 vrsti gnojiva
- Koristi se 21 prostor, 3 prostora na svakoj od 7 farmi u različitim dijelovima države.
- Opis: Primjenjujemo gnojiva na prostore, posadimo biljku i mjerimo rezultate.
- Osnovno pitanje: Kako ćemo raspodijeliti gnojivo po prostorima?
- Najgore što možemo učiniti je da stavimo npr. gnojivo 1 na 3 prostora na farmi A, gnojivo 2 na 3 prostora na farmi B, itd... - rezultat koji na taj način dobijemo nije vezan samo za gnojivo već i za sastav tla, klimatske uvjete i rezlike između farmi - loš odabir.
- Rješenje: sljedeća shema.

Primjer

	Farm A	Farm B	Farm C	Farm D	Farm E	Farm F	Farm G
Plot 1	1	1	1	2	2	3	3
Plot 2	2	4	6	4	5	4	5
Plot 3	3	5	7	6	7	7	6

Primjer

Fanova ravnina



Informacijska matrica

Definicija

Neka je Δ $1-(v,k,r)$ *dizajn, sa* $v \geq 2$. *Konkurencijska matrica dizajna* Δ *je* $v \times v$ *matrica* $\Lambda = (\lambda_{ij})$ *takva da* λ_{ij} *predstavlja broj blokova koji sadrže točke* x_i *i* x_j.

Definicija

Informacijska matrica od Δ *je*

$$F(\Delta) := rI_v - k^{-1}\Lambda,$$

gdje je Λ *konkurencijska matrica od* Δ.

Definicija

Svojstvene vrijednosti $\delta_0 \leq \delta_1 \leq ... \leq \delta_{v-1}$ *od* $F(\Delta)$ *nazivaju se* **kanonski faktori efikasnosti**.

Mjere efikasnosti za 1−dizajne

- Neka je Δ $1-(v,k,r)$ dizajn sa $v \geq 2$ i neka su $\delta_0 \leq \delta_1 \leq ... \leq \delta_{v-1}$ kanonski faktori efikasnosti.

Definicija

Ako Δ nije povezan (odgovarajući graf nije povezan), tada je $A_\Delta = D_\Delta = E_\Delta := 0$. U suprotnom, definiramo **mjere efikasnosti***:*

$$A_\Delta := (v-1)/\sum_{i=0}^{v-1} 1/\delta_i,$$

$$D_\Delta := (\prod_{i=0}^{v-1} \delta_i)^{1/\delta_i},$$

$$E_\Delta := \delta_1 = min\{\delta_0, ..., \delta_{v-1}\}.$$

- Ako je $k = v$ (Δ ima potpune blokove) tada je svaki od kanonskih faktora efikasnosti jednak 1 pa su tome jednake i mjere efikasnosti,
- ako je $k < v$ (tada Δ ima nepotpune blokove), želimo minimizirati gubitak informacije zbog blokova i želimo da gornje mjere efikasnosti budu čim bliže 1.
- Za testiranje pretpostavke eksperimenta pomoću dizajna, moramo usporediti mjere efikasnosti - koristi se GAP, paket DESIGN.

Statistička optimalnost

- Neka je Δ $1-(v,k,r)$ dizajn sa $v \geq 2$ i neka su $\delta_0 \leq \delta_1 \leq ... \leq \delta_{v-1}$ kanonski faktori efikasnosti.

Definicija

Δ je $A-$optimalan u klasi $\mathcal{C}$, klasi $1-(v,k,r)$dizajna koji sadrže Δ, ako je $A_\Delta \geq A_\Gamma$, za svaki $\Gamma \in \mathcal{C}$.

- Slično se definiraju D-optimalnost i E-optimalnost.

Definicija

Δ je Schur-optimalan u klasi $1-(v,k,r)$dizajna koji sadrže Δ, ako za svaki $\Gamma \in \mathcal{C}$, sa kanonskim faktorima efikasnosti $\gamma_0 \leq ... \leq \gamma_{v-1}$, vrijedi $\sum_{i=0}^{l} \delta_i \geq \sum_{i=0}^{l} \gamma_i$, za $l = 0, ..., v-1$.

Giovagnoli i Wynn, 1981. su dokazali:

- Schurov optimalan dizajn ne mora postojati u zadanoj klasi $\mathcal{C}$ $1-(v,k,r)$ dizajn, ali ako postoji taj dizajn je optimalan u klasi $\mathcal{C}$ sa visokim stupnjem statističkog kriterija optimalnosti uključujući A, D i E optimalnost.
- Ako je Δ $2-(v,k,\lambda)$ dizajn, tada su kanonski faktori efikasnosti jednaki $v(k-1)/((v-1)k)$, iz čega slijedi da je Δ Schur optimalan u klasi svih $1-(v,k,\lambda(v-1)/(k-1))$ dizajna. (2−dizajni sa svojstvima koji nam trebaju u eksperimentima ne moraju postojati.)

Latinski kvadrati

Definicija

Latinski kvadrat reda n je $n \times n$ polje L, čija svaka pozicija sadrži neki od elemenata $n-$članog skupa Ω, tako da svaki redak i svaki stupac tog polja sadrži svaki element točno jednom.

Definicija

Dva latinska kvadrata L_1 i L_2, reda n, sa skupom simbola Ω_1 i Ω_2 su ortogonalna ako za svaki $(\alpha_1, \alpha_2) \in \Omega_1 \times \Omega_2$, postoji jedinstven par (i, j) takav da je α_1 (i, j)-ta pozicija u L_1 i α_2 (i, j)-ta pozicija u L_2.

1	2	3
3	1	2
2	3	1

4	5	6
5	6	4
6	4	5

Definicija

Skup od m međusobno ortogonalnih latinskih kvadrata reda n, MOLS(n) je skup od m latinskih kvadrata $\{L_1, ..., L_m\}$ *takvih da je* $\forall 1 \leq i < j \leq m$ L_i *ortogonalan na* L_j.

- Neka je $n > 1$. Skup MOLS reda n ima veličinu najviše $n - 1$. Egzistencija skupa MOLS reda n i veličine $n - 1$ je ekvivalentno egzistenciji projektivne ravnine reda n.

Semi-latinski kvadrati

Definicija

$(n \times n)/k$ semi-latinski kvadrat je $n \times n$ polje S, čije su pozicije $k-$podskupovi od $nk-$skupa Ω, skupa simbola za S, takvih da se svaki simbol nalazi u točno jednoj poziciji u svakom retku i točno jednoj poziciji u svakom stupcu od S. Pozicija, $i-$ti redak i $j-$ti stupac se zove $(i,j)-$ta pozicija od S i zapisuje kao $S(i,j)$.

- Pretpostavit ćemo da je $n > 1$, $k > 0$ (da izbjegnemo trivijalan slučaj).

Primjer

$(5 \times 5)/3$ *semi-latinski kvadrat*

1 2 11	3 4 12	5 6 13	7 8 14	9 10 15
3 5 15	1 7 11	2 9 12	4 10 13	6 8 14
4 6 14	2 8 15	1 10 11	3 9 12	5 7 13
7 9 13	5 10 14	3 8 15	2 6 11	1 4 12
8 10 12	6 9 13	4 7 14	1 5 15	2 3 11

- $(n \times n)/1$ semi-latinski kvadrat je latinski kvadrat reda n.
- Semi-latinski kvadrati imaju mnoge primjene u statistici - agrikulturalni eksperimenti, testiranje kupaca i autentifikacija poruka.

Definicija

s-terostruka inflacija (s-fold inflation) od $(n \times n)/k$ *semi-latinskog kvadrata dobiva se tako da svaki simbol semi-latinskog kvadrata zamijenimo sa* s *novih simbola i kao rezultat dobijemo* $(n \times n)/(ks)$ *semi-latinski kvadrat.*

Primjer

$(3 \times 3)/2$ *semi-latinski kvadrat dobiven pomoću* 2*-terostruka inflacije latinskog kvadrata reda* 3*.*

1 4	2 5	3 6
3 6	1 4	2 5
2 5	3 6	1 4

Definicija

Superpozicija $(n \times n)/k$ *semi-latinskog kvadrata sa* $(n \times n)/l$ *semi-latinskim kvadratom (sa disjunktnim skupom simbola) dobiva se superponiranjem prvog kvadrata s drugim i kao rezultat dobivamo* $(n \times n)/(k + l)$ *semi-latinski kvadrat.*

Primjer

$(3 \times 3)/2$ *semi-latinski kvadrat dobiven kao superpozicija od* 2 *međusobno ortogonalna latinska kvadrata reda* 3*.*

1 4	2 5	3 6
3 5	1 6	2 4
2 6	3 4	1 5

Osnovni dizajn

Neka je S $(n \times n)/k$ semi-latinski kvadrat.

- Ako zanemarimo retke i stupce koji čine strukturu od S dobivamo osnovni (underlying) dizajn $\Delta(S)$, čije su točke simboli od S i čiji su blokovi pozicije od S.
- $\Delta(S)$ je $1-(nk, k, n)$ dizajn.
- Bailey, (1992) - S je optimalan s obzirom na kriterij statističke optimalnosti ako je $\Delta(S)$ optimalan s obzirom na isti kriterij u klasi osnovnih dizajna $(n \times n)/k$ semi-latinskih kvadrata.
- Cheng i Bailey, (1991) - dokazali su da je superpozicija MOLS (sa parovima disjunktnih skupova) A, D i E optimalan.
- Bailey (1992) - dokazano - za svaki $s \geq 1$ $s-$terostruka inflacija superpozicije potpunih skupova MOLS je A, D i E optimalna.

Uniformni semi-latinski kvadrati

Definicija

$(n \times n)/k$ semi-latinski kvadrat S je uniforman ako se svaki par ulaza od S, koji se ne nalaze u istom retku i istom stupcu, sijeku u konstantnom broju $\mu = \mu(S)$ točaka.

Primjer

Uniformni $(3 \times 3)/2$ semi-latinski kvadrat S sa $\mu(S) = 1$.

1 4	2 5	3 6
3 5	1 6	2 4
2 6	3 4	1 5

Lema

Ako je S uniforman $(n \times n)/k$ semi-latinski kvadrat tada je $\mu(S) = k/(n-1)$ i posebno, $n-1$ dijeli k.

Dokaz.
Neka je S uniformni $(n \times n)/k$ semi-latinski kvadrat i neka su $i, j \in \{1, ..., n\}$.
Prebrojavamo u dva smjera broj trojki (i', j', α), takve da je $i', j' \in \{1, ..., n\}, i' \neq i, j' \neq j$ i $\alpha \in S(i,j) \cap S(i', j')$.
Dobivamo $(n-1)^2 \mu(S) = k(n-1)$ odakle slijedi tvrdnja leme.

∎

Lema

Ako je S uniforman semi-latinski kvadrat tada je $s-$terostruka inflacija od S također uniformna. Ako su S i T oboje uniformni semi-latinski kvadrati (sa disjunktnim skupom simbola) tada je superpozicija od S i T također uniformna.

Teorem

$(n \times n)/(n-1)$ semi-latinski kvadrat S je uniforman ako i samo ako je S superpozicija od $n-1$ MOLS reda n.

Dokaz.

- Pretpostavimo da je S uniforman $(n \times n)/(n-1)$ semi-latinski kvadrat. Tada je $\mu(S) = 1$.
- Svake dvije pozicije od S koje su različite se sijeku u 0 ili 1 točki tj. svaki par različitih simbola od S se javlja na istoj poziciji najviše jednom.
- Bailey je pokazala da $(n \times n)/(n-1)$ semi-latinski kvadrat sa ovim svojstvom mora biti superpozicija od $n-1$ MOLS reda n.

- Obratno, pretpostavimo da je S superpozicija od $n-1$ MOLS reda n.
- Promatramo pozicije $S(i,j)$ i $S(i',j')$ za koje vrijedi da je $i \neq i'$ i $j \neq j'$.
- Tada vrijedi $|S(i,j) \cap S(i',j')| \leq 1$ jer u suprotnom bi postojala dva ili više simbola iz semi-latinskog kvadrata koja bi se javila zajedno u više od jedne pozicije.
- Prema tome, svaki od $n-1$ simbola od $S(i,j)$ se mora javiti u retku i',
- nikoja dva se ne mogu javiti zajedno u niti jednom drugom retku i niti jedan se ne može javiti u stupcu j.
- Stoga vrijedi $|S(i,j) \cap S(i',j')| = 1$

■

- Uniformne semi-latinske kvadrate možemo promatrati kao generalizacija potpunog skupa MOLS
- Budući da je $\mu-$terostruka inflacija uniformnog semi-latinskog kvadrata uniformna, egzistencija uniformnog $(n \times n)/((n-1)\mu)$ semi-latinskog kvadrata za svaki $\mu > 0$ je ekvivalentna egzistenciji potpunog skupa MOLS reda n. Takav skup postoji ako je n prost broj.

Statistička važnost semi-latinskih kvadrata

Teorem

Neka je S uniforman $(n \times n)/k$ semi-latinski kvadrat. Tada je S Schur-optimalan. To znači da je $\Delta(S)$ Schur-optimalan u klasi osnovnih dizajna $(n \times n)/k$ semi-latinskih kvadrata.

R.A.Bailey, Design of Comparative Experiments, Cambridge University Press, 2008.

C.C.Lindner, C.A. Rodger, Design Theory, CRC Press, 2009.
P.Cameron, Permutation Groups, Cambridge University Press, 1999.