Ανάλυση διακύμανσης (Μέρος 3ο)

31/3/2017

a x b Παραγοντικό Πείραμα (1)

- Όταν θέλουμε να μελετήσουμε την επίδραση (στη μεταβλητή απόφασης) δύο παραγόντων, έστω A και B, με ***a*** στάθμες ο A και ***b*** στάθμες ο B, μπορούμε να σχεδιάσουμε ένα ***a* x b** παραγοντικό πείραμα
- Σε ένα τέτοιο πείραμα, κάθε στάθμη του ενός παράγοντα συνδυάζεται με κάθε στάθμη του άλλου παράγοντα, δηλαδή, συνολικά έχουμε ***a · b*** διαφορετικές επεμβάσεις (διαφορετικούς συνδυασμούς από δύο στάθμες, μία για κάθε παράγοντα)

a x b Παραγοντικό Πείραμα (2)

- Θεωρούμε ότι για κάθε επέμβαση παίρνουμε ίδιο αριθμό παρατηρήσεων, έστω $\boldsymbol{r}$, και επομένως ότι έχουμε $\boldsymbol{a \cdot b \cdot r}$ παρατηρήσεις
- Αν $\boldsymbol{r > 1}$, πέραν της επίδρασης του παράγοντα A και της επίδρασης του παράγοντα B, μπορούμε να ελέγξουμε και αν, στη μεταβλητή απόκρισης υπάρχει στατιστικά σημαντική επίδραση που να οφείλεται στην αλληλεπίδραση των παραγόντων A και B

a x b Παραγοντικό Πείραμα (3)

- Αν μπορούμε να υποθέσουμε ότι οι δύο παράγοντες Α και Β επιδρούν στη μεταβλητή απόκρισης **ανεξάρτητα**, σχεδιάζουμε και εκτελούμε ένα ***a* x *b*** παραγοντικό πείραμα με μία παρατήρηση για κάθε επέμβαση (***r* = 1**)

- Αν μπορούμε να υποθέσουμε ότι οι δύο παράγοντες Α και Β δεν επιδρούν στη μεταβλητή απόκρισης ανεξάρτητα, δηλαδή υπάρχει επίδραση του ενός παράγοντα στη συμπεριφορά του άλλου, σχεδιάζουμε και εκτελούμε ένα ***a* x *b*** παραγοντικό πείραμα με περισσότερες από μία παρατηρήσεις για κάθε επέμβαση (***r* > 1**)

a x b Παραγοντικό Πείραμα (4)

- Οι έλεγχοι που πραγματοποιούνται για ***r* > 1** είναι οι εξής:
  - Για την αλληλεπίδραση των <u>δύο παραγόντων A και B</u>
    - $\mathbf{H_{0\gamma}}$: οι παράγοντες A και B δεν αλληλεπιδρούν
    - $\mathbf{H_{1\gamma}}$: οι παράγοντες A και B αλληλεπιδρούν
  - Για την κύρια επίδραση του <u>παράγοντα A</u>
    - $\mathbf{H_{0a}}$: ο παράγοντας A δεν επιδρά, ή αλλιώς, δεν υπάρχουν διαφορές μεταξύ των **a** μέσων του παράγοντα A
    - $\mathbf{H_{1a}}$: ο παράγοντας A επιδρά, ή αλλιώς, τουλάχιστον δύο από τους **a** μέσους του παράγοντα A διαφέρουν
  - Για την κύρια επίδραση του <u>παράγοντα B</u>
    - $\mathbf{H_{0b}}$: ο παράγοντας B δεν επιδρά, ή αλλιώς, δεν υπάρχουν διαφορές μεταξύ των **b** μέσων του παράγοντα B
    - $\mathbf{H_{1b}}$: ο παράγοντας B επιδρά, ή αλλιώς, τουλάχιστον δύο από τους **b** μέσους του παράγοντα B διαφέρουν

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με r>1 παρατηρήσεις ανά παρέμβαση (1)

Πηγή μεταβλητότητας	Β.Ε.	Άθροισμα τετραγώνων SS	Μέσο άθροισμα τετραγώνων MS	Κριτήριο F	Περιοχή απόρριψης
Παράγοντας A	a-1	$SSA = \sum_{i=1}^{a} \frac{A_i^2}{br} - \frac{G^2}{n}$	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MSA}{MSE}$	$F_A \geq F_{a-1;ab(r-1);a}$
Παράγοντας B	b-1	$SSB = \sum_{j=1}^{b} \frac{B_j^2}{ar} - \frac{G^2}{n}$	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$	$F_B \geq F_{b-1;ab(r-1);a}$
Αλληλεπίδραση A και B	(a-1)(b-1)	$SSA(AB) = \sum_{ij} \frac{(AB)_{ij}^2}{r} - \frac{G^2}{n} - SSA - SSB$	$MS(AB) = \frac{SS(AB)}{(a-1)(b-1)}$	$F_{AB} = \frac{MS(AB)}{MSE}$	$F_{AB} \geq F_{(a-1)(b-1);ab(r-1);a}$
Σφάλμα ή Εντός των δειγμάτων	ab(r-1)	$SSE = SST_{ot} - SSA - SSB - SS(AB)$	$MSE = \frac{SSE}{ab(r-1)}$		
Ολική	abr-1	$SST_{ot} = SSA + SSB + SS(AB) + SSE$			

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με r>1 παρατηρήσεις ανά παρέμβαση (2)

- **όπου**
  - ***a* οι στάθμες του παράγοντα A και *b* οι στάθμες του παράγοντα B**
  - $G = \sum_{ijh} y_{ijh}$, **το γενικό άθροισμα όλων των *a* · *b* · *r* παρατηρήσεων**
  - $A_i, i = 1, 2, \ldots k$ **το άθροισμα όλων των παρατηρήσεων στην *i* στάθμη του παράγοντα A**
  - $\mathrm{B}_j, j = 1, 2, \ldots b$ **το άθροισμα όλων των παρατηρήσεων στην *j* στάθμη του παράγοντα B**
  - $(\mathrm{AB})_{ij}$ **το άθροισμα των *r* παρατηρήσεων που πήραμε από την i στάθμη του παράγοντα A και την j στάθμη του παράγοντα B**
  - $SSTot = \sum_{ijh} y_{ijh}^2 - \frac{G^2}{n}$
  - **Μεταβλητότητα μεταξύ των κελιών**
    - $SSCells = \sum_{ij} \frac{(AB)_{ij}^2}{r} - \frac{G^2}{n} = SSA + SSB + SS(AB)$
  - $SSTot = SSCells + SSE = SSA + SSB + SS(AB) + SSE$

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με r>1 παρατηρήσεις ανά παρέμβαση (3)

- Υποθέσεις
  - Τα $\boldsymbol{a \cdot b}$ δείγματα είναι ανεξάρτητα τυχαία δείγματα (όλες οι παρατηρήσεις $\boldsymbol{y_{ijh}}$ είναι μεταξύ τους ανεξάρτητες)
  - Κάθε δείγμα προέρχεται από κανονικό πληθυσμό
  - Οι $\boldsymbol{a \cdot b}$ πληθυσμοί έχουν κοινή διακύμανση $\boldsymbol{\sigma^2}$ (ομοσκεδαστικότητα)

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με r>1 παρατηρήσεις ανά παρέμβαση (4)

- Αν $\boldsymbol{\varepsilon_{ijh}}$ η απόκλιση της $\boldsymbol{Y_{ijh}}$ από τη μέση τιμή της $\boldsymbol{\mu_{ij}}$, του πληθυσμού ***ij***, τότε οι προηγούμενες υποθέσεις είναι ισοδύναμες με τις υποθέσεις ότι
  - τα **σφάλματα** (**θεωρητικά υπόλοιπα**) $\boldsymbol{\varepsilon_{ijh}}$ είναι μεταξύ τους ανεξάρτητα
  - ακολουθούν κανονικές κατανομές με μέση τιμή **0** και κοινή διακύμανση $\boldsymbol{\sigma^2}$, δηλαδή ότι $\boldsymbol{\varepsilon_{ijh} \sim N(0, \sigma^2)}$

- Οι υποθέσεις αυτές ελέγχονται όπως στο εντελώς τυχαιοποιημένο σχέδιο

a x b Παραγοντικό Πείραμα (5)

- Οι έλεγχοι που πραγματοποιούνται για ***r* = 1** είναι οι εξής:
  - Για την κύρια επίδραση του <u>παράγοντα A</u>
    - **H_{0a}**: ο παράγοντας A δεν επιδρά, ή αλλιώς, δεν υπάρχουν διαφορές μεταξύ των **a** μέσων του παράγοντα A
    - **H_{1a}**: ο παράγοντας A επιδρά, ή αλλιώς, τουλάχιστον δύο από τους **a** μέσους του παράγοντα A διαφέρουν
  - Για την κύρια επίδραση του <u>παράγοντα B</u>
    - **H_{0b}**: ο παράγοντας B δεν επιδρά, ή αλλιώς, δεν υπάρχουν διαφορές μεταξύ των **b** μέσων του παράγοντα B
    - **H_{1b}**: ο παράγοντας B επιδρά, ή αλλιώς, τουλάχιστον δύο από τους **b** μέσους του παράγοντα B διαφέρουν

<u>Σημείωση:</u> κάνουμε την παραδοχή ότι δεν υπάρχει επίδραση στη μεταβλητή απόκρισης που να οφείλεται στην αλληλεπίδραση των δύο παραγόντων

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με μια παρατήρηση (r=1) ανά παρέμβαση (1)

Πηγή μεταβλητότητας	Β.Ε.	Άθροισμα τετραγώνων SS	Μέσο άθροισμα τετραγώνων MS	Κριτήριο F	Περιοχή απόρριψης
Παράγοντας A	a-1	$SSA = \sum_{i=1}^{a} \frac{A_i^2}{b} - \frac{G^2}{n}$	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MSA}{MSE}$	$F_A \geq F_{a-1;(a-1)(b-1);a}$
Παράγοντας B	b-1	$SSB = \sum_{j=1}^{b} \frac{B_j^2}{a} - \frac{G^2}{n}$	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$	$F_B \geq F_{b-1;(a-1)(b-1);a}$
Σφάλμα ή Εντός των δειγμάτων	(a-1)(b-1)	$SSE = SST_{ot} - SSA - SSB$	$MSE = \frac{SSE}{(a-1)(b-1)}$		
Ολική	ab-1	$SST_{ot} = SSA + SSB + SSE$			

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με μια παρατήρηση (r=1) ανά παρέμβαση (2)

- **όπου**
  - ***a* οι στάθμες του παράγοντα A και *b* οι στάθμες του παράγοντα B**
  - $G = \sum_{ij} y_{ij}$, **το γενικό άθροισμα όλων των *a* · *b* παρατηρήσεων**
  - $A_i, i = 1, 2, \ldots k$ **το άθροισμα όλων των παρατηρήσεων στην *i* στάθμη του παράγοντα A**
  - $\mathrm{B}_j, j = 1, 2, \ldots b$ **το άθροισμα όλων των παρατηρήσεων στην *j* στάθμη του παράγοντα B**
  - $SSTot = \sum_{ij} y_{ij}^2 - \frac{G^2}{n}$
  - $SSE = SSTot - SSA - SSB$

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με μια παρατήρηση (r=1) ανά παρέμβαση (3)

- Υποθέσεις
  - Δεν υπάρχει αλληλεπίδραση των δύο παραγόντων
  - Τα $\boldsymbol{a} \cdot \boldsymbol{b}$ δείγματα είναι ανεξάρτητα τυχαία δείγματα (όλες οι παρατηρήσεις $\boldsymbol{y_{ij}}$ είναι μεταξύ τους ανεξάρτητες)
  - Κάθε δείγμα προέρχεται από κανονικό πληθυσμό
  - Οι $\boldsymbol{a} \cdot \boldsymbol{b}$ πληθυσμοί έχουν κοινή διακύμανση $\boldsymbol{\sigma^2}$ (ομοσκεδαστικότητα)

Έλεγχος ανάλυσης διακύμανσης για το a x b παραγοντικό πείραμα με μια παρατήρηση (r=1) ανά παρέμβαση (4)

- Αν $\boldsymbol{\varepsilon_{ij}}$ η απόκλιση της $\boldsymbol{Y_{ij}}$ από τη μέση τιμή της $\boldsymbol{\mu_{ij}}$, του πληθυσμού ***ij***, τότε οι προηγούμενες υποθέσεις είναι ισοδύναμες με τις υποθέσεις ότι
  - τα **σφάλματα** (**θεωρητικά υπόλοιπα**) $\boldsymbol{\varepsilon_{ij}}$ είναι μεταξύ τους ανεξάρτητα
  - ακολουθούν κανονικές κατανομές με μέση τιμή **0** και κοινή διακύμανση $\boldsymbol{\sigma^2}$, δηλαδή ότι $\boldsymbol{\varepsilon_{ij} \sim N\left(0, \sigma^2\right)}$

- Οι υποθέσεις αυτές ελέγχονται όπως στο εντελώς τυχαιοποιημένο σχέδιο

Παράδειγμα 1

- Ένας ερευνητής ενδιαφέρεται να διερευνήσει αν και πώς οι δύο παράγοντες, θερμοκρασία, έστω Α (με τρεις στάθμες A_1, A_2 και A_3, αντίστοιχα, 10°C, 20°C και 30°C) και ένταση φωτισμού, έστω Β (με τρεις στάθμες B_1, B_2 και B_3) επιδρούν στην ανάπτυξη του συγκεκριμένου είδους φυτών, σε επίπεδο σημαντικότητας 5%.

Ένταση φωτισμού (παράγοντας Β)	Επίπεδο θερμοκρασίας (παράγοντας Α)			
		A1	**A2**	**A3**
	B1	310	480	400
	B2	320	540	340
	B3	410	550	250

Παράδειγμα 2

- Ένας ερευνητής για να διερευνήσει αν επηρεάζεται το αποτέλεσμα μιας χημικής αντίδρασης από το είδος του καταλύτη και από τη θερμοκρασία, καθόρισε τέσσερις στάθμες καταλύτη (A_1, A_2, A_3 και A_4) και τρεις στάθμες θερμοκρασίας (B_1, B_2, και B_3) και για κάθε συνδυασμό καταλύτη – θερμοκρασίας, εκτέλεσε το πείραμα μία φορά
- Θεωρώντας ότι οι δύο παράγοντες δεν αλληλεπιδρούν θα κάνει τον έλεγχο σε επίπεδο σημαντικότητας 5%

		Καταλύτης (παράγοντας Α)			
Θερμοκρασία (παράγοντας Β)		**A1**	**A2**	**A3**	**A4**
	B1	53	59	58	50
	B2	57	65	62	60
	B3	52	62	54	52