

Statistikë e aplikuar



UNIVERSITETI I EJL
ЈНЕ УНИВЕРЗИТЕТ
SEE UNIVERSITY

Disenji eksperimental dhe analiza e variancës

Faton Berisha

Kapitulli 10

Disenji eksperimental dhe analiza e variancës



Disenji eksperimental dhe analiza e variancës

- 10.1 Nocionet themelore të disenjit eksperimental
- 10.2 Analiza e njëanshme e variancës
- 10.3 Disenji i bloqeve të rastësishme
- 10.4 Analiza e dyanshme e variancës

10.1 Nocionet themelore të disenjit eksperimental

- Në shumë studime statistikore identifikohet një ndryshore me interes: **ndryshorja e varur**.
- Mbledhen të dhëna që tregojnë se si një ose më tepër faktorë (**ndryshoret e pavarura**) ndikojnë në ndryshoren e varur.
- Në qoftë se mund të kontrollohen vlerat e ndryshoreve të pavarura, themi se të dhënat janë **eksperimentale** (përkundër të dhënave **obsrevuese**).
- Vlerat e të dhënave eksperimentale quhen **trajtime**.
- Për të mbledhur të dhëna në një eksperiment, trajtime të ndryshme u shoqërohen objekteve të ndryshme (njerëz, vetura...): **njësi eksperimentale**.

Nocionet themelore të disenjit eksperimental (Vazhdim)

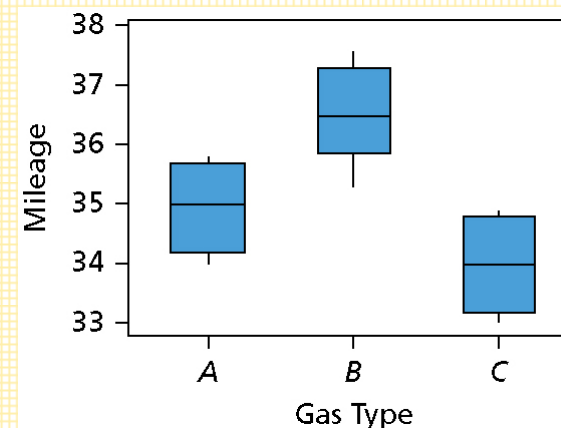
- Kur analisti kontrollon trajtimet e zbatuara dhe mënyrën si u shoqërohen njësive eksperimentale kemi të bëjmë me: **eksperiment të disenjuar.**
- Në një **disenj eksperimenti plotësisht të rastësishëm** trajtimeve u shoqërohen mostra të rastësishme të pavarura.

Shembull: Rasti i kilometrazhit të veturave

Qëllimi: Të krahasohet dhe vlerësohet efekti i **trjatimeve** të ndryshme në **ndryshoren e varur**.

Example 10.1 A ndryshon MPG e veturave me tipin e karburantit?

<u>Tipi A</u>	<u>Tipi B</u>	<u>Tipi C</u>
$x_{A1}=34.0$	$x_{B1}=35.3$	$x_{C1}=33.3$
$x_{A2}=35.0$	$x_{B2}=36.5$	$x_{C2}=34.0$
$x_{A3}=34.3$	$x_{B3}=36.4$	$x_{C3}=34.7$
$x_{A4}=35.5$	$x_{B4}=37.0$	$x_{C4}=33.0$
$x_{A5}=35.8$	$x_{B5}=37.6$	$x_{C5}=34.9$



Ndryshorja e varur: Shpenzimi i veturave (në MPG)
Trajtimet: Karburant i tipeve – A, B, C

10.2 Analiza e njëanshme e variancës

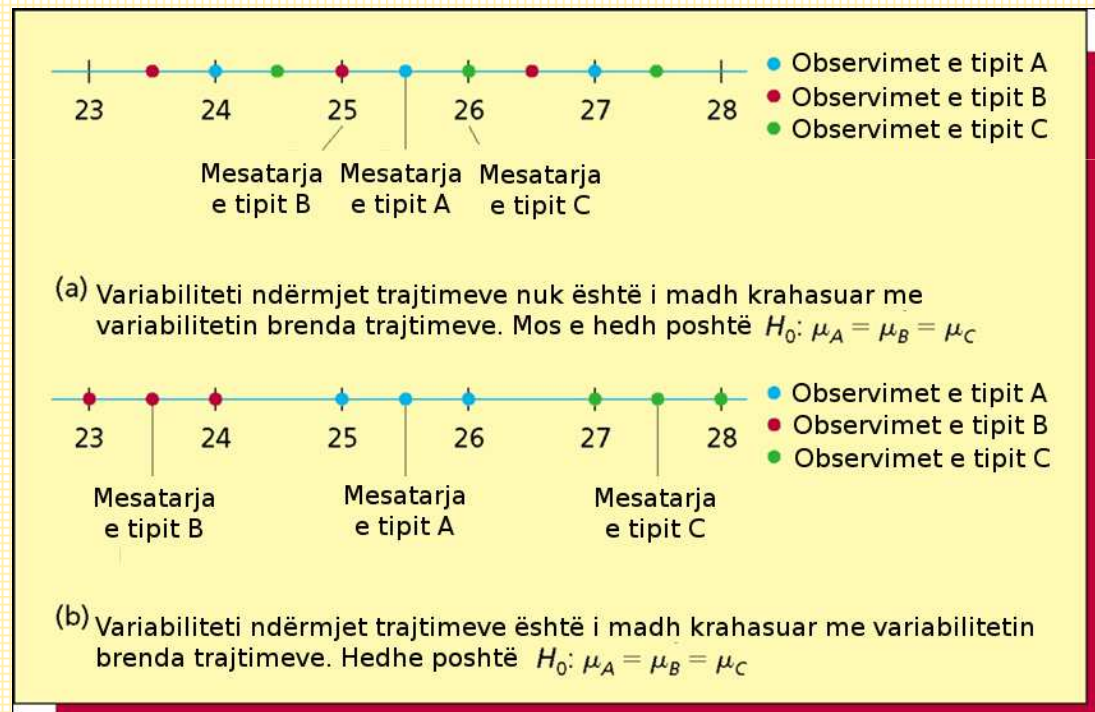
A ka ndryshime në mesataret $\mu_1, \mu_2, \dots, \mu_p$ e ndryshores së varur të shoqëruar me p trajtime?

$$H_0: \mu_1 = \mu_2 = \dots = \mu_p$$

H_a : Së paku dy nga

$\mu_1, \mu_2, \dots, \mu_p$ ndryshojnë

Ose, a është variabiliteti ndërmjet trajtimeve i madh krahasuar me variabilitetin brenda trajtimeve?



Ndarja e variabilitetit total në ndryshoren e varur

Variabiliteti total = Variabilitetit ndërmjet trajtimeve + Variabiliteti brenda trajtimeve

Shuma totale e katrorëve = Shuma e katrorëve të trajtimeve + Shuma e katrorëve të gabimeve

SSTO = SST + SSE

$$\sum_{i=1}^p \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2 = \sum_{i=1}^p n_i (\bar{x}_i - \bar{x})^2 + \sum_{i=1}^p \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

F për ndryshim ndërmjet mesatareve të trajtimeve

$H_0: \mu_1 = \mu_2 = \dots = \mu_p$ (nuk ka efekt të trajtimeve)

H_a : Së paku dy nga $\mu_1, \mu_2, \dots, \mu_p$ ndryshojnë

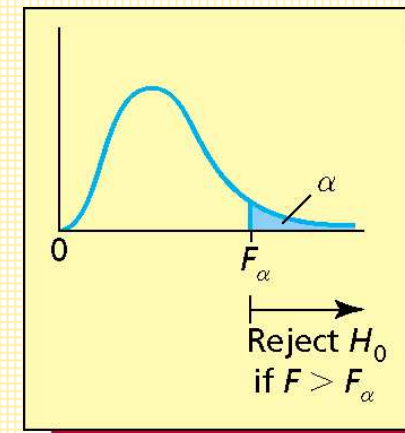
Statistika e testit:

$$F = \frac{MST}{MSE} = \frac{SST/(p-1)}{SSE/(n-p)}$$

Hedh poshtë H_0 në qoftë se

$$F > F_\alpha \quad \text{ose}$$

$$p\text{-vlera} < \alpha$$



F_α mbështetet në $p-1$ mbi $n-p$ shkallë lirie.

Tabela e analizës së njëanshme të variancës

Burimi	Shkallët e lirisë	Shuma e katrorëve	Mesatarja e katrorëve	F statistika
Treatments	$p-1$	SST	$MST = \frac{SST}{p-1}$	$F = \frac{MST}{MSE}$
Error	$n-p$	SSE	$MSE = \frac{SSE}{n-p}$	
Total	$n-1$	SSTO		

Shembulli 10.5 Rasti i kilometrazhit të veturave

<i>Burimi i variacionit</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p-vlera</i>	<i>F krit</i>
Ndërmjet grupeve	17.0493	2	8.5247	12.7424	0.0011	3.8853
Brenda grupeve	8.0280	12	0.6690			
Totali	25.0773	14				

Krahasimi në çifte, Intervalet individuale

Intervali individual i $100(1 - \alpha)\%$ besueshmërisë për $\mu_i - \mu_h$

$$(\bar{x}_i - \bar{x}_h) \pm t_{\alpha/2} \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_h} \right)} \quad t_{\alpha/2} \text{ mbështetet në } n-p \text{ shkallë lirie.}$$

Shembull 10.6 Rasti i kilometrazhit të veturave
(A kundrejt B , $\alpha = 0.05$)

$$(34.92 - 36.56) \pm 2.179 \sqrt{0.669 \left(\frac{1}{5} + \frac{1}{5} \right)} = -1.64 \pm 1.127 \\ = [-2.7671, -0.513]$$

Grupet	Nurmi	Mesatarja	Varianca	MSE
Tipi A	5	34.92	0.587	0.669
Tipi B	5	36.56	0.723	<i>n-p</i>
Tipi C	5	33.98	0.697	12

Krahasimi në çifte, Intervalet simultane të Tukey

Intervalet simultane të Tukey të $100(1 - \alpha)\%$ besueshmërisë për $\mu_i - \mu_h$

$$(\bar{x}_i - \bar{x}_h) \pm q_\alpha \sqrt{\frac{MSE}{m}}$$

q_α është pika e sipërme e α përqind të rangut të studentit me p dhe $(n-p)$ nga Tabela A.9.

m është madhësia e përbashkët e mostrës.

Shembull 10.6 Rasti i kilometrazhit të veturave
(A kundrejt B, $\alpha = 0.05$)

$$(34.92 - 36.56) \pm 3.77 \sqrt{\frac{0.669}{5}} = -1.64 \pm 1.379$$
$$= [-3.019, -0.261] \text{ (A v B)}$$

Grupet	Nurmi	Mesatarja	Varianca	MSE
Tipi A	5	34.92	0.587	0.669
Tipi B	5	36.56	0.723	$n-p$
Tipi C	5	33.98	0.697	12

$$[-0.439, 2.319] \text{ (A v C)}$$

$$[1.201, 3.959] \text{ (B v C)}$$

Vlerësimi i mesatareve individuale të trajtimeve

Intervali i $100(1 - \alpha)\%$ besueshmërisë për μ_i

$$\bar{x}_i \pm t_{\alpha/2} \sqrt{\frac{MSE}{n_i}}$$

$t_{\alpha/2}$ mbështetet në $n - p$ shkallë lirie.

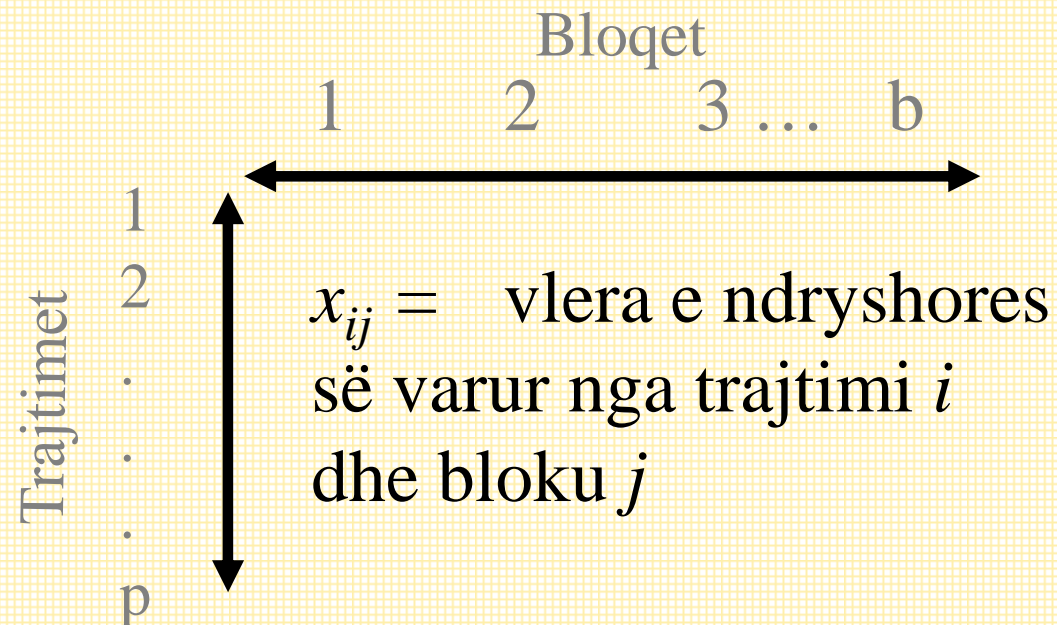
Shembull 10.6 Rasti i kilometrazhit të veturave
(Tipi B , $\alpha = 0.05$)

$$36.56 \pm 2.179 \sqrt{\frac{0.669}{5}} = 36.56 \pm 0.797 = [35.763, 37.357]$$

Grupet	Nurmi	Mesatarja	Varianca	MSE
Tipi A	5	34.92	0.587	0.669
Tipi B	5	36.56	0.723	$n-p$
Tipi C	5	33.98	0.697	12

10.3 Blok disenji i rastësishëm

Një **blok disenj i rastësishëm** krahason p trajtime (p.sh., metoda prodhimi) në b bloqe (ose njësi eksperimentale; p.sh., operatorë makine). Si përgjithësim i disenjit të diferencave të çiftëzuara, ky disenj kontrollon variabilitetin në çdo njësi eksperimentale duke krahasuar secilin trajtim në të njëjtat njësi eksperimentale (jo të pavarura).



Shembull: Blok disenji i rastësishëm

Shembull 10.7

Rasti i kutisë defektive nga kartoni

	Bloku			
Trajtimi	Operatori i makinës			
Met. prodh.	1	2	3	Mes. trajt.
1	9	10	12	10.3333
2	8	11	12	10.3333
3	3	5	7	5.0000
4	4	5	5	4.6667
Mes. bloq.	6.00	7.75	9.00	7.5833

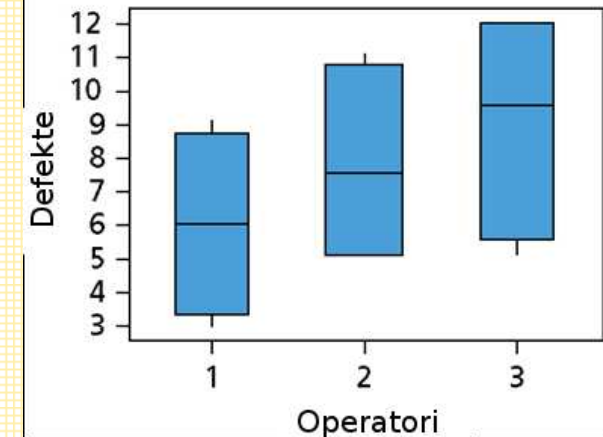
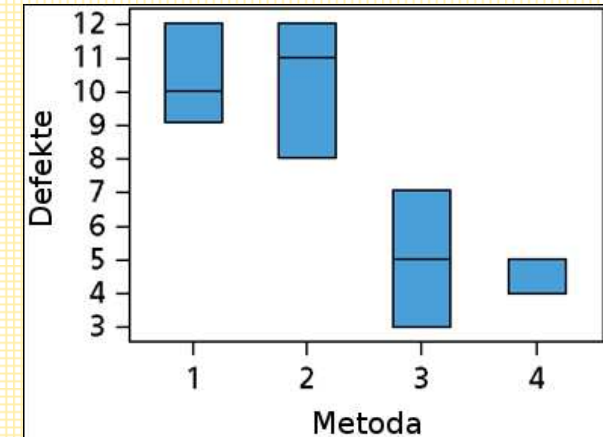


Tabela ANOVA për disenjin e bloqeve të rastësishme

Burimi	Shkallë lirie	Shuma e katrorëve	Mesatarja e katrorëve	F statistika
Trajtime	$p-1$	SST	$MST = \frac{SST}{p-1}$	$F(\text{trt}) = \frac{MST}{MSE}$
Bloqe	$b-1$	SSB	$MSB = \frac{SSB}{b-1}$	$F(\text{blk}) = \frac{MSB}{MSE}$
Gabimi	$(p-1)(b-1)$	SSE	$MSE = \frac{SSE}{(p-1)(b-1)}$	
Totali	$pb - 1$	SSTO		

F testi për efektet e trajtimeve

H_0 : Nuk ka dallime ndërmjet efekteve të trajtimeve

H_a : Së paku dy efekte të trajtimeve dallojnë

Statistika e testit:

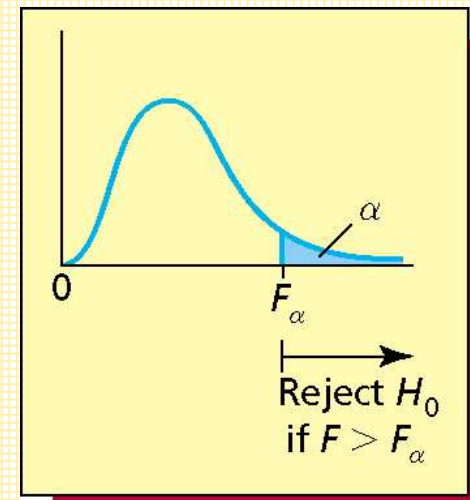
$$F = \frac{MST}{MSE} = \frac{SST/(p - 1)}{SSE/[(p - 1)(b - 1)]}$$

Hedh poshtë H_0 në qoftë se

$$F > F_{\alpha} \quad \text{ose}$$

$$p\text{-vlere} < \alpha$$

F_{α} mbështetet në $p-1$ mbi $(p-1)(b-1)$ shkallë lirie.



F testi për efektet e bloqeve

H_0 : Nuk ka dallim ndërmjet efekteve të bloqeve

H_a : Së paku dy efekte të bloqeve dallojnë

Statistika e testit:

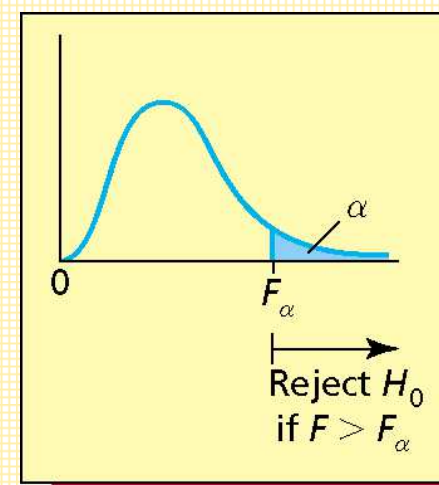
$$F = \frac{MSB}{MSE} = \frac{SSB/(p - 1)}{SSE/[(p - 1)(b - 1)]}$$

Hedh poshtë H_0 në qoftë se

$$F > F_\alpha \quad \text{ose}$$

$$p\text{-vlera} < \alpha$$

F_α mbështetet në $b-1$ mbi $(p-1)(b-1)$ shkallë lirie.



Shebull: ANOVA me bloqe të rastësishme

Shembull 10.7. Rasti i kutive defektive nga kartoni

Output i Megastat

ANOVA table

Source	SS	df	MS	F	p-value
Treatments	18.17	2	9.083	14.22	.0053
Blocks	90.92	3	30.306	47.43	.0001
Error	3.83	6	0.639		
Total	112.92	11			

	Bloku			
Trajtimi	Operatori i makinës			
Met. prodh.	1	2	3	Mes. trajt.
1	9	10	12	10.3333
2	8	11	12	10.3333
3	3	5	7	5.0000
4	4	5	5	4.6667
Mes. bloq.	6.00	7.75	9.00	7.5833

$$s = \sqrt{MSE}$$

$$= \sqrt{0.639} = 0.7994$$

Vlerësimi i diferencave të trajtimeve nën bloqet e rastësishme. Intervalet individuale

Intervalet individuale të $100(1 - \alpha)\%$ besueshmërisë për $\mu_{i\bullet} - \mu_{h\bullet}$

$$(\bar{x}_{i\bullet} - \bar{x}_{h\bullet}) \pm t_{\alpha/2} s \sqrt{\frac{2}{b}}$$

$t_{\alpha/2}$ mbështetet mbi $(p-1)(b-1)$ shkallë lirie.

Shembull 10.8 Rasti i kutive defektive nga kartoni (4 kundrejt 1)

$$(4.6667 - 10.3333) \pm 2.447(0.7994) \sqrt{\frac{2}{3}} = -5.6666 \pm 1.5971 \\ = [-7.2637, -4.0695]$$

	Bloku			
Trajtimi	Operatori i makinës			
Met. prodh.	1	2	3	Mes. trajt.
1	9	10	12	10.3333
2	8	11	12	10.3333
3	3	5	7	5.0000
4	4	5	5	4.6667
Mes. bloq.	6.00	7.75	9.00	7.5833

$t_{.025}$ me $(3-1)(4-1) = 6$ shkallë lirie.

Vlerësimi i diferencave të trajtimeve nën bloqet e rastësishme. Intervalet simultane

Intervalet simultane të Tukey të $100(1 - \alpha)\%$ besueshmërisë për $\mu_{i\cdot} - \mu_{h\cdot}$

$$(\bar{x}_{i\cdot} - \bar{x}_{h\cdot}) \pm q_{\alpha} \frac{s}{\sqrt{b}}$$

q_{α} është pika e sipërme e α përqindjes i rangut të studentit për p dhe $(p-1)(b-1)$ nga tabela A.9.

Shembull 10.8 Rasti i kutive defektive nga kartoni (4 kundrejt 1)

$$(4.6667 - 10.3333) \pm 4.90 \frac{0.7994}{\sqrt{3}} = -5.6666 \pm 2.2621$$

$$= [-7.9278, -3.4054]$$

	Bloku			
Trajtimi	Operatori i makinës			
Met. prodh.	1	2	3	Mes. trajt.
1	9	10	12	10.3333
2	8	11	12	10.3333
3	3	5	7	5.0000
4	4	5	5	4.6667
Mes. bloq.	6.00	7.75	9.00	7.5833

$q_{.05}$ për 4 dhe 6.