

4.2 Үлкен іріктеулер үшін корреляция коэффициентін есептеу

Мақсаты. Үлкен іріктеулер үшін корреляция коэффициентін есептеу тәсілдерімен танысу.

Корреляция коэффициентін есептеу және анықтау тәсілдері үлкен және кіші іріктеулерде сандық белгілер үшін бір бірінен айтарлықтай ажыратылады. Іріктеу мәліметтерінің ұйымдастыру формасы, яғни $n > 30$, корреляциялық тор болып табылады, бұндағы варианттардың айырмасы екі белгіні есепке алумен жүргізіледі (x, y) . Варианттардың ыдырау және тығыздық дәрежелері бойынша вариациялық кластардың торлары арқылы корреляция коэффициентінің шамасы үлкен, орташа және кіші болатындығын алдын ала болжауға болады.

Егер варианттар тар эллипс түзсе, онда байланыс үлкен болады. Егер варианттар торлардың көбінде орналасса, онда байланыс жоқ немесе мағынасыз.

Корреляциялық торды құрғанда әрбір өзгермелі белгіге қажет: 1) класс аралық шаманы және класс шекараларын үйлестіріп жинап, класс санын анықтау; 2) бір белгінің кластарын корреляциялық тордың сол жақ ұшына дейін жоғарыдан төменге қарай, ал екінші белгінің класын – жоғарғы жағы бойынша солдан оңға қарай үйлестіру. Кластар сызықпен бөлінген. Көлденеңінен және тігінен сызықтар бір бірімен түйісіп, корреляциялық тордың ұяшығын түзеді.

Үлкен іріктеулер бойынша сандық белгілердің арасындағы корреляция коэффициентін нақты есептеу үшін мына формуланы қолданамыз:

$$r = \frac{\sum p a_x a_y - (b_x b_y n)}{\sigma_x \sigma_y n} \quad (38)$$

Корреляция коэффициентін есептеуде класс аралық шамасындағы орташа квадраттық ауытқу алынып, мына формуламен шығарылады:

$$\sigma = \pm \sqrt{\frac{\sum p a^2}{n} - b^2} \quad (39)$$

Корреляция коэффициентін есептеу үшін, корреляциялық торға көлденеңінен төрт және тігінен төрт графалар қосу қажет. Бір көлденең және тігінен біреу графалар класс варианттарының бойына жасалынады; олар шартты ортадан (а) класс ауытқуларын жазу үшін керек. Қалған үш көлденең және үш тігінен графалар тордың төменгі жағында (у қатары үшін) және оның оң жағында (х қатары үшін) орналасады. Бұнда жазылады: жиіліктер (р), ауытқитын жиілік туындылары (ра), квадраттық ауытқу жиілік туындылары (ра²). Тік графалардың басына біртіндеп: $p_x, p_x a_x, p_x a_x^2$, ал көлденең графа басына (сол жақ) $p_y, p_y a_y, p_y a_y^2$ жазылады.

Сосын қосылған графаларды есептеу және толтыру жүргізіледі.

Мысал. Қызыл горбатов тұқымы сиырларының тірі салмағы (х) және төс ұстамы (у) арасындағы корреляция коэффициентін 4.2.1 кестесінде берілген мәліметтер бойынша есептеу қажет.

Кесте 4.2.1 Қызыл горбатов тұқымы сиырларының тірі салмағы және төс ұстамы

№	х, кг	у, см	№	х, кг	у, см	№	х, кг	у, см	№	х, кг	у, см
1	489	184	26	440	185	51	445	180	76	450	180
2	467	186	27	524	194	52	420	179	77	450	181
3	462	185	28	447	179	53	491	191	78	387	171
4	441	182	29	430	179	54	450	181	79	374	171
5	473	186	30	485	187	55	406	172	80	360	167
6	491	190	31	440	173	56	417	174	81	545	191
7	545	196	32	488	187	57	434	177	82	454	184
8	433	183	33	439	180	58	432	177	83	467	186
9	488	191	34	445	180	59	505	193	84	454	184
10	539	196	35	504	191	60	534	195	85	441	178
11	440	182	36	550	203	61	473	188	86	434	180
12	475	186	37	495	190	62	441	179	87	519	192
13	411	178	38	536	192	63	556	197	88	488	187
14	488	189	39	426	186	64	486	185	89	441	178
15	426	177	40	388	169	65	535	196	90	456	180
16	390	172	41	407	176	66	460	183	91	400	173
17	482	192	42	425	184	67	469	180	92	420	178
18	391	174	43	390	172	68	421	180	93	403	172
19	470	185	44	418	179	69	520	195	94	390	170
20	421	180	45	465	189	70	445	183	95	442	179
21	429	176	46	391	174	71	384	173	96	445	177
22	439	180	47	365	167	72	488	182	97	429	176
23	442	183	48	383	172	73	500	190	98	425	176
24	490	189	49	427	183	74	432	178	99	457	182
25	426	177	50	448	186	75	475	185	100	493	185

Көкірек өлшемінің лимиті 167-203 см тең, ал тірі салмақ лимиті - 365-559 кг. Бірінші белгі үшін класс аралық шама $K_y = \frac{203-167}{12} \approx 4,0$ тең. Екінші белгі үшін класс аралық шама $K_x = \frac{559-365}{12} \approx 20$ тең.

Осы алынған сандардан, бірінші кластың шекарасының басталуы көкірек өлшемі бойынша 166, ал кластың барлығы – 10. Бірінші кластың шекарасының басталуы трі салмағы бойынша 360, ал кластың барлығы – 10. Корреляциялық торды құру қажет.

Кесте 4.2.2 Корреляциялық тордың формасы

	у-төс ұстамы
--	--------------

х-тірі салмақ	166-169	170-173	174-177	178-181	182-185	186-189	190-193	194-197	198-201	202-205
360-379										
380-399										
400-419										
420-439										
440-459										
460-479										
480-499										
500-519										
520-539										
540-559										

Корреляциялық тор сызылғаннан кейін осы тордың ұяшықтарын жиіліктермен толтырамыз. Бұл варианттар тасымалы вариациялық қатардағы кластары бойынша жүргізіледі, тек бұл жағдайда әрбір особь бойынша бір көрсеткішке емес, екі көрсеткішке назар аударамыз.

Варианттар тасымалынан кейін корреляциялық тор мынадай түрде болады (кесте 4.2.3).

Корреляциялық тордың ұяшықтары бойынша алынған варианттардың орналасуы, тірі салмақ пен төс ұстамы арасындағы тікелей байланыстың бар екендігін көрсетеді, яғни варианттар сол жақтан, төмен, оңға қарай орналасқандықтан, бұл төс ұстамының өзгеруінен тірі салмақтың өзгергендігін көрсетеді.

Кесте 4.2.3 Корреляциялық торды толтыру

х/у	166-169	170-173	174-177	178-181	182-185	186-189	190-193	194-197	198-201	202-205
360-379	∴	·								
380-399	·		∴							
400-419		∴	∴	∴						
420-439		·			∴					
440-459			·	∴						
460-479				·	∴					
480-499					∴					

500-519								••			
520-539								•			
540-559								•	••		•

Варианттардың бір сызық бойында шоғырлануы (тар сопақтықта), осы белгілер арасындағы үлкен өзара тәуелділіктің бар екендігін аңғартады. Егерде варианттар солдан төменгі оң жақ жоғарғы бұрышқа баратын сызықта орналасса, онда ол теріс (қайтымды) байланысты көрсетеді.

4.2.4 кесте Корреляция коэффициентін есептеу

X	a	y											p_x	$p_x a_x$	$p_x a_x^2$
		166-169	170-173	174-177	178-181	182-185	186-189	190-193	194-197	198-201	202-205				
		-3	-2	-1	0	1	2	3	4	5	6				
360-379	-4	2	1	I				II				3	-12	48	
380-399	-3	1	6	2								9	-27	81	
400-419	-2		3	2	2							7	-14	28	
420-439	-1			7	10	3						20	-20	20	
440-459	0			1	12	8	1					22	0	0	
460-479	1				1	4	6					11	11	11	
480-499	2					3	5	5	1			14	28	56	
500-519	3			III				4	IV			4	12	36	
520-539	4							1	5			6	24	96	
540-559	5							1	2		1	4	20	100	
p_y		3	10	12	25	18	12	11	8	0	1	100	22	476	
$p_y a_y$		-9	-20	-12	0	18	24	33	32	0	6	72			

$\sum y a_y^2$		27	40	12	0	18	48	99	128	0	36	408		
----------------	--	----	----	----	---	----	----	----	-----	---	----	-----	--	--

Бірақ варианттар корреляциялық тордың ұяшықтарында шашырап орналасады, бұл кезде байланыстың сипатымен дәрежесін көзбен анықтау қиын, сондықтан ең дұрысы бұл байланысты айқын сандық шамамен анықтап, яғни корреляция коэффициентін есептеп шығаруға болады.

Сызылған корреляциялық торға көлденеңінен және тігінен төрт графадан қосамыз. Арықарай қосылған графаларды толтыру және есептеу жүргізіледі; Есептеу жүрісі 4.2.4. кестесін қарағанда түсінікті болады.

$\sum r a_x a_y$ табу үшін корреляциялық торды класс бойында жатқан нольдік ауытқулармен, төрт квадрантқа қою сызықтармен бөлу қажет, сосын жиілігі бар, әрбір класс бойынша a_x ті a_y ауытқуын көбейтіп, және осы кластардың түйісуіндегі туындысын ұяшыққа жазу керек. $a_x a_y$ туындысын сәйкес ұяшық жиілігіне көбейту тиіс, соның нәтижесінде $r a_x a_y$ шамасы табылады. $\sum r a_x a_y$ санау жеке әрбір квадрант бойынша жүргізіледі.

$\sum r a_x a_y =$ (I квадрант = 24+9+8+36+12+6+4+7=106; II квадрант = 0; III квадрант = -3; IV квадрант = 4+6+12+20+30+36+12+15+8+80+40+30=293) = 396.

Екі белгі үшін b , b^2 , σ шамасын есептеу әдетте (вариациялық қатардағы сияқты) жүргізіледі

$$b_x = \frac{72}{100} = 0,72; \quad b_y = \frac{22}{100} = 0,22.$$

$$b_x^2 = 0,72^2 = 0,51; \quad b_y^2 = 0,22^2 = 0,05.$$

$$\sigma_x = \pm \sqrt{\frac{408}{100} - 0,51} = \pm 1,89;$$

$$\sigma_y = \pm \sqrt{\frac{476}{100} - 0,005} = \pm 2,17.$$

$$r = \frac{396 - (0,72 \cdot 0,22 \cdot 100)}{1,89 \cdot 2,17 \cdot 100} = \frac{380,2}{410,6} = +0,93.$$

Алынған корреляция коэффициенті +0,93 1-ге жақын, ол оның горбатов тұқымының қызыл сиырларының тірі салмағы мен төс ұстамының арасындағы өте үлкен дұрыс байланыстың бар екендігін көрсетеді.

Ішінара зерттеулердің корреляция коэффициенті, басқа іріктеулер сияқты, өзіндік қатесі болады. Кіші іріктеулер ($n > 100$) үшін корреляция коэффициентінің қатесін мына формуламен есептейді:

$$m_r = \pm \frac{1-r^2}{\sqrt{n}} \quad (40)$$

Аз санды іріктеулер үшін корреляция коэффициентінің қатесі:

$$m_r = \pm \frac{1-r^2}{\sqrt{n-2}} \quad (41)$$

Корреляция критериінің (t_r) дұрыстығы мына формуламен шығарылады:

$$t_r = \frac{r}{m_r} \quad (42)$$

Корреляция дұрыстығы бостандық дәрежесі санын (v) есепке ала отырып, Стюдент кестесі бойынша анықталады (кесте 2.9.1). t_r үшін бостандық дәрежесінің саны тең:

$$v = n - 2.$$

Егер t_r есептелгені ($t_r > t_{st}$) мәні кестедегіден тең немесе артық болса, онда корреляция коэффициенті дұрысырақ.

Біздің мысалымыз үшін:

$$m_r = \pm \frac{1 - 0,93^2}{100} = \pm 0,0135,$$

$$t_r = \frac{0,93}{0,0135} = 68,9.$$

Бұл жерде t_r кесте мәнінен артық t барлық ықтималдық деңгейінде: $P_{0,95} = 1,6$; $P_{0,99} = 2,0$; $P_{0,999} = 3,4$. Бұндай критерий жоғары деп аталады және ол корреляцияның үлкен дұрыстығын қоштайды.

Тапсырма 1. Енелерінің тірі салмағы (x) және бұзауларының туғандағы тірі салмағы (y) арасындағы корреляция коэффициентін есептеңіз

X	Y	X	Y	X	Y
438	38	513	47	468	29
502	41	439	45	492	45
456	37	487	39	398	42
380	20	395	28	415	29
479	45	493	50	438	33
500	48	480	49	450	42
405	26	475	44	395	30
463	48	390	23	423	23
412	28	453	35	485	48
483	45	487	38	426	32
446	44	413	31	487	46

Тапсырма 2. Орыстың күрең тұқымды биелерінің шоқтықтағы биіктігі (x) және төс ұстамы (y) арасындағы корреляция коэффициентін есептеңіз

x	y	x	Y	x	y
161	176	149	178	156	170
160	175	155	180	156	176
150	167	150	169	154	172
256	170	156	175	154	176
164	187	152	167	152	172
157	180	155	178	147	160
157	172	157	180	155	179
156	178	149	164	152	174
159	178	154	173	155	171

155	164	155	181	158	175
166	182	150	168	154	172
152	178	158	181	160	183
155	172	155	170	160	185
155	179	156	170	154	180
154	175	160	180	152	164
152	163	164	184	151	164
159	175	155	182	155	180
152	165	148	166	155	171
157	180	160	175	157	171
160	186	155	170	154	169
154	175	159	181	152	169
152	187	150	171	163	190
158	182	152	173	153	173
149	160	149	165	163	186
154	180	155	172	155	168

Бақылау сұрақтары.

1. Корреляция коэффициентінің дұрыстығын қалай есептейді?
2. Корреляциялық қатынас көрсеткіші қандай жағдайда қолданылады және ол қалай сипатталады?
3. Корреляция коэффициентіндегі белгі нені көрсетеді?
4. Теріс корреляцияны қалай түсінуге болады? Теріс корреляцияға мысал келтіріңіз.