

3.4 К.Пирсонның X^2 критеріі («хи-квадрат»)

Мақсаты. X^2 тәсілін меңгеру, оны есеп шығаруда қолдану.

Теориялық жиіліктер қандай да нақты есептелсе де, олар, әдетте, эмпириялық жиілік қатарымен сәйкес келмейді. Осыдан есептелген эмпириялық жиіліктерді салыстыру немесе күтілетін жиіліктердің дұрыстығын немесе олардың арасындағы кездейсоқтық ажырасуды бақылау қажеттілігі туындайды.

Эмпириялық және теориялық жиіліктерді, сандық және сапалық белгілерді салыстыру үшін К.Пирсон (1900) «хи-квадрат» (X^2) критеріін немесе сәйкестілік критеріін қолдануды ұсынды.

«Хи-квадрат» критерий формуласына эмпириялық ($P_{\text{эмп}}$) және теориялық ($P_{\text{теор}}$) жиіліктер арасындағы айырмашылық квадратты теориялық жиілікке ($P_{\text{теор}}$) бөлгендегі бөлшек жиынтығы жатады:

$$X^2 = \sum \frac{(P_{\text{эмп}} - P_{\text{теор}})^2}{P_{\text{теор}}} \quad (35)$$

X^2 белгісі бір санның квадраты емес, ол тек осы формуламен анықталған нәтиже шамасын көрсетеді. Себебі, эмпириялық жиілік ауытқуы квадратталып, X^2 критеріінің шамасы әрқашан оң. Возводятся

Эмпириялық жиіліктер есептелген немесе күтілетін $\Sigma(P_{\text{эмп}} - P_{\text{теор}}) = 0$ жиіліктермен толықтай тура келгенде және X^2 нольге тең болады. Егерде $\Sigma(P_{\text{эмп}} - P_{\text{теор}}) \neq 0$, бұл есептелген жиіліктердің эмпириялық жиілік қатарымен сәйкес келмейтіндігін байқатады. Бұндай жағдайда X^2 критеріінің мағынасын бағалау қажет, яғни теориялық түрде 0 ден ∞ дейін өзгеруі мүмкін.

Бұл нақты алынған шаманы $X^2_{\text{ф}}$ оның критикалық мағынасымен ($X^2_{\text{ст}}$) салыстыру арқылы жүзеге асады. Нольдік гипотеза, яғни ұйғарым, эмпирикалық және теориялық немесе күтілетін жиіліктер арасындағы ажырасу кездейсоқтықпен сипатталады, опровергается, егер $X^2_{\text{ф}} \geq X^2_{\text{ст}}$ бостандық санының дәрежесі және мағынасының қабылданған деңгейі үшін.

X^2 критикалық мағынасының әртүрлі мағына деңгейі және бостандық санының дәрежесі 2 кесте қосымшада берілген.

X^2 кездейсоқ шамасының ықтималдық мағынасының таралуы үздіксіз және ассиметриялы. Ол бостандық дәрежесінің санына байланысты және бақылау санының өсуіне қарай қалыпты таралуға жақындайды. Сондықтан X^2 критеріін дискретті бөлінуді бағалауда қолдану кейбір келеңсіздіктермен килігіп, ол оның шамасына, әсіресе кіші іріктеулерде әсерін тигізіп жатады.

Іріктеудің ең нақты бағалануын алу үшін, вариациялық қатарға бөлінуі 50 варианттан кем болмауы керек. Хи-квадрат критерийін дұрыс қолдану, шеткі кластардағы варианта жиіліктерінің 5 кем болмағаны; ал егерде ол 5 кем болса, онда олар көршілес кластардың жиіліктерімен бірігеді, яғни 5 көп немесе тең болатындай жиынтық шамасын құру керек. Жиіліктердің бірігуіне сәйкес кластардың да саны азаяды. Бостандық санының дәрежесі бостандық вариациясының шектеу санын есепке ала отырып, кластардың екінші реттік саны бойынша бекітіледі.

Мысал. Қояндарды жұқпалы аурулардан сақтандыру үшін жаңа препаратты сынау нәтижесін бағалау керек. 50 қоянның 20-сы профилактикалық препарат алды (тәжірибелік топ), ал 30-ы алған жоқ (бақылау). Тәжірибелік топта 7 особь ауруға шалдықты.

Тәжірибе нәтижелері препараттың профилактикалық әсерін немесе кездейсоқ себеп екендігін дәлелдейді ме?

Анық түйіндеме жасау үшін, тәжірибенің барлық мәліметтерін кестеге енгізіп, оларға сай өңдеуден өткізу керек.

Кесте 3.4.1 Екі топтағы қояндар арасындағы айырмашылық дұрыстығын анықтауды есептеу критерийінің сәйкестігі

Жануарлар топтары	Ауырған жануарлар саны		Сау жануарлар саны		Топтағы барлық жануарлар
	Бақыланатыны (Ф)	Теориялық күтілетіні (Т)	Бақыланыатыны (Ф)	Теориялық күтілетіні (Т)	
Тәжірибелік	7	8,4 (Т ₁)	13	11,6 (Т ₂)	20
Бақылау	14	12,6 (Т ₃)	16	17,4 (Т ₄)	30
Барлығы	21	21	29	29	50

Тәжірибелік және бақылау топтарындағы ауыру және сау жануарлары үшін теориялық күтілетін жиілікті – Т санау керек; $T_1 = \frac{20 \cdot 21}{50} = 8.4$; бақылау тобындағы жануарлар санының күтілетін шамасынан – Т₁ азайту арқылы Т₂ табуға болады, яғни $T_2 = 20 - 8,4 = 11,6$; $T_3 = \frac{30 \cdot 14}{50} = 12.6$; $T_4 = 30 - 12,6 = 17,4$.

Барлық шамаларды формулаға қоя отырып (35) аламыз:

$$x^2 = \frac{(7 - 8.4) - 1(2)^2}{8.4} + \frac{(13 - 11.6) - 1(2)^2}{11.6} + \frac{(14 - 12.6) - 1(2)^2}{12.6} + \frac{(16 - 17.4)^2}{17.4} = 0.26$$

Хи-квадраттың кестедегі есептелген мағынасына салыстыру жүргізу үшін, бостандық дәрежесінің санын, яғни класс санының бірлікке аздығын білу қажет. Төрттаңбалы кесте бойынша есептегенде бостандық дәрежесінің саны бірге тең. Біздің мысалымызда алынған хи-квадраттың стандартты мағынасын салыстыра отырып, табамыз, біздің есептеген шамамыз (0,26) кесте бағанының барлық стандарттық мағынасынан аз,

бостандықтың бір дәрежесіне сай келеді. Ендеше, препараттың профилактикалық әсері дәлденді деп санауға болмайды.

Тапсырма 1. Қара масақты персидтік бидайды қызыл масақтымен будандастырғанда бірінші ұрпақтағы барлық өсімдіктер қара масақты болып шықты, ал екінші ұрпақта, яғни гибридті тұқымдарды сепкен кезде, 154 қара масақтыға, 40 қызыл масақтыға және 15 ақ масақты өсімдіктерге ажырады. Осы жағдайдағы ажырауды ұйғарым көмегімен, күтілетін қатынасқа 12:3:1 сай келетіндігін тексеріңіз

Тапсырма 2. Индло-3-пропиогидросомды қышқылдың қорғаныс әрекетін зерттеуде, қояндарға Ауески ауыруын экспериментті түрде жұқтырғанда 20 особьтан 8-і тірі, 12-і өлді, ал терапиялық тиімділігін зерттегенде – 17 қоянның 6-ы тірі, 11-і өлді. Осы препараттың терапиялық және қорғаныс әсерінің тиімділігінің гипотезасын тексеріңіз.

Бақылау сұрақтары.

1. Сәйкес келу критеріі (хи-квадрат) дегеніміз не және оны генетикалық зерттеулерде қалай қолданады?
2. Сапалы белгілердің тұқым қууын зерттегенде хи-квадратты қалай қолданады?
3. Жануарлардың екі тобы арасындағы айырмашылық дұрыстығын анықтауда хи-квадратты қалай қолданады?