Модуль III.

№19 дәріс: Корреляциялық талдау және оның қолданылуы.

Мақсаты: студентитерге корреляциялық байланыстың мәнін, оны зерттеу әдістерін және корреляция коэффициенттерін есептеп, бағалау ережелерін түсіндіру; биологиялық және экологиялық деректерде корреляциялық талдауды дұрыс қолдану дағдыларын қалыптастыру.

Күтілетін нәтиже: студент Пирсон мен Спирмен коэффициенттерін таңдап, есептейді, олардың мәнділігін бағалайды және нәтижені дұрыс түсіндіреді.

Әдістері: дәріс.

Дәрістің топтама-тәсімі (тірек конспектісі немесе тезистер).

1. Корреляция (корреляциялық байланыс) ұғымы және түрлері.
2. Корреляциялық тәуелділікті зерттеу әдістері және корреляция коэффициенттері.
3. Корреляция коэффициентінің қасиеттері, қателігі және статистикалық мәнділігі.
4. Корреляция коэффициентін есептемейтін жағдайлар және қолданылуы.

1. Корреляция (корреляциялық байланыс) ұғымы және түрлері.

Көрсеткіштер мен белгілер арасындағы байланы)с көптеген ғалымдардың назарын аударды. Олардың біріне әсер етіп, екіншісін басқаруға мүмкін болады. Мысал ретінде топырақтың тамыр қабатындағы азот мөлшері мен өнімділік арасындағы байланысты атауға болады.

Белгілер арасындағы байланыс бірнеше түрлі болады. Функционалдық байланыста бір белгінің әр мәніне екіншісінің белгілі бір ғана мәні сәйкес келеді, мысалы, шеңбердің радиусы мен ауданы. Стохастикалық байланыста бір белгің мәнінің өзгеруі таралу заңдылықтарының өзгеруіне әкеледі. Статистикалық байланыста бір белгінің мәні орташа алғанда екінші белгінің мәніне байланысты өзгеретін байланыс.

Биологияда белгілер көптеген факторға тәуелді, сондықтан байланыс статистикалық, яғни корреляциялық сипатта болады: бір белгі өзгергенде екіншісінің орташа мәні өзгереді. Мысалы, балықтың ұзындығы артқан сайын салмағы орта есеппен артады, бірақ ұзындығы бірдей екі балықтың салмағы әртүрлі болуы мүмкін.

Корреляция латынның *correlation* деген сөзінен шыққан, ол өзара бір-бірімен байланыстық, қатынастылық, тәуелділік деген мағынаны білдіреді. «Корреляция» (co-relation) терминін ғылымға Фрэнсис Гальтон (Francis Galton, 1822–1911) енгізді, ал оны математикалық тұрғыдан Карл Пирсон (Karl Pearson) жетілдірді. Түрлері: бағыты бойынша оң (х артқанда у артады) және теріс (х артқанда у кемиді); пішіні бойынша сызықтық және сызықтық емес; белгі санына қарай жұптық және көпше. Мысалы, су температурасы көтерілгенде еріген оттегі мөлшері азаяды: бұл теріс корреляция.

Статистикалық зерттеулерде объектілердің әр түрлі қасиеттерін сипаттайтын белгілер арасындағы байланысты анықтау өте маңызды. Кездейсоқ көрсеткіштер арасында байқалады.

2. Корреляциялық тәуелділікті зерттеу әдістері және корреляция коэффициенттері.

Корреляциялық тәуелділіктерді зерттеу кестелік, графикалық және аналитикалық әдістер арқылы жүргізіледі. Кестелік әдісте корреляциялық кесте құрылады: жолдарға бір белгінің, бағандарға екіншісінің интервалдары қойылып, ұяшықтарға ортақ жиілік жазылады. Графикалық әдісте шашырау диаграммасы салынады: нүктелердің жайылуы бағытты, тығыздығы байланыстың күшін көрсетеді. Дегенмен есептеулерде графикалық тәсіл өте көп жұмысты қажет етеді. Сондықтан аналитикалық есептеулер қолданылады. Аналитикалық әдіс коэффициентті түрлі есептеулерге негізделген.

Пирсон коэффициенті (1) формуласымен есептеледі және екі белгі де сандық, таралуы нормалға жақын болғанда қолданылады.

$$r = \Sigma(x - \bar{x})(y - \bar{y}) / \sqrt{[\Sigma(x - \bar{x})^2 \cdot \Sigma(y - \bar{y})^2]} \qquad (1)$$

Спирменнің (Charles Spearman, 1904) рангілік коэффициенті (2) формуласымен есептеледі, мұндағы d — рангілер айырмасы. Ол реттік деректерде, нормал емес таралуда және ерекше мәндер болғанда қолайлы.

$$r_s = 1 - 6\Sigma d^2 / [n(n^2 - 1)] \qquad (2)$$

Мысалы, су ластануының балдық бағасы мен түр саны арасындағы байланысты Спирмен коэффициентімен бағалаймыз.

3. Корреляция коэффициентінің қасиеттері, қателігі және статистикалық мәнділігі.

Корреляция коэффициентінің мәні −1 мен +1 аралығында жатады (-1 ≥ r ≤ +1); таңбасы байланыс бағытын, абсолюттік мәні байланыс күшін көрсетеді.

Егер корреляция коэффициенті оң болса, бір көрсеткіштің артуы екінші көрсеткіштің де артуына әкеледі, ал кері байланыста бір көрсеткіштің артуы, екінші көрсеткіштің кемуіне әкеледі.

Шамамен |r| < 0,3 болса байланыс әлсіз, 0,3–0,7 орташа, 0,7-ден жоғары күшті деп бағаланады. r = 0 сызықтық байланыстың жоқтығын білдіреді, бірақ сызықтық емес байланыс болуы мүмкін.

Көрсеткіштер арасындағы байланыстың күшті немесе әлсіздігін сипаттау үшін *детерминация коэффициенті* қолданылады. Детерминация коэффициенті корреляция коэффициентінің квадратына (r^2) тең, кейде процентпен (r^2*100) де белгілейді. Коэффициент өлшем бірлігіне тәуелсіз, ал r^2 (детерминация коэффициенті) бір белгі өзгерісінің қанша үлесі екіншісімен түсіндірілетінін көрсетеді.

Басқа да биометрлік көрсеткіштер тәрізді корреляция коэффициентінің қателігі есептеледі. Корреляция коэффициентінің қателігі (3) формуласымен есептеледі.

Корреляция коэффициентінің қателігі:

$$m_r = \frac{1 - r^2}{\sqrt{n-1}} \qquad (3)$$

Мәнділік критерийі $t = r/m_r$, еркіндік дәрежесі n − 2: егер t кесте мәнінен үлкен болса, байланыс мәнді. Мысалы, n = 20 және r = 0,60 болса, m_r ≈ 0,19, t ≈ 3,18; 18 еркіндік дәрежесінде $t_{0,01}$ = 2,88 болғандықтан, байланыс p < 0,01 деңгейінде мәнді. Сенімділік аралығы Фишердің z-түрлендіруі арқылы есептеледі.

Егер біріккен таралу қалыпты болса, онда статистикалық қорытынды сызықтық корреляцияның іріктемелік коэффициентіне негізделеді, басқа жағдайларда Кендалл және Спирменнің рангілік корреляция коэффициенті анықталады, ал сапалық белгілер үшін хи-квадрат критерийі қолданылады.

4. Корреляция коэффициентін есептемейтін жағдайлар және қолданылуы.

Корреляция коэффициентін әрдайым есептеуге болмайды. Байланыс сызықтық емес болса (мысалы, оптимумы бар U-тәрізді қисық), r нөлге жақын шығып, байланысты жасырады. Бір-екі ерекше мән коэффициентті күрт өзгертеді, сондықтан алдымен график салу керек: Фрэнсис Анском (Francis Anscombe, 1973) төрт жиынтықтың r мәні бірдей (≈ 0,82) бола тұра, графиктері мүлде өзгеше екенін көрсетті. Біртекті емес топтар (мысалы, еркек және ұрғашы даралар) араласса, жалған байланыс шығуы мүмкін.

Белгі диапазоны тар немесе n өте аз болса, нәтиже сенімсіз. Ең маңыздысы: корреляция себеп-салдарды дәлелдемейді. Балмұздақ сатылымы мен суға батып кеткендер саны бірге артады, себебі екеуі де ыстық ауа райына байланысты. Қолданылуы: селекцияда белгілер байланысы, экологияда температура мен түр алуантүрлілігі, физиологияда доза мен әсер. Қазіргі бағдарламаларда (R-дағы cor.test, Python-дағы scipy.stats) корреляциялық

матрица мен жылу картасы құрылады; көп жұпты салыстырғанда түзету (мысалы, Бонферрони) қажет.

Бақылау сұрақтары:

1. Корреляция коэффициенті қандай байланысты көрсетеді?
2. Корреляция термині нені білдіреді және қандай жағдайда қолданылады?
3. Көрсеткіштер арасындағы байланыс теріс болуы мүмкін бе? Кері корреляциялық байланыс нені білдіреді?
4. Корреляциялық тәуелділіктерді зерттеудің кестелік, графикалық және аналитикалық әдістеріне сипаттама беріңіз.
5. Корреляци коэффициентінің қатесі қалай анықталады?
6. Корреляциялық қатынас 0,785-ке тек болған жағдайда детерминация коэффициенті нешеге тең болады?
7. Детерминация коэффициенті деген не? Детерминация коэффициенті 0,796-ға тең болған жағдайда корреляциялық қатынас неге тең болады?
8. Корреляция коэффициентін есептемейтін жағдайларды атаңыз.
9. Қауымдастықтың биологиялық өнімділігі мен шаруашылықтық өнімділігің арасындағы корреляциялық қастынас 0,79-ға тең. Бұл екі көрсеткіштің арасындағы байланыс қандай екенін білдіредді: а) әлсіз; б) қалыпты; в) тығыз; г) өте тығыз.

Әдебиеттер:

1) Аубакиров Х.Ә. Биометрия: оқулық / Х.Ә.Аубакиров. – А.: Дәуір, 2011. – 165-176 б.; 337-339 беттер.

2) Раманқұлова А.А. Биологиялық статистика: оқу құралы. – Алматы: Эверо, 2023. – 169-187 бет.

3) Шоқаманов Ю.К., Белгібаева Қ.Қ. Статистика . – Алматы: Экономика. – 540 бет.

4) Хили Д.Ф. Статистика негіздері: әлеуметтік зерттеу құралы. – Алматы : «Ұлттық аударма бюросы» қоғамдық қоры, 2020. – 576 бет.

5) Биологиядағы бағдарламамен қамтамасыз етілген статистикалық әдістер: оқу-әдістемелік құралы / А.М. Калимағамбетов, З.М.Бияшева, А.В. Ловинская, С.Б. Дәулетбаева. – Алматы: Қазақ университеті, 2019. – 58-64 б.