Модуль I.

№3 дәріс: Биологиялық деректер және оның түрлері.

Мақсаты: Білім алушыларға биологиялық деректердің табиғаты, құрылымы және негізгі түрлері туралы жүйелі түсінік қалыптастыру; биологиялық айнымалыларды өлшеу шкалалары бойынша жіктеуге, сапалық және сандық деректерді ажыратуға, бір өлшемді, екі өлшемді және көп өлшемді деректерді статистикалық талдауға дайындауға үйрету.

Әдістері: дәріс.

Дәрістің топтама-тәсімі (тірек конспектісі немесе тезистер).

1. Биологиялық деректер ұғымы, ерекшеліктері және оларды жинау әдістері.

2. Биологиялық деректердің өлшеу шкалалары бойынша жіктелуі: номиналдық, реттік, интервалдық және қатынастық деректер.

3. Сапалық және сандық деректер: дискретті және үздіксіз айнымалылар.

4. Бір өлшемді, екі өлшемді және көп өлшемді биологиялық деректер және оларды статистикалық талдауға дайындау.

1. Биологиялық деректер ұғымы, ерекшеліктері және оларды жинау әдістері.

Биологиялық деректер – тірі ағзалардың құрылысы, қызметі, дамуы, тұқымқуалаушылығы, таралуы және қоршаған ортамен байланысы туралы бақылау, өлшеу немесе тәжірибе нәтижесінде алынған мәліметтер жиынтығы. Биометрияда деректер зерттеу нысандарын сипаттайтын белгілердің тіркелген мәндері ретінде қарастырылады. Мысалы, өсімдіктің биіктігі, жапырақ саны, тұқым массасы, жануардың дене температурасы, қандағы гемоглобин мөлшері, популяциядағы даралар саны және судағы еріген оттек концентрациясы – биологиялық деректер.

Биометриядағы зерттеу нысаны бақылау нәтижесінде сандық немесе сапалық деректер алынған кез келген биологиялық объект болуы мүмкін. Биологияда математикалық статистиканы қолдану биологиялық объектілердің өзгермелі және өзгеруге бейім екендігіне байланысты. Мысалы, өсімдіктің бір түріндегі тұқымдар саны, бір популяциядағы жануарлар дараларының мөлшері, ұқсас тіршілік ету ортасы түрлерінің саны, адам қанындағы химиялық заттардың концентрациясы, бір түрдегі организмдердегі гаметалар саны, популяциялардағы гетерозиготалар саны және т. б.

Статистика – кездейсоқ шамалардың тәжірибелік талдауы. Белгісіз кездейсоқ шамаларға көптік бақылаудлардың (үлгі) негізінде баға бере аламыз. Деректер, қағида бойынша, жеке бақылауларға негізделген. Бұл бақылаулар немесе өлшеулер іріктеменің ең аз бірлігінде жасалған. Бақыланатын белгісіз кездейсоқ шаманың үлестірілім заңы *жалпы* деп аталады.

Үлгі – X1...Xn сандар тізбегі, тұрақты жағдайда тәжірибенің n есе қайтаналуымен алынған, мысалы ол n түрлі түрлерге тән мән болуы мүмкін.

Биологиялық нысандарды зерттеу зерттеушіні қызықтыратын немесе қажетті белгілер бойынша жүргізіледі. Белгілер – бір затты екіншісіен ажыратуға мүмкіндік беретін қасиет. Мысалы зерттелетін белгі – ағаштың биіктігі, м. Олардың әрқайсы өлшеудің ең аз бірлігі болып саналады. Ал өлшенген ағаш биіктігінің көрсеткіштері статистикалық жиын болып саналады. Өлшенген ағаш биіктігі әртүрлі болғандықтан, өлшенген белгіні *айнымалы*, ал ағылшын тілінен аударылған кітаптарда *варианта* деп те атайды.

Биометриялық әдістің негізгі объектісі болып *өзгермелі белгі* саналады. Өзгермелі белгі мәнінің жеке дараларда әр қилы өзгеруі *вариация* деп, ал оның жеке сандық мәндері *варианта* (*айнымалы немесе деректер*) деп аталады.

Барлық белгілер өзгермелі болады, алайда олардың барлығын өлшеу мүмкін емес.

2. Биологиялық деректердің өлшеу шкалалары бойынша жіктелуі: номиналдық, реттік, интервалдық және қатынастық деректер.

Өлшеу шкаласы – айнымалы мәндерінің қандай мағына беретінін және оларға қандай математикалық амалдар қолдануға болатынын анықтайтын жүйе. Өлшеу шкалаларын ғылыми тұрғыдан жүйелеуге американдық психолог және психофизик Стэнли Смит Стивенс елеулі үлес қосты. Ол 1946 жылы өлшеудің номиналдық, реттік, интервалдық және қатынастық төрт шкаласын ұсынды. Шкаланы дұрыс анықтау статистикалық әдісті дұрыс таңдаудың негізгі шарты болып саналады.

Номиналдық шкалада мәндер тек нысандарды топтарға бөледі. Мысалы, өсімдіктердің тіршілік формалары: ағаш, бұта, бұташық және шөптесін өсімдік. Мұнда бір санат екіншісінен «үлкен» немесе «кіші» емес. Түрлерді 1, 2, 3 сандарымен кодтау бұл сандарды сандық көрсеткішке айналдырмайды.

Реттік шкалада санаттардың табиғи тәртібі бар, бірақ олардың аралықтары тең деп есептелмейді. Мысалы, өсімдіктің зақымдануы «жоқ – әлсіз – орташа – күшті» деп бағалануы мүмкін. «Күшті» зақымдану «әлсізден» жоғары, бірақ олардың арасындағы айырманы дәл екі есе деп айтуға болмайды.

Интервалдық шкалада көршілес мәндердің айырмашылығы мағыналы. Мысалы, 10 °C пен 20 °C арасындағы айырмашылық 20 °C пен 30 °C арасындағы айырмашылыққа тең. Бірақ 20 °C температураны 10 °C-тан екі есе жылы деуге болмайды, өйткені Цельсий шкаласындағы нөл жылудың толық жоқтығын білдірмейді.

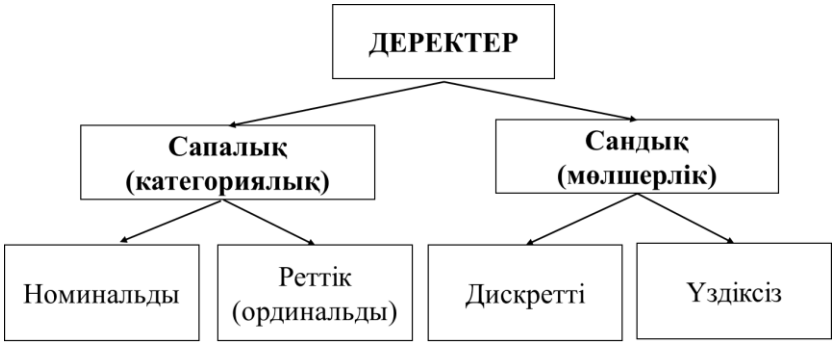
Қатынастық шкалада нақты нөлдік нүкте болады. Мысалы, тұқым массасы 20 мг және 10 мг болса, бірінші тұқымның массасы екіншісінен екі есе жоғары деп айтуға болады. Ағзаның ұзындығы, биомассасы, жасы, популяция тығыздығы және жауын-шашын мөлшері осы шкалаға жатады.

Өлшеу шкаласы қолданылатын статистикалық әдісті шектейді: номиналдық деректер үшін жиілік талдауы мен χ^2 критерийі, реттік деректер үшін Спирмен корреляциясы немесе Манн–Уитни критерийі, ал сандық интервалдық және қатынастық деректер үшін параметрлік әдістер қолданылуы мүмкін. Экологиялық зерттеулерде номиналдық, реттік және қатынастық шкалалар жиі кездеседі.

3. Сапалық және сандық деректер: дискретті және үздіксіз айнымалылар.

Статистикалықтәдісті қолданбас бұрын, зерттелетін деректердің қай түрге жататынын анықтап алу керек.

Деректердің (белгілердің) екі тобын ажыратады: сандық және сапалық. Сапалық (атрибутты) деректердің нақты бір мәндік көрсеткіштері болмайды. Олар жиынтықтағы әрбір нысанға тән белгілі бір сапаны сипаттайды. Оларды көбінесе визуальды бағалайды.



Сапалық белгілерді тікелей өлшеу мүмкін емес, оларды зерттеліп отырған топтың жеке мүшелерінде бар болуы мүмкін. Сапалық белгілер негізінен сөз арқылы сипатталады.

Сапалық деректер өз кезегінде номинальды (категориялық немесе атаулы) және реттік (ординальды, рангілік) деректер.

Номинальды деректер өзара ерекше және реттелмейтін категориялармен өрнектеледі. Мысалы жынысы (♀/♂) немесе қан тобы (A/B/AB/0), шаш немесе көз түсі, ұлты, жынысы. Олар альтернативті болуы мүмкін, мысалы ер/әйел, ауру/сау, қара/ақ т.б.

Реттік немесе ординальды деректер өзара ерекше және реттелген болады. Оларды санмен көрсетуге болмайды, бірақ реттеп ранжировать етуге болады: өсуі немесе кемуі бойынша, үлкеннен кішіге қарай немесе белгілі бір белгінің көрінісіне қарай орташа, әлсіз, күшті. Өмір сүру сапасы (нашар, қанағаттанарлық, жақсы, өте жақсы), тіршілік жағдайы (қалыпты, төмендеген, көтеріңкі, өте жоғары) т.б. Ранжировать ету үшін белгілі бір шкала қолданылады (МООС шкаласы), сондықтан бұл көрсеткішті *рангілік* деп те атайды. Кейде ранг сөзбен де беріледі, мысалы сукцессия стадиялары: пионерлік, өтпелі, климакстық.

Сандық белгілерді тікелей өлшеуге немесе санауға болады, мәндерін санмен өрнектейді. Олар өз кезегінде дискретті (үзілісті өзгермелі) және үздіксіз өзгермелі деректерге бөлінеді. Дискретті сандық деректер бүтін сандар (тұтас мән) болады, олар тек қана санау нәтижесінде белгілі бір есепке алынатын сандық мәндер. Мысалы, ағаштар саны, класстағы оқушылар саны, жанұядағы адам саны, құс жұмыртқаларының саны, гүл күлтесінің саны, қоңыз денесіндегі дақтар саны т.б.

Үздіксіз деректер белгілі бір диапозон аралығында үздіксіз өзгеріспен сипатталады, мәндер саны шексіз, өлшеу дәлдігіне қарай бөлшек түрінде де беріледі. Мәндер диапозондағы кез келген мәнді қабылдайды. Мысалы дене салмағы кг-ға дейінгі дәлдікпен немесе микрограммға дейін өлшенеді, адам бойы метр немесе см бойынша өлшенеді. Ал ұсақ жануарлар салмағын миллишрамм, микронға дейін өлшейді.

Тура өлшеу немесе санау арқылы алынған деректер алғашқы деп аталады.

Сандық және сапалық белгілерден басқа белгілі бір формула арқылы анықталатын туынды белгілер де болады. Оларға санның белгілі бір үлесі, процент және индекстер жатады, мысалы Шеннон, Жаккар т.б. индекстер.

4. Бір өлшемді, екі өлшемді және көп өлшемді биологиялық деректер және оларды статистикалық талдауға дайындау.

Деректерді зерттелетін айнымалылар саны бойынша бір өлшемді, екі өлшемді және көп өлшемді деп жіктеуге болады. Бір өлшемді деректерде әрбір бақылау бірлігі үшін бір айнымалы ғана тіркеледі. Мысалы, 50 өсімдіктің тек биіктігі өлшенсе, бұл бір өлшемді деректер жиыны болады. Мұндай деректердің таралуы, орташа шамасы, медианасы, дисперсиясы, стандарттық ауытқуы және өзгеру коэффициенті есептеледі. Гистограмма мен қораптық диаграмма деректердің таралуын көрсетуге мүмкіндік береді.

Екі өлшемді деректерде әр нысан бойынша екі айнымалы өлшенеді. Негізгі мақсат – олардың арасындағы байланысты немесе екі топтың айырмашылығын зерттеу. Мысалы, өсімдіктің биіктігі мен биомассасы, су температурасы мен еріген оттек мөлшері, жануардың жасы мен дене массасы қатар тіркеледі. Сандық айнымалылардың байланысы корреляция және регрессия арқылы талданады. Корреляцияның болуы себеп-салдарлық байланысты автоматты түрде дәлелдемейтінін ескеру қажет. Мәселен, ауа температурасы жоғарылаған кезде өсімдіктің өсу жылдамдығы да артуы мүмкін, бірақ бұл байланыс ылғал, жарық немесе қоректік заттар сияқты үшінші факторлардың әсерінен өзгеруі ықтимал.

Көп өлшемді деректерде бір бақылау бірлігіне үш немесе одан да көп айнымалы сәйкес келеді. Мысалы, көл экожүйесін зерттеу кезінде температура, pH, тұздылық, фосфат, нитрат, хлорофилл мөлшері және әрбір гидробионт түрінің саны тіркеледі. Мұндай мәліметтер деректер матрицасы түрінде беріледі: жолдар – бақылау бірліктері, бағандар – айнымалылар.

Бақылау сұрақтары:

1. Бір өлшемді, екі өлшемді, көп өлшемді деректер сипаттамасы.
2. Сапалық және сандық деректердің айырмашылығы.
3. Туындаған (екіншілей) деректерге мысал келтіріңіз.
4. Атрибутивті белгілер деп қандай белгілер айтылады?
5. Бір өлшемді және екі өлшемді деректер үшін қандай статистикалық әдістер қолданылады?4)

Әдебиеттер:

1) Аубакиров Х.Ә. Биометрия: оқулық / Х.Ә.Аубакиров. – А.: Дәуір, 2011. – 119-120; 123-130 б.

2) Раманқұлова А.А. Биологиялық статистика: оқу құралы. – Алматы: Эверо, 2023. – 9-14 бет.

3) 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. – 15-18 б.

4) Sokal, Robert Reuven; Rohlf, F. James. Biometry: The Principles and Practice of Statistics in Biological Research. 4th ed. New York: W. H. Freeman, 2012.

5) Zar, Jerrold Howard. Biostatistical Analysis. 5th ed. Upper Saddle River: Pearson Prentice Hall, 2010.