«ЗАТВЕРДЖУЮ»
Завідувач кафедрою
інформатики та
кінезіології

_____І.П. Занєвский
(підпис,ініціали,прізвище)

_____200__р

ЛЕКЦІЯ №2
з навчальної дисципліни
«КОМП'ЮТЕРНА ТЕХНІКА І М АТЕМАТИЧНА СТАТИСТИКА»

Тема: *Статистичні таблиці та графіки*

Навчальний потік

для студентів першого року навчання факультету спорту, фізичного виховання, здоров'я людини і туризму

Навчальна мета: Ознайомити студентів з попередньою обробкою статистичних даних, поданням статистичної інформації в табличному та графічному вигляді, а також з характеристиками варіаційного ряду.

Виховна мета: Розвивати логічне мислення, увагу, пам'ять, спостережливість, поглиблювати знання з математичної статистики як науки, розширювати кругозір з інформатики, математичної статистики в галузі ФКіС.

Навчальні питання і розподілення часу:

Вступ _____ -10…хв.

1. Попередня обробка статистичних даних._____15 хв.

2. Табличне та графічне подання інформації.____20 хв.

3. Характеристики варіаційного ряду._____20хв.

4. Інтервали, частоти й накопичені характеристики варіаційного ряду.____15хв.

Заключення та відповіді на запитання _____- 10 хв.

Навчальна література

1. Ільків О.С. Матвіїв В.І. Інформатика та комп'ютерна техніка (з елементами математичної статистики): Навч. посіб. –Львів: ЛДУФК. 2010.
2. Дибкова Л.М. Інформатика та комп'ютерна техніка: Навч. Посіб. – Київ: Академвидав, 2005.
3. Литвин І.І. Інформатика: теоретичні основи і практикум. – Львів: Новий світ, 2004.
4. Глинський Я.М. Практикум з інформатики. – Львів: Деол, 2005.
5. Русіло П.О., Заневський І.П. Математична статистика. Обробка і аналіз результатів спортивних вимірів. - Львів, 1995.
6. Статистика. Підручник за ред. С.С. Герасименка. Київ: КНЕУ, 2000.
7.Основы математической статистики. Под ред. В.С.Иванова, -М.: ФиС, 1990.
8.Глинський Я.М. Інформатика: інформаційні технології. - Львів: Деол, 2003.

Додаткова

1. Петров П.К. Информационные технологии в физической культуре и спорте. – М.: ФКиС, 2008.
2. Височанський В.С., Кардаш А.І., Костів О.В., Черняхівський В.В. "Елементи інформатики", - Львів: "Світ", 1990.
3. Андреев А.Г. и др. Microsoft Windows XP: Home Edition и Professional. Русские версии / Под общ. ред. А.Н. Чекмарева. – СПб.: БХВ - Петербург, 2003.
4. Руденко В.Д., Макарчук О.М., Патланжоглу М.О. Практичний курс інформатики / За ред. Мадзігога В.М. - К.: Фенікс, 1997.
5. Фигурнов В.Е. IBM PC для пользователя. - М.: Финансы и статистика, 1992.
6. Зайцева Т.И., Смирнова О.Ю. В сб.: Информационные технологии в образовании. – М., 2000.
7. Kosmol A., Kosmol I. Komputery - nowoczesne technologie w sporcie. - Warczawa: AWF, 1999
8. Волков В.Ю. Компьютерные технологии в образовательном процессе по физической культуре в вузе: Монография. – СПб.: СПбГТУ, 1997.
9. Макарова М.В. Електронна комерція. Посібник. Київ. Видавничий центр "Академія". 2002.

Тема 2. Статистичні таблиці та графіки

1.Попередня обробка статистичних даних.

2.Табличне та графічне подання інформації.

3.Характеристики варіаційного ряду.

4.Інтервали, частоти й накопичені характеристики варіаційного ряду.

Глибина статистичних знань і широта використання статистичних методик фахівцями в спорті та фізичному вихованні змінюється від країни до

країни. Чим вище економічне положення країни тим ширше там застосовується статистика. Яскравий приклад тому – Сполучені Штати Америки – країна всезагальної статистики, висновки якої узагальнює і координує всесвітньовідомий статистичний інститут Д. Гелапа, що має філіали в 19 країнах світу. Нерозуміння значення статистики означає фактичне сповзання будь-якого фахівця до наукового, діяльнісного і світоглядного примітивізму.

Знання математичної статистики і вміння її застосовувати потрібні не лише для успішного навчання, але також для ефективної роботи по спеціальності в будь-якій галузі, в тому числі і в галузі фізичної культури і спорту. Важливими причинами і визначальними факторами вивчення математичної статистики майбутніми фахівцями галузі фізичної культури і спорту є:

1. Статистичні підходи дають розуміння тієї логіки, яка лежить в основі розвитку фізичного виховання і спорту вищих досягнень, оскільки фактичні параметри такого розвитку мають імовірнісний характер.
2. Спеціалісти в галузі фізичного виховання і спорту вищих досягнень повинні вміти в рамках своєї компетенції обробляти результати записів, вимірювань, тестувань, обчислювати варіацію значень фізичних чи спортивних параметрів, можливі похибки спостережень і характер розсіювання вимірів.
3. Фахівець галузі фізичної культури і спорту повинен розуміти статистичні факти, що мають відношення до його роботи, передбачати зміни з тим, щоб будувати правильну стратегію в роботі з групами людей, що займаються фізичними вправами та спортсменів, щоб не допустити регресу чи небажаних наслідків. Тому фахівцям фізичної культури та тренерам-практикам необхідно знати, як інтерпретувати статистичні дані і робити на підставі їх аналізу адекватні висновки.

4. Вивчення математичної статистики сприяє розвитку науково-обґрунтованих поглядів, критичних, дедуктивних здібностей, необхідних в процесі навчання, при проведенні наукових досліджень та в практичній діяльності .

У процесі спостереження чи вимірювання якого-небудь показника отримують ряд чисел (вимірів). Результати вимірів, репрезентовані випадковими числами, називаються *вибірковою сукупністю* або просто *вибіркою.*

Вибірка – це сукупність однорідних величин, яка доступна для вимірів. Сукупність усіх значень, які можна було б отримати для вибірки, що вивчають, називається *генеральною (основною) сукупністю.*(Рис.1) Генеральна сукупність – це безмежна кількість однорідних вимірів.

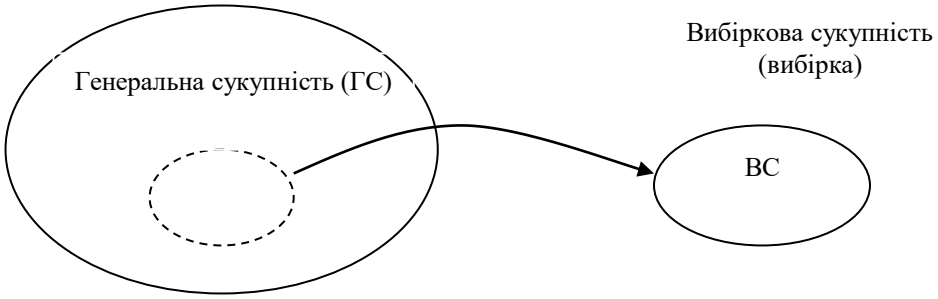


Рис.1.

Однією з основних характеристик вибірки є її **обсяг** (n), який визначається числом об'єктів спостереження. За вибірковими (експериментальними) даними оцінюють параметри, котрі дозволяють описати всю генеральну сукупність, визначають інтервал, у якому зі заданим рівнем достовірності (вірогідності) міститься істинне значення оцінюваного параметра, а потім перевіряються ті чи інші твердження і робляться висновки про властивості всієї генеральної сукупності.

Вибіркова сукупність має бути *репрезентативною*, тобто обсяг її повинен бути достатньо великим, наприклад, тридцять вимірів або більше.

За результатами вимірів вибірки оцінюють емпіричні середні величини, характеристики розсіювання й однорідність вибірки, порівнюють розподіл результатів з нормальним.

Вибіркова сукупність – це маса однорідних в певному відношенні елементів, що мають єдину якісну основу, але різняться між собою по певній ознаці і підлягають певному закону розподілу.

Сукупність може бути однорідною і неоднорідною.

Однорідна сукупність – якщо вивчається ознака, яка є загальною для всіх одиниць.

Неоднорідна сукупність – об'єднує явища різного типу.

Статистична ознака: певна властивість окремої одиниці сукупності.

Атрибутивна ознака: характеризує стан або якість певного елемента статистичної сукупності (наприклад: професія, національність, колір шкіри).

Кількісна ознака: виражає властивість певного елемента статистичної сукупності числом; значення цього числа, властиве даному елементу сукупності називають **варіантом.**

Ознака, яка приймає в межах певної сукупності різні значення називається **варіюючою**. Коливання значень ознаки від елемента сукупності до елемента називають **варіацією ознаки**. Кількісні ознаки поділяють на дискретні і неперервні.

Для відображення рівнів ознак застосовують наступні **шкали вимірювання:**

- метрична – це звичайна числова шкала, яку використовують для виміру фізичних величин (ваги, площі, часу тощо);
- порядкова – встановлює порядок елементів по певній ознаці (рейтинг спортсменів по спортивних досягненнях, рейтинг команд тощо);
- номінальна – шкала цифрових кодувань якісних об'єктів (наприклад, спортсмен чоловічої статі має код 1, жіночої статі – 0);

Відповідно до цього, виміри виконані за допомогою метричної шкали називають метричними, порядкової – порядковими або рейтинговими, номінальної – номінальними.

Статистичні закономірності встановлюють на основі дії **закону великих чисел** (ЗВЧ), математичне обґрунтування якого дає теорія ймовірностей. Зміст ЗВЧ в наступному: в масі індивідуальних подій загальна закономірність виявляється тим точніше, чим більше спостережень. Отже, **статистична закономірність** – це кількісна закономірність змін статистичної сукупності в просторі і часі. Вона властива не окремим елементом сукупності, а всій масі, - сукупності в цілому.

Будь-яке статистичне дослідження має три етапи:

- статистичне спостереження (збір первинного статистичного матеріалу шляхом реєстрації вимірів у спорті чи фізичній реабілітації);
- зведення і групування зібраних даних;
- аналіз і синтез отриманих зведених матеріалів.

В цілому статистичне дослідження базується на **статистичній методології** – це комплекс спеціальних, властивих лише статистиці засобів дослідження кількісних закономірностей в структурі, динаміці і у взаємозв'язках досліджуваних явищ.

Статистичне спостереження – це спланована, науково-організована реєстрація масових даних про досліджуваний процес і є першим етапом будь - якого статистичного дослідження.

В статистичній практиці застосовують три організаційні форми статистичного спостереження: звітність, спеціально-організовані спостереження, статистичні реєстри. Із спеціально-організованих спостережень в практиці статистики фізичного виховання і фізичної реабілітації застосовують: обліки, спеціальні обстеження, опитування. Статистичні реєстри дають відомості, що неперервно доповнюються, наприклад, про офіційно зареєстровані спортивні організації, товариства.

За ступенем охоплення одиниць сукупності статистичне спостереження буває суцільним (коли охоплює всіх обстежуваних спортсменів) і несуцільним (коли досліджується лише їх частка). Несуцільне спостереження має п’ять видів: основного масиву, вибіркове, монографічне, анкетне, моніторинг. В практиці статистики фізичного виховання і фізичної реабілітації широкого застосування набули монографічне спостереження (за видатними спортсменами) і моніторинг – систематичне спостереження за фізичним станом спортсмена, реабілітанта, тощо.

Другим етапом будь якого статистичного дослідження є зведення і групування даних, отриманих в результаті статистичного спостереження (групи спортсменів, чи окремих реабілітантів).

Зведення – це комплекс послідовних операцій по узагальненню конкретних одиничних фактів спостережень з метою виявлення типових рис і закономірностей, властивих досліджуваному об’єкту. Результати зведення подають, як правило у вигляді статистичних таблиць.

Зведення є простими (це простий підсумок результатів по певній ознаці) і груповими – це комплекс операцій, що включають групування, підрахунок групових і загальних підсумків.

Групування **–** це процес систематизації первинних даних з метою добування інформації, котра в них міститься.

Ранжуванням називають розміщення результатів вимірів у порядку зростання або спадання.

Вибірку великого обсягу розбивають на інтервали, кожен з яких містить певний діапазон значень досліджуваної ознаки.

Наближено число інтервалів k можна знайти за **формулою Стерджеса**: $k = 1 + 3{,}32 \cdot \lg n$, де n – обсяг вибірки,

lg – символ десяткового логарифма.

Ширину кожного з інтервалів (***крок інтервалу***) визначають за такою формулою:

$$h = (X_{max} - X_{min}) / k$$

де X_{max} і X_{min} – максимальне та мінімальне значення результатів вимірів.

Нижня границя групування першого інтервалу ($X_{н1}$) вибирається рівною X_{min}. Нижня границя наступних інтервалів є одночасно верхньою границею попередніх інтервалів ($X_{ні} = X_{в(i-1)}$). Верхня границя останнього інтервалу дорівнює X_{max}.

Значення нижніх і верхніх границь обчислюються за такими формулами:

$$X_{ні} = X_{min} + (i-1)h$$

$$X_{ві} = X_{min} + i \cdot h$$

де i=1, 2,…,k – номер інтервалу.

Середина кожного інтервалу

$$X_{ср\ i} = X_{min} + \frac{(2j-1)}{2} h$$

Величина, котра визначається числом результатів вимірів вибірки, які містяться в кожному інтервалі групування, називається ***частотою інтервалів***. Частота інтервалу позначається символом – f_j. Сума частот усіх інтервалів завжди дорівнює обсягу вибірки.

Накопичена частота інтервалу – це число, отримане послідовним додаванням частот у напрямку від першого інтервалу до останнього. Накопичена частота (Σf_j) отримується простим послідовним додаванням частот попередніх інтервалів. Накопичена частота в останньому інтервалі дорівнює обсягу вибірки.

Частістю (відносною частотою) називається відношення частоти до обсягу вибірки $F_j = f_j/n$. Сума частостей усіх інтервалів завжди дорівнює одиниці.

Накопиченою частістю називається відношення накопиченої частоти до обсягу вибірки: $\Sigma F_i = \Sigma f_i/n$. Накопичена частість в останньому інтервалі дорівнює одиниці.

Таблиця, заповнена відповідними значеннями, розглянутими вище, є ***інтервальним варіаційним рядом*** вимірів (табл. 1).

Варіаційним рядом вимірів називається подвійний числовий ряд, котрий показує, яким чином чисельні значення досліджуваної ознаки, пов'язані з їх повторністю у вибірці. Варіаційні ряди дають уявлення про характерні особливості змінності ознаки.

Таблиця 1

Таблиця інтервального варіаційного ряду

№ інтервалу i	Границі інтервалів **Нижня верхня** Xні Xві	Середина інтервалів Xc.i	Частота fi	Накопичена частота Σfi	Частість Fi	Накопи-чена частість Σ Fi
1	Xmin — Xmin +h	Xmin +h/2	f_1	f_1	Fi =fi/n	ΣFi = Σfi/n
2	Xmin +h — Xmin +2h	Xmin +3h/2	f_2	$f_1 + f_2$		
3	Xmin +2h — Xmin +3h	Xmin +5h/2	f_3	$f_1 + f_2 + f_3$		
...	...	...	...	...	…	...
k	Xmin +(k-1)h — Xmin +kh = Xmax	Xmin +(2k-1)h/2	f_k			

Трактування графічного методу представлення статистичних даних як особливої знакової системи - штучної знакової мови - пов'язане з розвитком семіотики, науки про знаки і знакові системи.

Знак в семіотиці служить символічним виразом деяких явищ, властивостей або відносин.

Знакові системи, що існують в семіотиці, прийнято розділяти на мовні і немовні.

Немовні знакові системи дають уявлення про явища навколишнього світу (наприклад, шкала вимірювального приладу, висота стовпчика ртуті в термометрі і т. д.).

Окрім сигнальних функцій мовні знакові системи виконують також завдання зіставлення сукупності явищ, їх аналізу. Характерно, що в цих

системах поєднання знаків набуває сенс тільки тоді, коли їх об'єднання проводиться по певних правилах.

У мовних знакових системах розрізняють природні і штучні системи знаків, або мов.

З погляду семіотики людська мова, виражена знаками-буквами, складає природну мову.

Штучні мовні системи використовуються в різних галузях життя і техніки. До них відносяться системи математичних, хімічних знаків, алгоритмічні мови, графіки і ін.

Не виключаючи природної мови, штучні, або символічні мови спрощують виклад спеціальних питань у певній області знань.

Таким чином, статистичний графік - це креслення, на якому статистичні сукупності, такі, що характеризуються певними показниками, описуються за допомогою умовних геометричних образів або знаків. Представлення даних таблиці у вигляді графіка справляє сильніше враження, ніж цифри, дозволяє краще осмислити результати статистичного спостереження, правильно їх тлумачити, значно полегшує розуміння статистичного матеріалу, робить його наочним і доступним. Це, проте, зовсім не означає, що графіки мають лише ілюстративне значення. Вони дають нове знання про предмет дослідження, будучи методом узагальнення початкової інформації.

При побудові графічного зображення слід дотримуватися ряду вимог. Перш за все графік повинен бути достатньо наочним, оскільки весь сенс графічного зображення як методу аналізу в тому і полягає, щоб наочно зобразити статистичні показники. Крім того, графік повинен бути виразним, дохідливим і зрозумілим.

Графік складається з графічного образу і допоміжних елементів. Графічний образ – це сукупність ліній, фігур, крапок, якими зображені статистичні дані. Діаметричні знаки, малюнки або образи, що використовуються при побудові статистичних графіків, мають багато видів.

Це крапки, відрізки прямих ліній, знаки у вигляді фігур різної форми, штрихування або забарвлення (круги, квадрати, прямокутники та ін.). Ці знаки застосовуються для порівняння статистичних величин, що зображають абсолютні і відносні розміри порівнюваних сукупностей. Порівняння на графіці проводиться по деяких вимірюваннях: площі або довжині однієї із сторін фігури, місцезнаходженню крапок, їх густині, густині штрихування, інтенсивності або кольору забарвлення.

Допоміжні елементи включають загальний заголовок, умовні позначення, осі координат, шкали з масштабами і числову сітку.

Словесні пояснення (експлікація графіка) поміщених на графіці геометричних образів, різних по їх конфігурації, штрихуванню або кольору, дозволяють в думках перейти від геометричних образів до явищ і процесів, зображених на графіку.

У статистичних графіках найчастіше застосовується система прямокутних координат, але є і графіки, побудовані за принципом полярних координат (кругові графіки).

Коли графік будується в прямокутних координатах, на горизонтальній осі абсцис і вертикальній осі ординат в певному порядку розташовуються характеристики статистичних ознак явищ, що зображаються, або процесів, а в полі графіка розміщуються геометричні знаки, складові, сам графік. Поле графіка - це простір, в якому розташовуються геометричні ознаки, що створюють графік.

Ознаки, що розташовуються на осях координат, можуть бути якісними і кількісними.

Одне з важливих завдань статистичного графіка - це його композиція: відбір статистичного матеріалу, вибір способу зображення, тобто формату графіка. Розмір графіка повинен відповідати його призначенню. Для статистичних графіків зручні формати із співвідношенням сторін поля $1:\sqrt{2}$. Але у багатьох випадках зручна квадратна форма графіка.

У заголовку (назві) графіка визначається завдання, яке вирішується за допомогою графіка, дається характеристика місця і часу, до якого відноситься графік.

Написи уздовж масштабних шкал указують, в яких одиницях вимірюються ознаки. Цифри значень кожного параметра проставляються у пригранічних відмітках масштабних шкал.

Масштабна шкала - лінія (на статистичному графіку зазвичай пряма), що несе на собі масштабні відмітки з їх числовими позначеннями. Краще робити ці позначення тільки на відмітках, що відповідають круглим числам: у такому разі проміжні відмітки читають шляхом відліку від найближчого числа, позначеного на масштабній шкалі. Згідно масштабним відміткам на діаграмному полі відкладають розміри явищ або процесів, що зображаються. Масштабні відмітки розташовуються на шкалі рівномірно (шкала рівномірна, арифметична) або нерівномірно (шкала функціональна, шкала логарифмічна).

Шкала функціональна – масштабна шкала, де числові значення помічених крапок виражають значення аргументу, а розташування цих крапок відповідає рівномірно розподіленим значенням деякої функції того ж аргументу. З функціональних шкал в статистичних графіках застосовують головним чином логарифмічну шкалу. При цьому, якщо розглядаються дві величини, то така шкала може бути застосовна до обох або тільки до однієї з них (“напівлогарифмічний” графік або масштаб). Відстані між крапками, нанесеними по числових відмітках логарифмічної шкали, відповідають різниці логарифмів відповідних чисел і, отже, характеризують співвідношення між числами.

За способом побудови статистичні графіки діляться на діаграми і статистичні карти. Діаграми - найбільш поширений спосіб графічних зображень. Це графіки кількісних відносин. Види і способи їх побудови різноманітні. Діаграми застосовуються для наочного зіставлення в різних аспектах (просторовому, тимчасовому і ін.) незалежних один від одного

величин: територій, населення і так далі При цьому порівняння досліджуваних совокупностей проводиться за якою-небудь істотною варіюючою ознакою. Статистичні карти - графіки кількісного розподілу по поверхні. По своїй основній меті вони близько примикають до діаграм і специфічні лише в тому відношенні, що є умовними зображеннями статистичних даних на контурній географічній карті, тобто показують просторове розміщення або просторову поширеність статистичних даних. Геометричні знаки, як було сказано вище, - це або крапки, або лінії або площини, або геометричні тіла. Відповідно до цього розрізняють графіки точкові, лінійні, площинні і просторові (об'ємні).

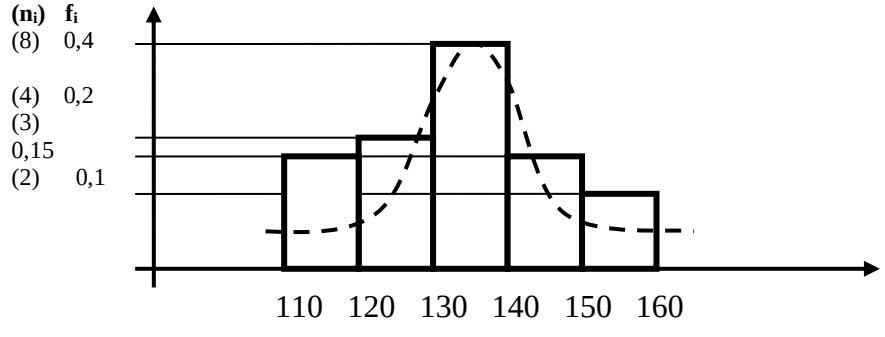
При побудові точкових діаграм як графічні образи застосовуються сукупності крапок, при побудові лінійних - лінії. Основний принцип побудови всіх площинних діаграм зводиться до того, що статистичні величини зображаються у вигляді геометричних фігур і, у свою чергу, підрозділяються на стовпчикові, смугові, кругові, квадратні і фігурні.

Статистичні карти по графічному образу діляться на картограми і картодіаграми.

Залежно від кола вирішуваних завдань виділяють діаграми порівняння, структурні діаграми і діаграми динаміки.

Для збільшення наочності та спрощення аналізу варіаційних рядів використовують їх графічне зображення у вигляді *гістограми, полігона частот і кумуляти*.

Гістограма – це стовбчата діаграма, яка використовується для графічного зображення розподілу змінних ознак і складається зі суміжних прямокутників. Основа кожного прямокутника дорівнює ширині інтервалу групування h, а висота – пропорційна частотам інтервалу.



Полігон частот (полігон розподілу) утворюється ламаною лінією, котра з'єднує точки, що відповідають на осі абсцис – середнім значенням інтервалів групування, а на осі ординат – частоті цих інтервалів.

Кумулята *(полігон накопичених частот або частостей)* отримується при з'єднанні відрізками прямих точок на координатній площині, абсциса яких відповідає верхнім границям інтервалів групування, а ордината – накопиченим частотам або накопиченим частостям, вираженим у відсотках.

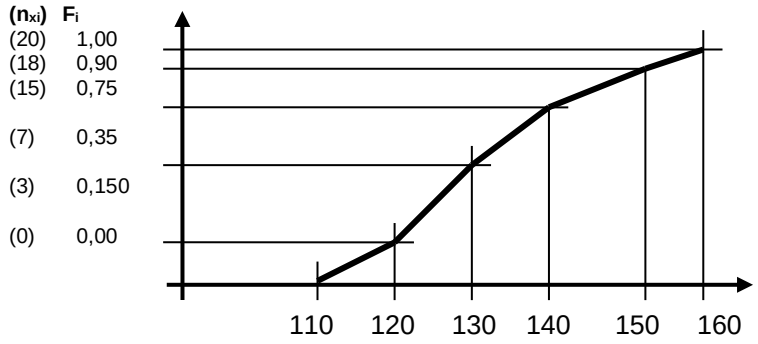


Рис.2. Полігон накопичених частостей - кумулята систолічного тиску.

Питання для самоконтролю:

1. Назвіть причини і визначальні фактори вивчення математичної статистики.
2. Дайте визначення поняттю "***Вибірка***".
3. Що таке **обсяг вибірки**.
4. Дайте визначення поняттю ***Групування***.
5. Дайте визначення поняттю ***Ранжування.***
6. Правило Сиерджеса і його застосування.

7. За якою формулою визначають ***ширину*** кожного з інтервалів (***крок інтервалу***).
8. Яким чином обчислюють ***частоту*** інтервалів?
9. Яким чином обчислюють ***накопичену частота інтервалу?***
10. За якою формулою обчислюють ***частість (відносну частоту)*** ?
11. Дайте визначення поняттю **Статистичний графік**.
12. Якими можуть бути ознаки. що розташовуться на осях координат статистичних графіків?
13. Який графік називається ***Гістограма?***
14. Який графік називається **Полігон частот?**
15. Який графік називається **Кумулята**?

Навчально-матеріальне забезпечення

Мультимедійний проектор_____

Самостійна робота:

Попередня статистична обробка спортивних результатів.

Лекцію розробили: к.пед. н., доц. О.С. Ільків
(посада,вчений ступінь,вчене звання,підпис, ініціали, прізвище)

Обговорено на засіданні кафедри: інформатики та кінезіології
(назва кафедри)

Протокол від _____20_____р. № _____