Статистика

Хоноруван .ас.д-р Пенко Данков

Ръководство по статистика – Величко, Тодор , Любомир , Поля, 2009 20лв

Основи на статистиката – Величко , Тодор, Любомир, Абагар 2009

Справочник статистика 2008

Могат да се ползват на изпита

Курсова работа към края на семестъра включва задачи и тестове

Статистическо изучаване

1. Същност – статистическото изучаване е общността от операциите, процедурите и фазите от получаването на първичните сведения за отделните случаи на масовото явление до получаването на неговите обобщаващи характеристики. То се състои от три взаимосвързани и обособени фази   
   а) Статистическо наблюдение – регистрират се първични сведения за единиците на масовото явление  
   б) Статистическа групировка – осъществява преход от индивидуални сведения за отделните единици към данни за цялата съвкупност  
   в) Статистически анализ – обхваща всички операции с които от получените абсолютни данни се изчисляват производни статистически величини
2. Основни статистически понятия   
   а) Статистическа съвкупност – това е общност от единици чрез които се проявява масовото явление в определени пространствени и времеви граници. Характеризира се с еднородност и конкретност. Съвкупността притежава обем и вътрешна структура. Съвкупностите биват :   
   - моментни и периодни –   
   - генерални и представителни –   
   - реални и хипотетични –   
   б) Статистическа единица – това е отделна елементарна и неделима форма на проява на изучаваното масово явление. Определя се ясно и недвусмислено с определени критерии. Единицата притежава качествени особености, те могат да имат материална форма, могат да бъдат събития или факти. Тези единици които нямат материален характер се наричат статистически случаи.  
   в) Статистически признак – качествена особеност на статистическата единица. Притежава конкретни индивидуални определения които могат да бъдат числови или словесни. Признаците биват вариационни (количествени) и категории (качествени). Вариационните признаци могат да бъдат прекъснати и непрекъснати   
   г) Измерителни скали -   
   - номинална – отчита признаците словесно   
   - ординална – отчита признаците словесно но в някаква степен   
   - рангова – подрежда единиците като им отрежда дадено място   
   - интервална – присвоява конкретни числови стойности  
   д) Обект на изучаване е масово явление   
   е) Обобщаващи числови характеристики – това са числови величини характеризиращи общото типичното и закономерното в статистическия признак отнасящо се до цялата съвкупност
3. Видове статистически изучавания   
   а) Изчерпателни изучавания – наблюдават се и се изучават всички единици  
   б) Частични   
   - Репрезентативно изучаване   
   - Изучаване на основният масив   
   - Статистическа оценка   
   - Статистическа анкета   
   - Статистическа монография

Статистическо наблюдение

1. Същност – планомерен, систематичен и научно организиран процес на регистрация на индивидуални сведения за статистическите единици. Осъществява се в две форми :   
   - статистическа отчетност  
   - специално организирани наблюдения   
   Всяко наблюдение се изпълнява по предварително съставен план. Планът определя необходимата организация която трябва да се създаде за да се осъществи наблюдението.
2. Единица за наблюдение – единицата от която се получават сведения. Отчетна единица, техническа единица.  
   Време на наблюдението -   
   Критичен момент -   
   Период на регистрация –
3. Място на наблюдението – териториалното разположение на съвкупността
4. Органи на наблюдението биват три вида :  
   а) Преброители (анкетьори)   
   б) Контрольори   
   в) Ръководители
5. Регистрация на единични сведения :  
   а) Самонаблюдение -   
   б) Кореспондентски начин -   
   в) Експедиционен начин –
6. Статистически формуляр – това са специално създадени и отпечатани листове в които се записват единичните сведения.
7. Източници на сведения – обекти или субектите от които се получава информация
8. Грешки при наблюдението   
   а) Аритметични и логически   
   б) Съзнателни и несъзнателни   
   в) Случайни и систематични

Статистически графични изображения

1. Същност – статистически и графични изображения (СГИ) са целесъобразна форма за нагледно опростено и ясно представяне (изобразяване) на статистическите величини. При тяхното построяване се използват различни геометрични средства (точки, линии, фигури, тела и др. ). Стилизирани образи, географски карти, схеми и др. Със СГИ се решават следните по важни задачи -   
   - характеризиране на разпределението, териториално разположение и развитието на едно или повече явления.  
   - сравняване на структурата на едно или повече явления  
   - характеризиране на връзки и зависимости  
   - сравняване на еднородни статистически величини -   
   - за характеризиране на продажби и планови задачи
2. Елементи на СГИ   
   - заглавие – то ясно точно и кратко характеризира обекта (статистическа съвкупност) мястото или времето на изследването   
   - скали – те представляват линии състоящи се от деления на които се нанасят числовите стойности на изобразяваните явления. Те биват праволинейни и криволинейни ( кръгови), равномерни (аритметични) неравномерни (логаритмични),   
   - мащаб – той показва на един отсек (разстоянието м/у две съседни деления) от съответната скала на единица площ, на фигура или точка, в графичното изображение какъв размер от изобразяваното явление отговаря   
   - графичен образ – той представлява съвкупност от точки различни геометрични фигури или стилизирани образи с помощта на които се изобразяват визуално изучаваните явления.   
   - легенда – тя представлява словесно изражение или отразено по друг начин обяснение на използваните основни означения на СГИ, чрез които се подпомага правилното възприемане на представения обект.
3. Видове и подвидове СГИ – според вида (формата) на графичния образ   
   а) Диаграми – СГИ се представят посредством различни геометрични форми или картинни изображения (стилизирани образи)  
   - линейни – явленията се представят нагледно с помощта на геометрични линии като най често се използва първи квадрант на правоъгълната координатна система  
   - плоскостни – те представят изследваните явления и процеси с различни двумерни фигури от геометрията – правоъгълници, триъгълници, квадрати и др.  
   - точкови – те служат за характеризиране на зависимости между две явления с помощта на точки   
   - стериограми – те представят изследваните явления и процеси с различни примерни фигури от стереометрията кубове, цилиндри пирамиди   
   - картинни – стилизиран образ   
   б) Картограми – използват се за характеризиране на териториалното разположение на единиците или статистическите съвкупности и представляват географски карти  
   в) Картодиаграми – съчетание на карти и диаграми   
   II Според функционалното им предназначение   
   - за изследване на динамиката (развитието) на изучаваните явления и процеси   
   - за характеризиране на структурата   
   - за характеризиране на структурата и динамиката  
   - за представяне на разпределението на единиците в съвкупностите  
   - за сравнение на две и повече едноименни статистически величини   
   - за графично представяне на териториални статистически редове   
   - за характеризиране на връзки и зависимости   
   - за характеризиране на стопански явления в това число изпълнение на планови задания

Статистическа групировка

1. Същност – статистическата групировка е процедура на систематизиране и обобщаване на единиците на съвкупността, качествено еднородни групи. Групировката се състои от три последователни операции   
   а) Образуване на групите – свързано е с природата на изучаваното явление, трябва да се постигне пълно обхващане на единиците.  
   б) Отнасяне на единиците в групите – извършва се на основата на конкретното определение на признака за всяка единица.  
   в) Изброяване на единиците в групите
2. Видове групировки   
   а) Според вида на групировъчният признак те са четири вида   
   - Вариационни – е по количествен признак, при дискретни прекъснати признаци се използват дискретни групировки. При непрекъснати признаци се използват интервални групировки.   
   - Категории  
   - Териториални   
   - Темпорални  
   Ширината на интервала се определя по 3 начина.   
   а) интервалите имат равна ширина. Тя се определя по следната формула h = k брой на групите, Х мах – максимална стойност на признака, Хмин – минимална стойност на признака  
   Пример – В едно предприятие трябва да се групират работниците в 5 групи според работната заплата. Минималната заплата в предприятието е 350лв. а максималната 600лв.   
   Когато не е известен броя на групите се използва следната формула   
    където N брой на единиците в съвкупността, lg – десетичен логаритъм.   
   В предприятието работят 100 работници -   
   б) Геометричен принцип – ширината на интервала нараства   
   в) Логически или целеви принцип – групите имат различна ширина и са съобразени с вътрешния строеж на съвкупността   
   Категориини – използват се при качествени признаци   
   Териториални съобразно териториалните подразделения   
   Темпорална групировка – по време на възникване на единиците. Според броя групировъчни различаваме прости и сложни групировки прости с една, сложни с 2 и повече.  
   Резултатите от групировката се представят в статистически таблици

Статистически анализ

1. Същност – състои се от действия, операции и методи за получаване на обобщаващи числови характеристики и тяхната смислова интерпретация за разкриване на закономерностите в развитието, структурата и взаимната обвързаност на явленията.
2. Етапи на анализа – статистическият анализ преминава през три етапа   
   а) подготвителен етап – набират се данните, проверяват се за пълнота и достоверност  
   б) Същински (основен) етап – обхваща обработката на данните   
   в) Заключителен – обхваща обобщаване интерпретиране на получените резултати и представянето им в подходяща форма
3. Видове анализ -   
   а) Според състоянието на явленията -   
   - статичен – състоянието  
   - динамичен – развитието   
   б) според времевата насоченост   
   - ретроспективен -   
   - диагностичен -   
   - прогностичен -   
   в) Според характера на анализа  
   - описателен   
   - сравнителен  
   - номографски – законно установяващ
4. Предпоставки за анализа –   
   а) Приоритет на качественият анализ   
   б) Пълнота на данните   
   в) Съпоставимост   
   г) Достоверност

Методи за представяне на информацията

1. Статистически таблици – това са форми получени от пресичането на редове и колонии, които съдържат систематизирана и нагледно представена статистическа информация
2. Елементи на таблиците  
   а) Формални елементи – заглавие, заглавен ред(антетка), челна колона, редове, колони и клетки  
   б) Логическите елементи на таблицата са   
   - статистически субект   
   - статистически предикат
3. Видове таблици
4. Оформяне на таблиците   
   а) мерни единици -   
   б) еднаква точност -   
   в) всички клетки са запълнени  
   г) посочва се източникът на данните
5. Статистически редове
6. Според целта с която се използват  
   - за популяризация на определени резултати   
   - за подпомагане на статистическият анализ  
   - за изясняване на определени изчислителни операции

Абсолютни и относителни величини

1. Статистически величини – това са числови характеристики на масовите явления, които се получават в резултат на наблюдението, групировката и анализа. Тези които се получават от наблюдението и групировката са първични величини, а при анализа чрез математически операции от първичните величини се получават производни величини. Статистическите величини са три вида :  
   - Абсолютни -   
   - Относителни -   
   - Средни –
2. Абсолютни величини – те изразяват размера на масовите явления и на техните съставни части. Конкретни са по време, място и съдържание и винаги са наименувани величини. Единиците за измерване са прости или сложни, натурални, основни или стойностни. Абсолютните величини са три основни вида:  
   а) Величини за обем – оказват броя на единиците в съвкупността  
   б) Величини за равнище – характеризират признаци на отделна единица  
   в) Величини за маса – характеризират признаци за цялата съвкупност. Масата се получава като произведение на обем и равнище
3. Относителни величини – това са производни величини, получени чрез отношение на две абсолютни, средни или други относителни величини. Основното действие е деление относителните величини се представят в три форми:  
   - Коефициента – показва колко пъти сравняваната величина е по голяма или по малка от базата  
   - Процентна форма – получава се чрез умножение по 100  
   - Промилна форма – умножава се по 1000. Получава се в промили
4. Видове относителни величини те са 6 вида -   
   а) Планови относителни величини – характеризират величината и изпълнението на планови задачи   
   б) Структурни относителни величини – показват вътрешният строеж на статистическата съвкупност  
   в) Интензивни относителни величини – показват силата (интензивността) с която една съвкупност поражда определен резултат   
   г)Динамични относителни величини – характеризират относителните изменения настъпващи с течение на времето  
   д) Териториални относителни величини – характеризират териториални различия   
   е) Координационни относителни величини – представляват отношение на две разноименни величини при което се получава логически издържан резултат

Средни величини

1. Обща характеристика – средните величини са обобщаващи числови характеристики характеризиращи общото, типичното и закономерното. Показват центъра на разпределение. Изразяват това определение на признака което е общо за цялата съвкупност тъй като е формирано от въздействието на главните трайно действащи фактори, при елиминирано влияние на случайните и несъществени фактори
2. Видове средни величини   
   а) Според характера на осредняваните величини – различаваме вариационни и хронологични средни.  
   Хронологичните средни биват моментни и периодни   
   б) Според начина на изчисляване -   
   - Алгебрични средни величини   
   - Неалгебрични (позиционни) средни величини  
   в) Според начина на включване на честотите   
   **- Претеглени средни – при групирани данни   
   - Непретеглени средни – при негрупирани данни**   
   г) Според обхвата си средните величини биват -  
   - Средни на генерална съвкупност   
   - Средни на извадка
3. Средна аритметична величина – използва се когато значенията на признака могат да се сумират. Използва се за осредняване на обеми равнища и маси. Полученият резултат е наименувана величина в същите мерни единици както и осредняваните величини. При интервални редове се използва трансформация в дискретни редове. Вземат се средите на интервалите при отворени интервали се възприема ширината на съседните
4. Средна хармонична величина – използва се в два случая   
   - Когато признакът е в обратна зависимост с определящото свойство   
   - Когато теглата са показател за маса  
   - непретеглен вариант
5. Средна геометрична величина – използва се при величини с определяща функция умножение. Това са относителни величини, индекси, темпове на развитие. Непретеглен вариант   
   Претеглен
6. Средна квадратична величина – използва се при осредняване на показатели на разсейване, при величини характеризиращи повърхнини или площи
7. Средна кубична – използва се при показатели за вместимост или обем
8. Мажорантност на средните величини -

Средни величини от динамични редове

1. Средна хронологична периодна величина
2. Средна хронологична моментна величина  
   пр. регистрираните безработни в бюрото по труда в една община през 2011 са били както следва :   
   - 01 01 – 68 души  
   - 01 02 - 75 души  
   - 01 03 – 82 души  
   - 01 04 – 93 души  
   - 01 05 – 101 души   
   - 01 06 – 113 души   
   - 01 07 – 102 души   
   да се изчисли средномесечният размер на безработните за полугодието   
   пр. ДМА на едно предприятие през 2011 са били както следва   
   0101 – 160 000  
   0102 – 190 000  
   0103 – 240 000  
   0109 – 264 000  
   3112 – 248 000

Позиционни средни величини

1. Медиана – средна на положение. Това е значението на признака на тази единица която разделя съвкупността на две равни части. Когато съвкупността съдържа четен брой единици медианата се получава като полусбор от значенията на признака на двете централни единици. При негрупирани данни медианата се посочва директно. За целта единиците се подреждат по значенията на признака. При групирани данни се използва следната формула …………..
2. Квартили – отбелязват се с Q1,Q2 е медиана,Q3.
3. Мода – показва най-срещаното. При групирани данни се използва следната формула. Модалната група е групата с най-много честоти.
4. Съотношение между средните величини

Статистическо разсейване

Единиците на всяка статистическа съвкупност се различават както от средната величина така и по между си. Тази не еднаквост на признака при отделните единици се нарича статистическо разсейване или вариация. Разсейването се поражда както от действието на постоянни трайни причини така и от влиянието на случайни фактори. Статистическото разсейване дава допълнителна характеристика на условията в които се е формирала средната величина.

1. Размах – изчислява се като разлика между най-голямото и най малкото значение на признака. В абсолютен размер е наименувана величина същите мерни единици както и признака. В относителна форма размаха се дава в %.
2. Средно аритметично отклонение – изчислява се чрез осредняване на абсолютните стойности на отклоненията на осредняваните величини от тяхната средна аритметична.
3. Средно квадратично отклонение – нарича се още стандартно отклонение. Представлява средна квадратична величина от отклоненията. В относителен размер се нарича коефициент на вариация.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ДА(хил.лв) | Фирми (бр.) | Хd | Xf | х - хср | (х - хср) 2 | (х - хср)2f |
| от 5 до 15 | 1 | 10 | 10 | -26 | 676 | 676 |
| 15 - 25 | 4 | 20 | 80 | -16 | 256 | 1024 |
| 25 - 35 | 7 | 30 | 210 | -6 | 36 | 252 |
| 35 - 45 | 12 | 40 | 480 | 4 | 16 | 192 |
| 45 - 55 | 6 | 50 | 300 | 14 | 196 | 1176 |
|  | 30 |  | 1080 |  |  | 3320 |

Квадратът на средното квадратично отклонение се нарича дисперсия. Дисперсията е единствената величина която няма мерна единица.

1. Квартилно отклонение -   
   пр. Разполагаме с данни на фирма   
   Q1 = 111 хил.лв, Me = 132 хил.лв, Q3 = 149 хил.лв
2. Средна разлика – изчислява се на основата на разликите между значенията на признака.
3. Съотношения между показателите за разсейване

Моменти, асиметрия и ексцес

1. Моменти – това са статистически величини, характеризиращи разпределението на съвкупността
2. Асиметрия – характеризира формата на разпределението   
   а) Измерители на асиметрията -   
   - Коефициент на пирсън   
   б) Юл -   
   в) Боули -   
   г) моментен коефициент на асиметрия –
3. Ексцес – характеризира връхната източеност на кривата.

Репрезентативно изучаване

1. Обща характеристика – програмиране изследване в редица области не е възможно да се постигне пълно обхващане на изучаваната съвкупност. В други случаи изчерпателното изучаване не е целесъобразно, тъй като е свързано с големи разходи на време и средства и кадрово осигуряване. Тогава се използва репрезентативния метод. Той дава възможност да се изследват съвкупности като се наблюдават сравнително малки части от тях наречени извадки. Целта е на основата на извадката да се направят обосновани съждения за цялата съвкупност. Съвкупността се характеризира със своите параметри. Извадката притежава характеристики. При извадково изучаване на основата на общи теоретични положения и като се спазват определени правила и процедури по характеристиките на извадката могат да се правят обосновани заключения за параметрите на генералната съвкупност. Заключенията имат вероятностен характер. В теорията на статистическите заключения има два дяла:  
   а) Статистическо оценяване -   
   б) Статистическа проверка на хипотези –
2. Репрезентативни извадки – за да бъде надеждна информацията,получена от извадката, тя трябва да е представителна по отношение на генералната съвкупност. Една извадка е представителна, когато е формирана по начин гарантиращ еднаква възможност на всяка единица от съвкупността да попадне в извадката. Такъв подбор се нарича случаен. Ако извадката е случайна характеристиките се отличават от параметрите на съвкупността само с размера на стохастичната грешка, ако подбора не е случаен се появява систематична грешка, която не се подава на контрол. Подборът може да бъде повторен или безповторен.
3. Видове подбор   
   а) Прост случаен подбор – използва се при сравнително малка съвкупност. Единиците се номерират, номерата се записват на подходящи носители, разбъркват се и се изтеглят толкова колкото са необходими за извадката  
   б) Механичен подбор – изготвя се списък на единиците от който се избират единици за извадката през определен интервал наречен крачка на подбора. Първата единица се определя случайно. Крачката на подбора се получава по тази формула където n – обем на извадката, N – обем. Една фирма има 100 работници. Да се направи извадка от 10 работника.   
   в) Райониран (стратифициран) подбор – генералната съвкупност се разделя на типични в някакво отношение групи. Единиците се избират случайно от съответните групи. Така се осигурява представителност в групите  
   г) Сериен (многостепенен) подбор – извадката се формира на две или повече степени, на първият етап се подбират случайно, големи съставни единици. От тях също случайно се избират единици за извадката.
4. Оценки – на основата на извадката се оценяват неизвестните параметри на генералната съвкупност. Оценките биват точкови и интервални
5. Интервално оценяване – интервалът които съдържа параметъра се нарича доверителен интервал. Вероятността с която се гарантира заключението се нарича гаранционна вероятност. Преминава се през три стъпки -   
   а) Изчислява се средна грешка  
   б) Изчисляване на доверителния интервал –
6. Обем на извадката –

Индекси

1. Същност – индексите са сложни относителни величини, характеризиращи измененията, настъпили в изследваните явления в териториален, планов или динамичен
2. Видове индекси
   1. според степента на обхващане на единиците на изучаваните съвкупности, индексите биват :  
      а) индивидуални   
      б) множествени
   2. според вида на изпълняваните функции, индексите биват   
      а) синтетични  
      б) аналитични
   3. Според характера на измерваните различия   
      а) динамични  
      б) териториални  
      в) планови
   4. Според формата на тяхното построяване, множествените индкси биват  
      а) агрегатни  
      б) осреднителни
   5. Според характера на индексираните величини   
      а) индекси на обем   
      б) индекси на равнище   
      в) индекси на маси
3. Функции на индексите   
   Измерват динамиката на явленията   
   Измерват териториалните различия на явленията   
   Измерват влиянието на факторите върху общото изменение на явленията
4. Динамични индекси на равнище,обем и маса   
   Използвани символи   
   i – индивидуален индекс  
   I – множествен   
   q – обем   
   p - цена  
   qp – маса
5. Производни коефициенти

Коефициент на определението К0 = r2 . 100  
коефициент на неопределеност   
КНО = 100 – К0   
Коефициент на акорелация

1. Измерване на зависимости при неинтервални скали
   1. Рангови корелационни коефициенти   
      rs =   
      Ed2 – сума от квадратите на разликите между ранговете  
      N – броя на наблюденията  
      Коефициент на Кендал  
      r k = където  
      S – сума от баловете   
      P – броя на съответствията между ранговете   
      Q – броя на несъответствията (инверсиите)
   2. Коефициенти на четириклетъчната корелация  
      Коефициент на асоциацията на Юл   
      r as =   
      Коефициент на контингенцията на Пирсън   
      r cont =   
      r col =
   3. Коефициенти на контингенцията (взаимната свързаност) при повече от две определения на признаците  
      Връзка между спомагателните коефициенти