Изпит ВПС - Втора част

**код - 7K8A**

**ЗАДАЧИ:**

**Въпрос 1.**

Направена е извадка с обем 4 и са получени стойности 10-к, 20-к, 10+к, 20+к, където к е 0. Намерете извадковата медиана

**Въпрос 2.**

За предстоящи избори за президент проучване показва, че 51% от допитаните ще гласуват за A. Да се напише съответната основна и алтернативна хипотеза за тестване дали A ще спечели изборите?

**Въпрос 3.**

Нека H0 : p = 0.1 H1 : p > 0.1 Тогава за проверка на хипотезата се използва какъв тест? (едностранен, лявостранен, дясностранен, и пр.)

**Въпрос 4.**

В трети курс в ПУ има 2000 студенти, които имат средното тегло 70 кг, модата 71 к, и стандартно отклонение 10 кг. Оказва се, че най-тежките десет студенти от курса прекъсват следването си. Кое от следните твърдения е вярно за новия състав на курса от 2990 студенти?

А) Новото средно тегло намалява.

Б) Новата мода на теглото се намалява

В) Стандартното отклонение на теглото не се променя

**Въпрос 5.**

Производител твърди, че в киселото мляко съдържанието на мазнините е поне 2% и го пакитира в пластмасови кофички по 200 гр. В случайна извадка от 25 кофички по 200 гр. се оказва, че всяка от тези кофички средно съдържа по 4.1 грама мазнини със стандартно отклонение 0.1 грама. Напишете основната хипотеза и алтернативата за проверката на подходящата хипотеза за съответната ситуация.

**Въпрос 6.**

Знае се, че височината на момчетата от 10 клас в даден регион е нормално разпределена. За да провери дали учениците са по-високи от 160 см се избират по случаен начин 100 ученици от 10 клас от съответния регион и се оказва, че те са средно високи 161 см с дисперсия 9 см. Кое разпределение ще използвате, за да построите 90% доверителен интервал на височината и защо?

**Задача 1.**

За да се провери дали дадена машина не пълни по-леки пакети от 2 кг се избират 10+A пакети по случаен начин и се притеглят и се намира, че те тежат 1900 гр със стандартно отклонение 100 гр., където A е 5. Знае се, че теглото е нормално разпределено.

А) Числото 2 е статистика или параметър?

Б) Доверителен интервал за статистика или за параметър се построява?

В) В статистиката се проверява хипотеза за статистика или за параметър?

Г) Кое разпределение ще използвате за построяване на доверителен интервал за средното тегло и защо?  
 Д) Постройте 92% доверителен интервал на средното тегло на пакетите.

Е) Има ли статистическо значимо основание да се счита, че пакетите са по-леки от 2 кг? Използвайте ниво на значимост 0,1 и следвайте процедурата: дефинирайте нулева хипотеза и алтернативата; определете статистиката и разпределението, което ще използвате и защо; напишете критичната област; направете извод и дайте интерпретация.

Ж) Използвайте p-стойността на теста, за да се провери статистически твърдението.

**РЕШЕНИЕ:**

**Въпрос 1.**

За да намерите извадковата медиана, трябва да подредите стойностите в извадката в нарастващ ред и да изберете средната стойност.

Във вашия случай, извадката съдържа следните стойности: 10-к, 20-к, 10+к, 20+к, където к е 0.

Подреждаме стойностите в нарастващ ред:

10-к, 10+к, 20-к, 20+к

Средната стойност в този случай е средната стойност на две средни стойности (10+к и 20-к), тъй като имаме четен брой стойности в извадката.

Така извадковата медиана е средноаритметичната средна на 10+к и 20-к:

(10+к + 20-к) / 2

Защото стойността на к е 0, можем да изчислим извадковата медиана:

(10 + 20) / 2 = 15

Извадковата медиана на тази извадка е 15.

**Въпрос 2.**

Решението за тестване на хипотезата се извършва чрез статистически анализ на резултатите от изборите, за да се сравнят фактическият процент на гласоподавателите, които гласуват за кандидат А, с предвидената стойност от 51%.

Нека предположим, че след изборите се получава фактически резултат, при който 48% от гласоподавателите гласуват за кандидат А. Тогава можем да проведем статистически тест, като използваме, например, z-тест, за да определим дали това е статистически значимо различие.

Стъпките за изпълнение на z-теста са както следва:

Формулираме нулевата хипотеза (H0): Процентът на гласоподавателите, които гласуват за кандидат А, е равен на 51% (p = 0.51).

Формулираме алтернативната хипотеза (H1): Процентът на гласоподавателите, които гласуват за кандидат А, е по-малък от 51% (p < 0.51).

Задаваме ниво на значимост (α) - обикновено 0.05.

Изчисляваме стандартната грешка (SE) по формулата:

SE = √(p \* (1 - p) / n), където n е общият брой допитани лица в проучването.

Изчисляваме z-статистиката по формулата: z = (p̂ - p) / SE, където p̂ е фактическата наблюдавана честота на гласоподавателите, които гласуват за кандидат А.

Сравняваме получената z-статистика с критичната стойност на z-разпределението, която е асоциирана със зададеното ниво на значимост (α). Ако получената z-статистика попада в критичната област (т.е. по-малка от критичната стойност), отхвърляме нулевата хипотеза и приемаме алтернативната. В противен случай, не отхвърляме нулевата хипотеза.

**Въпрос 3.**

Решението включва използване на едностранен тест, за да проверите дали параметърът p е по-голям от 0.1.

Стъпките за изпълнение на едностранния тест са както следва:

Формулираме нулевата хипотеза (H0): p = 0.1

.

Формулираме алтернативната хипотеза (H1): p > 0.1.

Задаваме ниво на значимост (α) - обикновено 0.05.

Избираме подходящ статистически тест за пресмятане на p-стойността (например, z-тест или t-тест), в зависимост от характеристиките на данните.

Изчисляваме статистиката на теста (например, z-статистика или t-статистика) въз основа на наблюдаваните данни.

Изчисляваме p-стойността - вероятността да получим статистиката на теста или по-екстремна стойност, ако нулевата хипотеза е вярна.

Сравняваме p-стойността с нивото на значимост (α). Ако p-стойността е по-малка от α, отхвърляме нулевата хипотеза в полза на алтернативната. В противен случай, не отхвърляме нулевата хипотеза.

**Въпрос 4.**

Твърдение А (Новото средно тегло намалява): Може да се промени или да остане същото, в зависимост от теглата на новите 10 студента.

Твърдение Б (Новата мода на теглото се намалява): Верно, модата ще се промени и ще бъде различна от 71 кг.

Твърдение В (Стандартното отклонение на теглото не се променя): Верно, премахването на най-тежките 10 студента няма да промени дисперсията и стандартното отклонение на теглото.

**Въпрос 5.**

Основната хипотеза (H0): Съдържанието на мазнини в киселото мляко е поне 2%.

Алтернативна хипотеза (H1): Съдържанието на мазнини в киселото мляко е по-малко от 2%.

Тук тестваме дали средното съдържание на мазнини в киселото мляко е по-малко от 2%. Нулевата хипотеза предполага, че производителското твърдение е вярно и съдържанието на мазнини е поне 2%, докато алтернативната хипотеза предполага, че съдържанието на мазнини е по-малко от 2%.

Тестваме тази хипотеза с помощта на статистическия тест, като използваме средното съдържание на мазнини от извадката от 25 кофички.

**Въпрос 6.**

Решението е да се използва разпределение на Стюдънт (t-разпределение) за построяване на 90% доверителен интервал за средната височина на момчетата от 10 клас. Тъй като параметрите на генералната съвкупност (средната височина и дисперсията) са неизвестни, избираме t-разпределението, което компенсира несигурността, свързана с това.

С размер на извадката от 100 ученика и предположението, че измерената средна височина е приближение на нормално разпределение, можем да използваме t-разпределението за построяване на доверителния интервал.

Избираме 90% доверителен интервал, което означава, че ще използваме 5% от вероятността в опашките на разпределението (2.5% в лявата и 2.5% в дясната опашка), оставяйки 90% от вероятността в средата. Това ни позволява да изчислим граници на доверителния интервал около измерената средна височина на 161 см.

Изпит ВПС - Втора част

**код - XM8K**

**ЗАДАЧИ:**

**Въпрос 1. (2 - 7K8A)**

**Въпрос 2. (5 - 7K8A)**

**Въпрос 3. (3 - 7K8A)**

**Въпрос 4. (1 - 7K8A)**

**Въпрос 5. (4 - 7K8A)**

**Въпрос 6. (6 - 7K8A)**

**Задача 1.**

**РЕШЕНИЕ:**