Проверка статистических гипотез

Оглавление

Завершить показ

Статистические гипотезы.

- Статистические гипотезы – это предположения исследователя о результатах измерений, выраженные в формализованном лаконичном виде.
- Статистические гипотезы разделяются на 4 типа.
- Статистические гипотезы
- Нулевые
- Альтернативные
- Направленные
- Ненаправленные

Оглавление

Статистические гипотезы

- **H_0** – нулевая гипотеза:
  Она делает предположение о том, что различия между сравниваемыми выборками отсутствуют. Её математический смысл состоит в том, что Хср.1 –Хср.2→0, т.е. различие между выборками стремится к нулю. На самом деле различия могут отклоняться от 0, но быть не достоверными или не доказанными.

Оглавление

Статистические гипотезы

- **H_1 (H_A)** – альтернативная гипотеза
(противостоящая нулевой гипотезе)

- Её смысл заключается в том, что различия между выборками есть и что они достоверны.

Оглавление

Статистические гипотезы

- Ненаправленная гипотеза – доказываем то, что выборки достоверно различаются, но не доказываем чем именно.

- Направленная гипотеза – под влиянием исследуемого фактора в определенном направлении (больше или, наоборот, меньше) изменяется исследуемый признак в экспериментальной выборке.

Оглавление

Критерии проверки гипотез

- Проверка статистических гипотез осуществляется с помощью статистического критерия (назовем его в общем виде К), являющегося функцией от результатов наблюдения.
- Статистический критерий - это правило (формула), по которому определяется мера расхождения результатов выборочного наблюдения с высказанной гипотезой $\mathbf{H_0}$.

Оглавление

Критерии проверки гипотез

- При проведении проверки статистических гипотез в первую очередь приходится решать задачи статистической проверки гипотез Ho:
- 1) принадлежности «выделяющихся» единиц исследуемой выборочной совокупности генеральной совокупности;
- 2) виде распределения изучаемых признаков;
- 3) величине средней арифметической и доли;
- 4) наличии и тесноте связи между изучаемыми признаками;
- 5) о форме корреляционной связи.

Оглавление

Критерии проверки гипотез

- При проверке гипотез имеется возможность совершить ошибку двоякого рода:
- а) ошибка первого рода - проверяемая гипотеза (нулевая гипотеза H_0) является в действительности верной, но результаты проверки приводят к отказу от нее;
- б) ошибка второго рода - проверяемая гипотеза в действительности является ошибочной, но результаты проверки приводят к принятию.

- В статистике в настоящее время имеется большое число критериев для проверки практически любых гипотез. Для построения статистического критерия, позволяющего проверить некоторую гипотезу, необходимо следующее:
- 1) сформулировать проверяемую гипотезу H_0. Наряду с проверяемой гипотезой формулируется также конкурирующая (альтернативная) гипотеза;
- 2) выбрать уровень значимости, контролирующий допустимую вероятность ошибки первого рода;
- 3) определить область допустимых значений и так называемую критическую область;
- 4) принять то или иное решение на основе сравнения фактического и критического значений критерия.

Оглавление

Критерии проверки гипотез

- Проверка статистических гипотез складывается из следующих этапов:
- - формулируется в виде статистической гипотезы задача исследования;
- - выбирается статистическая характеристика гипотезы;
- - выбираются испытуемая и альтернативная гипотезы на основе анализа возможных ошибочных решений и их последствий;
- - определяются область допустимых значений, критическая область, а также критическое значение статистического критерия (t, F) по соответствующей таблице;
- - вычисляется фактическое значение статистического критерия;
- - проверяется испытуемая гипотеза на основе сравнения фактического и критического значений критерия, и в зависимости от результатов проверки гипотеза либо отклоняется, либо не отклоняется.

Оглавление

Критерии проверки гипотез

- Уровнем значимости будет называться такое малое значение вероятности попадания критерия в критическую область при условии справедливости гипотезы, что появление этого события может расцениваться как следствие существенного расхождения выдвинутой гипотезы и результатов выборки. Обычно уровень значимости принимают равным 0,05 или 0,01.

Оглавление

Критерии проверки гипотез

- Принципом отношения правдоподобия, который позволяет построить критерий, наиболее мощный среди всех возможных критериев. Суть его сводится к выбору такого критерия К с известной функцией плотности f(k) при условии справедливости гипотезы Ho, чтобы при заданном условии значимости можно было бы найти критическую точку $K_{кр}$ распределения f(k), которая распределила бы область допустимых значений, в которой результаты выборочного наблюдения выглядят наиболее правдоподобными, и критическую область, в которой результаты выборочного наблюдения выглядят менее правдоподобными в отношении нулевой гипотезы H_0.

Оглавление

Критерии проверки гипотез

- Если такой критерий К выбран, и известна плотность его распределения, то задача проверки статистической гипотезы сводится к тому, чтобы при заданном уровне значимости рассчитать по выборочным данным наблюдаемое значение критерия $\mathbf{K}_{\text{набл}}$ и определить, является ли оно наиболее или наименее правдоподобным в отношении нулевой гипотезы $\mathbf{H_0}$.

Оглавление

Критерии проверки гипотез

- Рассмотрим использование критериев для проверки статистических гипотез на примере закона нормального распределения. Закон нормального распределения лежит в основе многих теорем и методов статистики
- - при оценке репрезентативности выборки (расчете ошибки выборки и распространении характеристик выборки на генеральную совокупность);
- - измерении степени тесноты связи и составлении модели регрессии;
- - построении и использование статистических критериев и др.

Оглавление

Проверка параметрических гипотез

- ***1этап.*** Пусть СВ Х имеет известный закон распределения F(x,q), где q-неизв. параметр. Требуется проверить гипотезу о том, что q равняется некоторому значению qо. эта гипотеза называется основной или нулевой. **H_0**:q=qо. кроме гипотезы **H_0** существует гипотеза **H_1** или На называемая альтернативной или конкурирующей.

- **H_1** :q^1qо, либо **H_1**:$q<$qо, либо **H_1** :$q>$qо . вид альтернативной гипотезы зависит от условий задачи.

Оглавление

Проверка параметрических гипотез.

- Сложной называется гипотеза, которая содержит несколько простых гипотез.
- $\mathbf{H_1}$:q>qo, $\mathbf{H_1}$:q=qi , qi >qo .

Оглавление

Проверка параметрических гипотез

- ***2этап***. Если отвергнута правильная гипотеза Но и принята альтернативная, то имеем ошибку первого рода. Если обозначить вероятность ошибки первого рода через α, то можно записать $P(\mathbf{H_1}) = \alpha$. α - называется уровнем значимости. Если отвергнута правильная гипотеза $\mathbf{H_1}$, а принята Но то говорят, что имеем ошибку второго рода. Обозначим вероятность ошибки второго рода через b, то можем сказать что $P\mathbf{H_1}(Ho) = b$.

$P\mathbf{H_0}(\mathbf{H_1}) = 1-\alpha . P\mathbf{H_1}(\mathbf{H_1}) = 1-b.$

Оглавление

Проверка параметрических гипотез

3этап. Критерий должен обладать свойствами:

- 1) он является функцией от наблюдаемых значений k=k(x1,…xn).

- 2) известен закон распределения критерия который табулирован если критерий имеет нормальный закон распределения то его обозначают через n. Если распределение Стюдинта то через T(t). Если распределение Фишера то через F. Если распределение с2 то через с2.

- 3) его значение позволяют судить о расхождении выборки с гипотезой $\mathbf{H_0}$.

Оглавление

Проверка параметрических гипотез

- ***4этап.*** Задают уровень значимости α=0,05; 0,01;0,001;0,005;(наиболее используемые 0,05 и 0,1) т.к. критерий k позволяет принять гипотезу $\mathbf{H_0}$, то рассматривают множество возможных значений. Из этого множества выбирают подмножества w так, чтобы выполнялось соотношение P(kÎw)= α.

Оглавление

Проверка параметрических гипотез

- Как определить критические точки k1 и k2. k1 и k2 находим по формуле: P(k1<k<k2)=ò¦(k|H0)-0|k=1- α =g. Где P(k<k1)= α /2; P(k>k2)= α /2. вычисляем значение статистики k по выборке сравнивают с критическим значение СВ величины k (это значение находим из таблицы и делаем вывод о принятии гипотезы **H_0**).

Оглавление

Критерий Пирсона, или критерий χ^2 (Хи-квадрат)

- Обозначим через X исследуемую случайную величину. Пусть требуется проверить гипотезу о том, что эта случайная величина подчиняется закону распределения . Для проверки гипотезы произведём выборку, состоящую из n независимых наблюдений над случайной величиной X.

Оглавление

Критерий Пирсона, или критерий χ^2 (Хи-квадрат)

- Для проверки критерия вводится статистика: $$\chi^2 = N \sum \frac{\left(P_i^{\text{emp}} - P_i^{\text{H}_0}\right)^2}{P_i^{\text{H}_0}},$$

где $P_i^{\text{H}_0} = F(x_i) - F(x_{i-1})$ – предполагаемая вероятность попадания в i –й интервал,

- $P_i^{\text{emp}} = \frac{n_i}{N}$ — соответствующее эмпирическое значение,
- ni — число элементов выборки из i-го интервала, N — полный объём выборки. Также используется расчет критерия по частоте, тогда:

$$\chi^2 = \sum \frac{\left(V_i - NP_i^{\text{H}_0}\right)^2}{NP_i^{\text{H}_0}},$$

- где Vi — частота попадания значений в интервал. Эта величина, в свою очередь, является случайной (в силу случайности) и должна подчиняться распределению x2.

Оглавление

Критерий Пирсона, или критерий χ^2 (Хи-квадрат)

Правило критерия:

- Если полученная статистика превосходит квантиль закона распределения х2 заданного уровня значимости α с (k-1) или с (k-p-1) степенями свободы, где k — число наблюдений или число интервалов (для случая интервального вариационного ряда), а p— число оцениваемых параметров закона распределения, то гипотеза H_0 отвергается. В противном случае гипотеза принимается на заданном уровне значимости α.

Оглавление