

ФКН МНаД: ТВиМС 2026.

Лист задач для самостоятельного решения #10.

Проверка статистических гипотез. Часть I.

1. Long time ago in a galaxy far, far away... Жители одной планеты знают, что на другой планете живут некие существа с необычно длинными шеями. Однако, из-за удаленности этой планеты и нарушенной коммуникации, научные базы данных не синхронизированы. Чтобы решить проблему команда биологов с первой планеты была отправлена заниматься научным туризмом. Ученые уверены, что средняя длина шеи этих существ — около 20 метров. Однако в ходе исследования было получено выборочное среднее 2.8 метров.

Пусть X — случайная величина, обозначающая длину шеи; дисперсию X считаем известной и верной, а сами длины — нормально распределёнными. Выборка насчитывает $n = 20$ существ при известной дисперсии $\text{Var}[X] = 0.16$. Проверьте гипотезы $H_0 : \mu = 20$ против $H_1 : \mu < 20$ на уровне значимости 1%.

Найдите границу критической области для этого теста в оригинальной шкале, т.е. при каком выборочном среднем длины шеи ученые посчитают, что им 'хватило доказательств' и будут отклонять нулевую гипотезу.

2. Исследователь проводил односторонний тест, но вместо использования (как должно было быть) верхней 5% критической точки стандартного нормального распределения, он использовал верхнюю 5% точку распределения Стьюдента с 6 степенями свободы. Каким на самом деле получился уровень значимости этого теста?
3. Случайная выборка из десяти студентов показала следующие значения времени (в часах), потраченного на изучение в неделю перед финальными экзаменами:

28 57 42 35 61 39 55 46 49 38.

Предположим, что исследуемая случайная величина имеет нормальное распределение.

- (a) Найдите выборочное среднее и выборочное стандартное отклонение.
- (b) Проверьте гипотезу о том, что математическое ожидание равно 40, против альтернативы, что оно больше.

Можно пользоваться следующим разбиением на подзадачи, чтобы лучше разобраться в теме:

- (a) Опишите распределение статистики $\bar{X}$ в этой задаче. Известна ли вам истинная дисперсия?
- (b) Сформулируйте нулевую и альтернативную гипотезы.
- (c) Опишите все случайные величины, необходимые для процедуры проверки, их свойства и распределения при нулевой гипотезе, когда мы уверены, что H_0 полностью верна.
- (d) Найдите границу критической области в оригинальной шкале и выполните проверку, используя $\alpha = 1\%, 5\%, 10\%$.
- (e) Выполните проверку с помощью правильного *score*, т.е. путём преобразования $\bar{x}$ в шкалу выбранного распределения.
- (f) Покажите, что результаты проверки в двух подходах идентичны.

4. Для людей, живущих в Калифорнии, решение о покупке страховки от землетрясений является критически важным. Статья в научном журнале от июня 1992 года исследовала множество факторов, которые жители Калифорнии учитывают при покупке страховки от землетрясений. Опрос показал, что только 133 из 337 случайно выбранных домохозяйств в округе Лос-Анджелес были защищены страховкой от землетрясений.
 - (а) Каковы подходящие нулевая и альтернативная гипотезы для проверки утверждения, что менее 40% жителей округа Лос-Анджелес были защищены страховкой от землетрясений?
 - (б) Предоставляют ли данные достаточные доказательства в поддержку нулевой гипотезы? (Используйте $\alpha = 0.10$)
5. Американская ассоциация больниц сообщает в Hospital Statistics, что средняя стоимость для общих общественных больниц на одного пациента в день в больницах США составляла \$951 в 1998 году. В том же году случайная выборка из 30 дневных затрат в больницах Нью-Йорка дала среднее значение \$1185. Предполагая истинное стандартное отклонение \$333 для больниц Нью-Йорка, предоставляют ли данные достаточные доказательства для заключения, что в 1998 году средняя стоимость в больницах Нью-Йорка превышала национальное среднее \$951? Выполните требуемую проверку гипотез на уровне значимости 5%.
6. Во время ночной смены в пятницу из производственной линии случайным образом было отобрано $n = 28$ мятных конфет и взвешено. Они имели средний вес $\bar{x} = 21.45$ граммов. Известно, что стандартное отклонение веса конфеты равно $\sigma = 0.31$ грамма.
 - (а) Проверьте нулевую гипотезу $\mu = 20$ против альтернативы $\mu > 20$ на уровне значимости 5%. Получите границу критической области в оригинальной шкале.
 - (б) Какое в данном случае получилось P -значение?
 - (с) Предположим, что вдруг *на самом деле* $\mu = 22$ (то есть верна альтернативная гипотеза). Также пусть K - граница критической области в оригинальной шкале из предыдущего пункта. Напишите, как будут выглядеть необходимые распределения статистик в таком случае (то есть распределения при $\mu = 22$). Найдите вероятность:

$$P_{H_1}(\bar{X} < K) = \beta$$

Это вероятность ошибки II рода, того, что мы в результате тестирования *не* отклоним нулевую гипотезу, когда она на самом деле *неверна*. Говоря более подробно, это вероятность такого случая, когда при верной альтернативной гипотезе мы волей случая получаем неубедительные данные, и, как следствие, не отклоняем нулевую гипотезу.