

ОП «Политология», 2017-18**Математика и статистика, часть 2****Проверка статистических гипотез. (06.06.2018)***А. А. Макаров, А. А. Тамбовцева*

Задача 1. Вася, изучив статистические методы, сконструировал свой индекс политических свобод. Случайным образом он выбрал 11 стран и выяснил, что среднее значение индекса свобод для стран в выборке равно 5, а выборочная дисперсия (несмещенная оценка) равна 9.

- a. Постройте 95%-ный доверительный интервал для среднего значения индекса политических свобод. Проинтерпретируйте его.
- b. Можно ли считать, что среднее значение индекса свобод для всех стран в мире равно 4? Сформулируйте подходящую статистическую гипотезу, которую необходимо проверить для ответа на этот вопрос. Сформулируйте одностороннюю альтернативную гипотезу (направление выберите исходя из данных). Какой критерий мы будем использовать для проверки гипотезы?
- c. Представьте, что вам необходимо построить критическую область для значений статистики выбранного критерия. Постройте критическую область, считая, что:
 - гипотеза проверяется на 10%-ном уровне значимости;
 - гипотеза проверяется на 5%-ном уровне значимости;
 - гипотеза проверяется на 1%-ном уровне значимости,
- d. Наблюдаемое значение статистики критерия оказалось равным 1.11. Какой вывод, касающийся нулевой гипотезы, можно сделать?
- e. Представьте, что теперь эту задачу нам пришлось бы решать, используя p-value. Что такое p-value в нашей конкретной задаче?
- f. Запишите, как бы выглядело p-value в этом случае. Давайте найдем его, используя R! Какой содержательный вывод, исходя из p-value, мы можем сделать (рассмотрите все три уровня значимости, предложенные в задаче)?
- g. Можно ли проверить нулевую гипотезу, используя доверительный интервал? Если да, то как?

Задача 2. Студент проверял гипотезу о равенстве доли сторонников консервативной партии в стране K 0.56. Наблюдаемое значение z-статистики оказалось равным 1.72.

- a. Постройте критическую область для статистики критерия, принимая уровень значимости равным: а) 5%. б) 10%. Есть ли основания отвергнуть нулевую гипотезу о равенстве доли сторонников консервативной партии 0.56 на 5%-ном уровне значимости? А на 10%-ном уровне значимости?
- b. Посчитайте p-value. Объясните, какой содержательный смысл имеет это значение, и каким должно оно быть, чтобы гипотеза о равенстве доли сторонников консерваторов не отвергалась на 5%-ном уровне значимости.

Задача 3. (R)

- a. Проверьте гипотезу о равенстве среднего возраста выживших пассажиров «Титаника» 30 годам: а) на 10%-ном уровне значимости; б) на 5%-ном уровне значимости.
- b. Можно ли утверждать, что средний возраст выживших пассажиров женского и мужского пола одинаков? Сформулируйте соответствующую задаче нулевую гипотезу, альтернативную гипотезу. Проверьте ее на 5%-ном уровне значимости. Сделайте статистический и содержательный вывод.

Задача 4. (R)

- а. Проверьте гипотезу о равенстве среднего возраста выживших пассажиров «Титаника» 30 годам: а) на 10%-ном уровне значимости; б) на 5%-ном уровне значимости.
- б. Можно ли утверждать, что средний возраст выживших пассажиров женского и мужского пола одинаков? Сформулируйте соответствующую задаче нулевую гипотезу, альтернативную гипотезу. Проверьте ее на 5%-ном уровне значимости. Сделайте статистический и содержательный вывод.

Задача 5. (R) Загрузите базу из файла *elect18.csv*, в котором сохранены результаты президентских выборов 2018 года по регионам.

- а. Постройте диаграмму рассеяния для переменных *Явка* (*turnout_perc*) и процент за Зюганова (*Zu*). Основываясь на графике, что можно сказать о характере связи между этими показателями? Прделайте то же самое для переменных *Явка* (*turnout_perc*) и процент за Путина (*Pu*).
- б. Посчитайте для этой пары показателей коэффициент корреляции Пирсона. Проинтерпретируйте его значение. Совпало ли оно с вашими ожиданиями, основанными на графике?
- с. Проверьте, является ли коэффициент корреляции статистически значимым. Сформулируйте нулевую и альтернативные гипотезы. Проверьте H_0 в R. Сделайте статистический и содержательный вывод.