

**Школа лингвистики, 2023-24 уч. год****Теория вероятностей****Z-тест и тест Стьюдента (9.04.2024)***И. В. Щуров, Д. А. Филимонов, Р. Я. Будылин*

**Задача 1.** С целью уменьшения количества прогулов, компания вводит режим плавающего графика. В прошлом (до ввода нового режима), каждый сотрудник прогуливал (отпрашивался, опаздывал и т.д.) в среднем 6.3 дня в год. Стандартное отклонение составляло 2.9 дней. После ввода нового режима проходит год и менеджмент пытается подвести итоги и понять, достигнута ли заявленная цель. Для этого случайно выбирается 100 сотрудников и запрашиваются данные по их прогулам. Выясняется, что среднее число прогулов снизилось до 5.5 дней, стандартное отклонение считаем неизменным. Можно ли на основании этих данных утверждать, что свободный график действительно поспособствовал уменьшению числа прогулов или полученный результат объясняется случайностью? Считаем, что рассматриваемый год ничем не отличается от предыдущих (нет других факторов, которые бы влияли на прогулы — например, неверно, что в этом году в среднем была более благоприятная для прогулов погода или экономическая ситуация). Сформулировать соответствующую нулевую гипотезу и альтернативу и применить адекватный статистический тест. Какой вариант теста логичнее брать: двусторонний, правосторонний или левосторонний? Проверить нулевую гипотезу на уровнях значимости  $\alpha = 5\%$  и  $\alpha = 1\%$ .

**Задача 2.** Решить задачу 1 в предположении, что исследуемая выборка состояла лишь из одного человека.

**Задача 3.** Решить задачу 1 в предположении, что стандартное отклонение составляло 29 дней.

**Задача 4.** Решить задачу 1 в предположении, что полученное после введения свободного графика среднее составляет 5.8 дней.

**Задача 5.** Решить задачу 1 в предположении, что стандартное отклонение считается неизвестным, а выборка состояла из 30 сотрудников. При этом несмещённая оценка дисперсии по выборке оказалась равной 6.25. Уровень значимости принять равным 5%.

**Задача 6.** Для оценки степени владения языком может использоваться оценка сложности, например количество именных групп на одно высказывание<sup>1</sup>. Для каждого испытуемого такая величина усредняется и потому ее можно считать нормально распределенной.

- (a) Пусть известно, что для взрослых в среднем этот показатель равен 1.7. Для нескольких детей одной возрастной группы получены следующие значения: 1.7; 1.9; 1.1; 1.4; 1.2. Можно ли считать, что эти дети владеют языком так же хорошо, как взрослые? Выбрать соответствующий критерий и правильную (двустороннюю или одностороннюю) альтернативу. Рассмотреть уровни значимости 10% и 5%
- (b) Пусть известно среднее для возрастной группы детей 8-9 лет и этот показатель равен 1.4. У вас есть выборка детей этой же возрастной группы, но из одного конкретного типа семей (нет понимания о влиянии этого фактора на развитие речи): 1.8; 1.9; 1.2; 2.7; 2.3. Можно ли считать, что эти дети владеют языком так же хорошо, как в общем дети их возраста? Выбрать соответствующий критерий и правильную

<sup>1</sup>Такую оценку использовали авторы работы [Huttenlocher et al., Cognitive Psychology, Vol. 45, no 3, 2002]

- (двустороннюю или одностороннюю) альтернативу. Рассмотреть уровни значимости 10% и 5%
- (с) Что если в условиях пункта б) добавить информацию, что это дети из семей, где хотя бы один родитель имеет высшее филологическое или лингвистическое образование? Как изменится ответ?