

Контрольная работа по статистике

Задание 1. Аудитор заинтересовался средним балансом банковского счета. Он случайно выбрал 75 банковских счетов и получил, что средний баланс составляет \$128. Считая, что баланс счетов имеет нормальное распределение со стандартным отклонением \$27, постройте 90%-й доверительный интервал для среднего значения баланса.

Задание 2. Университетская футбольная команда сыграла 28 матчей с разными командами и проиграла 12 из них. Постройте 90%-й доверительный интервал для доли проигрышей.

Задание 3. Менеджер уверял, что снабжение налажено без перебоев и запасы топлива на всех складах практически одинаковые, “стандартное отклонение 7 тонн, не больше” — утверждал он. Однако, у директора возникло подозрение, запасы на разных складах слишком сильно отличаются и он устроил ревизию. Для случайно выбранных 25 складов выборочное стандартное отклонение запасов составило 8.5 тонн. Предполагая, что распределение запасов топлива нормальное и одинаковое для всех складов, а также, что запасы на разных складах независимы, проверьте заявление менеджера как статистическую гипотезу с 5%-м уровнем значимости.

Задание 4. При дегустации кофе 235 человек из 400 предпочли кофе лидирующей на рынке марки. Проверьте с 1%-м уровнем значимости гипотезу, что больше 52 процентов людей предпочитают именно эту марку кофе.

Задание 5. Есть несколько различных наборов упаковок M&Ms. Они отличаются количеством драже различного цвета (пропорции драже каждого цвета для каждого набора приведены в таблице). При расфасовке драже сначала смешиваются в нужной пропорции, а потом насыпаются в упаковки по весу.

Открыв упаковку и пересчитав драже, мы обнаружили: 15 голубых, 32 коричневых, 7 зеленых, 19 оранжевых, 29 красных и 24 желтых.

С помощью критерия χ^2 -Пирсона, проверьте гипотезу (с уровнем значимости 5%) о том, что наша упаковка была из набора “Milk chocolate”. Повторите проверку для набора “Peanut”. Как вы думаете, из какого набора скорее всего была упаковка?

	голубые	коричневые	зеленые	оранжевые	красные	желтые
Milk chocolate	1	3	1	1	2	2
Peanut	2	2	1	1	2	2
Peanut butter	1	1	1	0	1	1
Almond	1	1	1	1	1	1

Критерий согласия χ^2 -Пирсона

Критерий согласия Пирсона проверяет гипотезу о том, что данная выборка имеет определенное **дискретное** распределение (критерий Колмогорова-Смирнова решает ту же задачу для непрерывного распределения).

Итак, пусть дана выборка X_1, \dots, X_N дискретных значений из множества $\{1, \dots, K\}$. Мы хотим проверить гипотезу H_0 о том, что распределение выборки X — дискретное распределение на $\{1, \dots, K\}$ с вероятностями $\{p_1, \dots, p_K\}$.

Известен следующий факт (теорема Пирсона): если нулевая гипотеза верна, то статистика

$$t_{\text{Pearson}} = \sum_{i=1}^K \frac{(n_i - Np_i)^2}{Np_i},$$

где n_i — количество элементов из X , равных i , имеет асимптотически распределение χ^2 с $K - 1$ -й степенями свободы.

Если же нулевая гипотеза неверна, то статистика t_{Pearson} (она имеет смысл расстояния между распределениями) стремится к бесконечности.

Таким образом получаем критерий:

H_0 : X имеет распределение $\{p_1, \dots, p_K\}$,

H_1 : X имеет другое распределение.

Критическая область: $S = (\chi_{1-\alpha, K-1}^2, \infty)$, где $\chi_{1-\alpha, K-1}^2$ — $1 - \alpha$ -квантиль распределения χ^2 с $K - 1$ -й степенью свободы.