

**Математика. Углубленный уровень. 2016 год. День 3. Статистика и вероятность.**

*Работа рассчитана на 1 час.*

*Для выполнения работы требуется графический калькулятор и буклет с формулами.*

*Все числовые ответы должны быть точными или даны с тремя значащими цифрами, если в задании не указано обратное.*

*Максимальный балл за эту часть работы составляет 60 баллов.*

*Верный ответ не является единственным условием получения максимального балла за задание. Ответы должны быть подкреплены решениями и пояснениями. В случае неверного ответа за задание можно получить некоторые баллы в случае наличия в работе верного метода решения. Поэтому рекомендуется записывать тексты решения полностью. Если при решении используется графический калькулятор, то это должно быть отражено в работе. Например, если для нахождения решения используется график, то его эскиз должен быть нарисован в работе.*

**Задача 1 (12 баллов).** Адам каждый день решает кроссворд в местной газете. Время  $X$  минут, необходимое Адаму для решения кроссворда, имеет нормальное распределение  $N(22, 5^2)$ .

**а) (3 балла)** Найдите  $a$  для которого в случайно выбранный день вероятность решить кроссворд быстрее чем за  $a$  минут равна 0,8.

**б) (3 балла)** Найдите вероятность того, что общее время, которое ему понадобится, чтобы решить пять случайных кроссвордов подряд, превышает 120 минут.

Беатрис также решает кроссворды в местной газете каждый день. Время  $Y$  минут, необходимое Беатрис для решения кроссворда, имеет нормальное распределение  $N(40, 6^2)$ .

**в) (6 баллов)** Найдите вероятность того, что в случайный день время, которое Беатрис затратит на решение кроссворда окажется более чем вдвое больше, чем время, которое потребуется Адаму. Считайте, что эти два времени независимы.

**Задача 2 (10 баллов).** Случайные величины  $X, Y$  имеют двумерное нормальное распределение с линейным коэффициентом корреляции  $\rho$ .

**а) (2 балла)** Сформулируйте подходящие гипотезы для исследования, являются ли  $X$  и  $Y$  независимыми.

Была получена случайная выборка из 10 наблюдений для  $X, Y$ , для нее был вычислен линейный коэффициент корреляции  $r$  равный 0,486.

**б) (7 баллов)** Определите  $P$ -значение.

Сформулируйте свое заключение с уровнем значимости 5%.

**в) (1 балл)** Объясните, почему уравнение линии регрессии  $y$  от  $x$  не должно использоваться для прогнозирования значения  $y$  соответствующего  $x = x_0$ , где  $x_0$  лежит в диапазоне значений выборки.

**Задача 3 (15 баллов).** Непрерывная случайная величина  $X$  принимает значения в интервале  $[0; \theta]$  с  $E(X) = \frac{\theta}{2}$  и  $Var(X) = \frac{\theta^2}{24}$ . Для определения неизвестного параметра  $\theta$  была сделана случайная выборка размера  $n$  из распределения  $X$ . Выборочное среднее обозначено  $\bar{X}$ ,  $U = k\bar{X}$  — несмещенная оценка для  $\theta$ .

**а) (3 балла)** Найдите значение  $k$ .

- б) (4 балла) Вычислите несмещенную оценку для  $\theta$ , используя случайную выборку

8.3, 4.2, 6.5, 10.3, 2.7, 1.2, 3.3, 4.3.

Дайте краткое объяснение, почему это не очень хорошая оценка для  $\theta$ .

- в) (8 баллов) Покажите, что  $Var(X) = \frac{\theta^2}{6n}$ .

Покажите, что  $U^2$  не является несмещенной оценкой для  $\theta^2$ .

Найдите несмещенную оценку для  $\theta^2$  через  $U$  и  $n$ .

**Задача 4 (11 баллов).** Владельца завода просят производить кирпич массой 2,2 кг. Менеджер по контролю качества хочет проверить, является ли средний вес кирпичей, произведенных в некоторый день, равным 2,2 кг.

- а) (2 балла) Сформулируйте гипотезы, позволяющие менеджеру по контролю качества проверить средний вес с помощью двухстороннего теста.

Менеджер берет случайную выборку из 20 кирпичей и определяет массу  $x$  кг каждого кирпича. В результате он получает следующую итоговую статистику  $\sum x = 42,0$ ,  $\sum x^2 = 89,2$ .

- б) (7 баллов) Вычислите несмещенные оценки для математического ожидания и дисперсии массы произведенных кирпичей.

Предполагая, что масса кирпичей имеет нормальное распределение, определите  $P$ -значение для результатов выше и сформулируйте вывод с точки зрения уровня значимости в 5%.

- в) (2 балла) Владелец лучше знаком с доверительными интервалами. Определите 95%-ный доверительный интервал для средней массы кирпичей, произведенных в этот день.

**Задача 5 (12 баллов).** Непрерывная случайная величина  $X$  имеет функцию распределения

$$f = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

Дискретная случайная величина  $Y$  определяется как целая часть от  $X$ , т.е. как наибольшее целое, не превосходящее  $X$ .

- а) (4 балла) Покажите, что распределение вероятности  $Y$  задается равенством  $P(Y = y) = e^{-y}(1 - e^{-1})$ ,  $y \in \mathbb{N}$ .

- б) (8 баллов) Покажите, что производящая функция вероятностей  $G(t)$  для  $Y$  задается равенством  $G(t) = \frac{1 - e^{-1}}{1 - e^{-1}t}$ .

Определите отсюда  $E(Y)$  с точностью до трех значащих цифр.