Дискретные случайные величины

Случайная величина - это величина, которая в результате опыта со случайным исходом принимает то или иное числовое значение, причем заранее неизвестно, какое именно. В случае конечного или счетного количества исходов случайная величина ξ называется дискретной.

Если известны значения $x_1, x_2, \ldots, x_N, \ldots$, которые принимает случайная величина ξ, а также вероятности $p_i = P\{\xi = x_i\},\ i = 1, \ldots, N, \ldots,\ p_1 + \ldots + p_N + \ldots = 1$, то говорят, что задан ее дискретный закон распределения

$$\xi : \begin{pmatrix} x_1 & x_2 & \ldots & x_N \\ p_1 & p_2 & \ldots & p_N \end{pmatrix}. \tag{1}$$

Распределение Бернулли

- При подбрасывании монеты возможны два исхода: успех или неудача, которые кодируются соответственно 1 и 0. Успех обычно ассоциируется с выпадением монеты строной "орел", которая выпадает с вероятностью $p = 0.5$. Неудача ассоциируется с выпадением стороны "решка", которая выпадает с вероятностью $q = 1 - p = 0.5$.
- Если при подбрасывании игрального кубика успехом считать выпадении шести очков, то $p = \frac{1}{6}$, $q = \frac{5}{6}$.

Случайная величина, принимающая два значения с вероятностями p и $q = 1 - p$, имеет распределение Бернулли.

Задача де Мере

Первое пари

Шевалье де Мере предлагал заключить пари на то, что за 4 подбрасывания игральной кости хотя бы один раз кубик выпадет стороной с шестью очками (событие B). Противоположное событие $\bar{B}$ означает, что за четыре попытки сторона с шестью очками так и не появится.

$$P(\bar{B}) = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = 0.48, \quad P(B) = 1 - 0.48 = 0.52.$$

Второе пари

При 24-х подбрасываниях двух кубиков хотя бы раз выпадут две шестерки (событие B).

$$P(\bar{B}) = \left(\frac{35}{36}\right)^{24} \approx 0.51, \quad P(B) = 1 - 0.51 = 0.49.$$

Биномиальный закон распределения

Независимость случайных величин

Случайные величины ξ со значениями $x_1, x_2, \ldots$ и η со значениями $y_1, y_2, \ldots$ независимы, если для любых i и j

$$P\{\xi = x_i, \eta = y_j\} = P\{\xi = x_i\} P\{\eta = y_j\}. \tag{2}$$

Случайная величина ξ, равная числу успехов в n независимых испытаниях с вероятностью успеха p, имеет биномиальный закон распределения

$$P_k = P\{\xi = k\} = C_n^k p^k (1 - p)^{n-k}, \; k = 0, 1, \ldots, n. \tag{3}$$

$P_n = p^n, \quad P_0 = q^n$

$P_1 = pq^{n-1} + qpq^{n-2} + q^2pq^{n-3} + \ldots + q^{n-1}p = C_n^1 p^1 q^{n-1}$

$P_{n-1} = qp^{n-1} + pqp^{n-2} + p^2qp^{n-3} + \ldots + p^{n-1}q = C_n^{n-1} p^{n-1} q^1$

Биномиальный закон распределения (пример)

число успехов k	варианты комбинаций	вероятность p_k
0	○ ○ ○○	$C_4^0 p^0 q^4 = \left(\frac{5}{6}\right)^4$
1	● ○ ○○ ○ ● ○○ ○ ○ ●○ ○ ○ ○●	$C_4^1 p^1 q^3 = 4 \cdot \frac{1}{6} \left(\frac{5}{6}\right)^3$
2	● ● ○○ ● ○ ●○ ● ○ ○● ○ ● ●○ ○ ● ○● ○ ○ ●●	$C_4^2 p^2 q^2 = 6 \cdot \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2$
3	● ● ●○ ● ● ○● ● ○ ●● ○ ● ●●	$C_4^3 p^3 q^1 = 4 \cdot \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^1$
4	● ● ●●	$C_4^4 p^4 q^0 = \left(\frac{1}{6}\right)^4$

Table: Биномиальный закон распределения при $n = 4$ и $p = 1/6$:
$P_0 = 0.482, P_1 = 0.386, P_2 = 0.116, P_3 = 0.015, P_4 = 0.001$

Наивероятнейшее число событий k_0

$P_n(k_0-1) \le P_n(k_0)$ и $P_n(k_0) \ge P_n(k_0+1)$. Из первого имеем $C_n^{k_0-1}p^{k_0-1}q^{n-k_0+1} \le C_n^{k_0}p^{k_0}q^{n-k_0} \iff$

$$\frac{n!}{(k_0-1)!(n-k_0+1)!}p^{k_0-1}q^{n-k_0+1} \le \frac{n!}{k_0!(n-k_0)!}p^{k_0}q^{n-k_0} \iff$$

$$\frac{q}{(n-k_0+1)} \le \frac{p}{k_0} \iff k_0 q \le p(n-k_0+1),$$

отсюда $k_0 \le p(n+1)$. Из второго неравенства имеем $C_n^{k_0}p^{k_0}q^{n-k_0} \ge C_n^{k_0+1}p^{k_0+1}q^{n-k_0-1} \iff$

$$\frac{n!}{k_0!(n-k_0)!}p^{k_0}q^{n-k_0} \ge \frac{n!}{(k_0+1)!(n-k_0-1)!}p^{k_0+1}q^{n-k_0-1} \iff$$

$$\frac{q}{(n-k_0)} \ge \frac{p}{(k_0+1)} \iff (k_0+1)q \ge p(n-k_0),$$

отсюда получаем второе неравенство $k_0 \ge np-q$. Таким образом,

$$np-q \le k_0 \le np+p.$$

Кубик подбрасывается $n = 16$ раз. Найти наивероятнейшее число появления очков, кратного трем, $p = \frac{1}{3}$.

$$\frac{16}{3} - \frac{2}{3} \leq k_0 \leq \frac{16}{3} + \frac{1}{3} \iff \frac{14}{3} \leq k_0 \leq \frac{17}{3}, \implies k_0 = 5.$$

Сколько раз необходимо бросить игральную кость, чтобы с вероятностью 0.9 шесть очков выпало хотя бы один раз?

Вероятность того, что в n испытаниях шесть очков выпадут хотя бы раз равна $1 - \left(\frac{5}{6}\right)^n$.

$$1 - \left(\frac{5}{6}\right)^n > 0.9 \iff \left(\frac{5}{6}\right)^n < \frac{1}{10} \iff \left(\frac{6}{5}\right)^n > 10,$$

отсюда $n > \frac{\log 10}{\log 6 - \log 5} = 12.63$.

Задачи на биномиальное распределение

В походе участвуют 12 байдарок. На пути встречается сложный порог. Проходить порог решаются с вероятностью 0.55 и с вероятностью 0.2 его проходят. Чему равна вероятность того, что порог благополучно пройдут не менее 5 байдарок?

Вероятность пройти порог равна $p = 0.55 \cdot 0.2 = 0.11$. Ответ $\sum_{k=5}^{12} C_{12}^{k} 0.11^{k} 0.89^{12-k} = 0.0065$.

Футбольная команда выигрывает на своем поле с вероятностью 0.51, а на чужом с вероятностью 0.43. Чему равна вероятность того, что в серии из 8 игра (4 на своем и 4 на чужом поле) 5 побед будет за этой командой?

$$\sum_{k=1}^{4} C_4^k 0.51^k 0.49^{4-k} C_4^{5-k} 0.43^{5-k} 0.57^{4-(5-k)} = 0.1914$$

Число $e = 2.718282\ldots$

$$e = \lim_{n\to\infty}(1+\tfrac{1}{n})^n = \lim_{p\to 0}(1+p)^{\frac{1}{p}}$$

Пример с начислением процентов. Если на один рубль 100% начисляется один раз в год, то в получаем $(1+1)^1 = 2$.
Если два раза в год, то $1+\frac{1}{2}+\frac{1}{2}(1+\frac{1}{2}) = (1+\frac{1}{2})^2 = 2.25$
Если три раза в год, то
$(1+\frac{1}{3})+\frac{1}{3}(1+\frac{1}{3})+\frac{1}{3}((1+\frac{1}{3})+\frac{1}{3}(1+\frac{1}{3})) =$
$(1+\frac{1}{3})^2+\frac{1}{3}(1+\frac{1}{3})^2 = = (1+\frac{1}{3})^3 = 2.370370$
Если 4 раза в год, то $(1+\frac{1}{4})^4 = 2.441406$
$(1+\frac{1}{100})^{100} = 2.704814$
$(1+\frac{1}{365})^{365} = 2.714567$

$$\lim_{p\to 0}(1-p)^{\frac{1}{p}} = e^{-1}$$

Выражение числа e через степенной ряд

$$e^{\lambda} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}$$

$$e^{\lambda} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^{n\lambda} = \lim_{n\to\infty} \left(1 + C_{n\lambda}^{1} \cdot \frac{1}{n} + C_{n\lambda}^{2} \cdot \frac{1}{n^2} + \ldots\right) =$$

$$= \lim_{n\to\infty} \left(1 + \frac{n\lambda}{n} + \frac{n\lambda(n\lambda - 1)}{2!n^2} + \ldots\right) = 1 + \lambda + \frac{\lambda^2}{2!} + \ldots$$

Основное свойство числа e

Инвариантность относительно дифференцирования.

$$(\mathrm{e}^{\mathrm{x}})' = \mathrm{e}^{\mathrm{x}}.$$

$$(\mathrm{e}^{\mathrm{x}})' = \lim_{\Delta\to 0} \frac{\mathrm{e}^{\mathrm{x}+\Delta} - \mathrm{e}^{\mathrm{x}}}{\Delta} = \lim_{\Delta\to 0} \frac{\mathrm{e}^{\mathrm{x}}(\mathrm{e}^{\Delta} - 1)}{\Delta} =$$

$$= \lim_{\Delta\to 0} \frac{\mathrm{e}^{\mathrm{x}}}{\Delta} \left(1 + \Delta + \frac{\Delta^2}{2!} + \frac{\Delta^3}{3!} + \ldots - 1\right) = \mathrm{e}^{\mathrm{x}}\,.$$

Распределение Пуассона

Вычислим вероятность P_k при $p \to 0$ и $n \to \infty$.

$$P_k = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k},$$

умножим и разделим P_k на n^k обозначим через $\lambda = np$.

$$P_k = \frac{n(n-1)\dots(n-k+1)}{k! n^k} (np)^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}, \qquad (4)$$

так как $e^{-1} = \lim\limits_{p \to 0} (1-p)^{\frac{1}{p}}$, при заданном k

$$\lim_{p \to 0} (1-p)^{n-k} = \lim_{p \to 0} \left((1-p)^{\frac{1}{p}} \right)^{pn} (1-p)^{-k} = e^{-\lambda},$$

$$\lim_{n \to \infty} \frac{n(n-1)\dots(n-k+1)}{n^k} = 1\,.$$

Случайная величина ξ, принимающая значения $k = 0, 1, 2, \dots$ с вероятностями (4), имеет распределение Пуассона. Так как $e^{\lambda} = \sum\limits_{k=0}^{\infty} \frac{\lambda^k}{k!}$, то $\sum\limits_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1\,.$

Задачи на распределение Пуассона

Известно, что 5% мужчин дальтоники. Какова вероятность того, что из 40 мужчин нет ни одного дальтоника?

$\mathrm{p} = 0.05$, $\mathrm{n} = 40$, $\lambda = 0.05 \cdot 40 = 2$. $\mathrm{P}_0 = \mathrm{e}^{-2} = 0.135$.

Человек, зашедший в магазин, покупает дорогую вещь с вероятностью 0.0055. Какова вероятность, что из 600 посетителей по крайней мере трое купят дорогую вещь?

$\mathrm{p} = 0.0055$, $\mathrm{n} = 600$, $\lambda = \mathrm{np} = 0.0055 \cdot 600 = 3.3$

$1 - \mathrm{P}_0 - \mathrm{P}_1 - \mathrm{P}_2 = 1 - \mathrm{e}^{-3.3}(1 + 3.3 + \frac{3.3^2}{2}) = 1 - 0.3594 = 0.6406$

Характеристики распределения: математическое ожидание

$$\xi : \begin{pmatrix} x_1 & \dots & x_n & \dots \\ p_1 & \dots & p_n & \dots \end{pmatrix}, \quad \sum_{k=1}^{n} p_k = 1\,. \tag{5}$$

Математическое ожидание (МО) или среднее значение дискретной случайной величины (5) имеет вид

$$\mathbb{E}\xi = \sum_{k=1}^{\infty} x_k p_k$$

Например, МО случайной величины, имеющей распределение Бернулли, равно $0 \cdot q + 1 \cdot p = p$, а для биномиального распределения (3)

$$\mathbb{E}\xi = \sum_{k=0}^{n} \frac{kn!}{k!(n-k)!} p^k q^{n-k} = np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} = np.$$

Математическое ожидание линейной комбинации

$$\mathbb{E}(a\xi + b) = \sum_{i=1}^{\infty}(ax_i + b)p_i = a\sum_{i=1}^{\infty}x_i p_i + \sum_{i=1}^{\infty} bp_i = a\mathbb{E}\xi + b. \quad (6)$$

МО суммы двух случайных величин ξ и η равно сумме их математических ожиданий.

$$\mathbb{E}(\xi + \eta) = \sum_i \sum_j (x_i + y_j)P\{\xi = x_i, \eta = y_j\} =$$

$$= \sum_i \sum_j x_i P\{\xi = x_i, \eta = y_j\} + \sum_i \sum_j y_j P\{\xi = x_i, \eta = y_j\} =$$

$$= \sum_i x_i \sum_j P\{\xi = x_i, \eta = y_j\} + \sum_j y_j \sum_i P\{\xi = x_i, \eta = y_j\} =$$

$$= \sum_i x_i P\{\xi = x_i\} + \sum_j y_j P\{\eta = y_j\} = \mathbb{E}\xi + \mathbb{E}\eta\,.$$

МО произведения двух независимых (2) случайных величин ξ и η равно произведению их математических ожиданий.

$$\mathbb{E}\xi\eta = \sum_{\mathrm{i}}\sum_{\mathrm{j}} \mathrm{x_i y_j P}\{\xi = \mathrm{x_i}, \eta = \mathrm{y_j}\} =$$

$$= \sum_{\mathrm{i}}\sum_{\mathrm{j}} \mathrm{x_i y_j P}\{\xi = \mathrm{x_i}\}\mathrm{P}\{\eta = \mathrm{y_j}\} =$$

$$= \sum_{\mathrm{i}} \mathrm{x_i P}\{\xi = \mathrm{x_i}\} \sum_{\mathrm{j}} \mathrm{y_j P}\{\eta = \mathrm{y_j}\} = \mathbb{E}\xi\mathbb{E}\eta\,.$$

Характеристика вариабельности данных дисперсия σ^2

$$\sigma^2 = \mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{E}(\xi^2 - 2\xi\mathbb{E}\xi + (\mathbb{E}\xi)^2) = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2\,, \quad (7)$$

σ стандартное отклонение
Если ξ имеет распределение Бернулли, то

$$\mathbb{E}\xi^2 = 0^2 \cdot \mathrm{q} + 1^2 \cdot \mathrm{p} = \mathrm{p},$$
$$\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \mathrm{p} - \mathrm{p}^2 = \mathrm{pq}\,.$$

Свойства дисперсии

Дисперсия суммы двух случайных величин ξ_1, ξ_2 равна

$$\begin{aligned} \mathrm{D}(\xi_1+\xi_2) = \mathbb{E}((\xi_1+\xi_2)-\mathbb{E}(\xi_1+\xi_2))^2 = \\ = \mathbb{E}((\xi_1-\mathbb{E}\xi_1)+(\xi_2-\mathbb{E}\xi_2))^2 = \\ = \mathbb{E}(\xi_1-\mathbb{E}\xi_1)^2+\mathbb{E}(\xi_2-\mathbb{E}\xi_2)^2+2\underbrace{\mathbb{E}(\xi_1-\mathbb{E}\xi_1)(\xi_2-\mathbb{E}\xi_2)}_{\mathrm{cov}(\xi_1,\xi_2)} \end{aligned}$$

Если величины ξ_1, ξ_2 независимы, то ковариация

$\mathrm{cov}(\xi_1,\xi_2) = \mathbb{E}(\xi_1-\mathbb{E}\xi_1)(\xi_2-\mathbb{E}\xi_2) = \mathbb{E}(\xi_1-\mathbb{E}\xi_1)\mathbb{E}(\xi_2-\mathbb{E}\xi_2) = 0.$

дисперсия величины $\xi_1+\ldots+\xi_n$ равна npq.

Свойства дисперсии

Дисперсия линейного преобразованияи $a\xi + b$ не зависит от сдвига и отличется на a^2

$$\mathbb{D}(a\xi + b) = \mathbb{E}(a\xi + b - \mathbb{E}(a\xi + b))^2 = a^2\mathbb{E}(\xi - \mathbb{E}\xi)^2 = a^2\mathbb{D}\xi$$

Заметим, что дисперсия константы равна нулю.

МО и дисперсия распределения Пуассона

$$p_k = P\{\xi = k\} = \tfrac{\lambda^k}{k!} e^{-\lambda}$$

Математическое ожидание и дисперсия случайной величины ξ равны λ, так как

$$\mathbb{E}\xi = \sum_{k=0}^{\infty} k p_k = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda\,.$$

$$\mathbb{E}\xi^2 = \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=1}^{\infty} \frac{k\lambda^k}{(k-1)!} e^{-\lambda} \overset{k=t+1}{=} \sum_{t=0}^{\infty} \frac{(t+1)\lambda^{t+1}}{t!} e^{-\lambda} = \lambda^2 + \lambda,$$

$$\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda\,.$$