ТВиМС

Содержание

1 Случайные величины

Определение 1. Случайной величиной ξ называется функция, заданная на множестве Ω, принимающая значения в $\mathbb{R}$.

Задать случайную величину, значит указать все ее реализации и соответственные вероятности.

Определение 2. Индикатором события A называется случайная величина:

$$\mathbb{I}(A) \sim \begin{pmatrix} 0 & 1 \\ 1 - \mathbb{P}(A) & \mathbb{P}(A) \end{pmatrix}$$

Определение 3. Законом распределения случайной величины называется некоторое правило, позволяющее однозначно определить значение вероятности по значению случайной величины.

1.1 Числовые характеристики случайных величин

Определение 4. Математическим ожиданием дискретной случайной величины, если оно существует, называется число:

$$\mathbb{E}(\xi) = \sum_{i=1}^{n} \omega_i \cdot \mathbb{P}(\xi = \omega_i)$$

Определение 5. Дисперсией случайной величины называется $\mathbb{D}(\xi) = \mathbb{E}(\xi - \mathbb{E}(\xi))^2$.

Теорема 1.1. $\mathbb{E}(a\xi + b\eta + c) = a\mathbb{E}(\xi) + b\mathbb{E}(\xi) + c;\ a, b, c \in \mathbb{R}$

Доказательство.

$$\begin{aligned} &\mathbb{E}(a\xi + b\eta + c) = \\ =&\sum_{i=1}^{n} \widehat{\omega}_i \cdot \mathbb{P}(a\xi + b\eta + c = \widehat{\omega}_i) = \\ =&c + \sum_{i=1}^{n} \widehat{\omega}_i^c \cdot \mathbb{P}(a\xi + b\eta = \widehat{\omega}_i^c) = \\ =&c + \sum_{i=1}^{n} \omega_i^\xi \cdot \mathbb{P}(a\xi = \omega_i^\xi) + \sum_{i=1}^{n} \omega_i^\eta \cdot \mathbb{P}(b\eta = \omega_i^\eta) = \\ =&c + a\mathbb{E}(\xi) + b\mathbb{E}(\eta) \end{aligned}$$

∎

Теорема 1.2. Дисперсия случайной величины ξ может быть вычислена, как $\mathbb{D}(\xi) = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2$

Доказательство.

$$\begin{aligned} &\mathbb{D}(\xi) = \mathbb{E}(\xi - \mathbb{E}(\xi))^2 = \\ =&\mathbb{E}(\xi^2 - 2\xi\mathbb{E}(\xi) + (\mathbb{E}(\xi))^2) = \\ =&\mathbb{E}(\xi^2) - 2(\mathbb{E}(\xi))^2 + (\mathbb{E}(\xi))^2 = \\ =&\mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 \end{aligned}$$

∎

Определение 6. Стандартным отклонением случайной величины ξ называется $\sigma(\xi) = \sqrt{\mathbb{D}(\xi)}$.

1.1.1 Распределение Бернулли

Определение 7. Случайная величина ξ распределена по Бернулли, если ее распределение суть индикатор.

$$Ber(p) \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

1.1.2 Биномиальное распределение

Определение 8. Случайная величина ξ распределена биномиально, если она моделирует схему испытаний Бернулли или является суммой бернулиевых случайных величин.

$$B(p,\, n) \sim \begin{pmatrix} 0 & \cdots & k & \cdots & n \\ (1-p)^n & \cdots & C_n^k p^k (1-p)^{n-k} & \cdots & p^n \end{pmatrix}$$

Теорема 1.3. Математическое ожидание биномиально распределенной случайной величины ξ может быть вычислено, как $\mathbb{E}(\xi) = np$.

Доказательство.

$$\begin{aligned} \mathbb{E}(\xi) = & C_n^1 p q^{n-1} + 2C_n^2 p^2 q^{n-2} + \ldots + kC_n^k p^k q^{n-k} + \ldots + nC_n^n p^n = \\ = & np \cdot (C_{n-1}^0 q^{n-1} + C_{n-1}^1 p q^{n-2} + \ldots + C_{n-1}^{k-1} p^{k-1} q^{n-k} + \ldots + C_{n-1}^{n-1} p^{n-1}) = \\ = & np \cdot (q+p)^{n-1} = \\ = & np \end{aligned}$$

■

Теорема 1.4. Дисперсия независимых случайных величин линейна: $\mathbb{D}(\xi + \eta) = \mathbb{D}(\xi) + \mathbb{D}(\eta)$

Лемма 1.5. Дисперсия биномиально распределенной случайной величины ξ может быть вычислена, как $\mathbb{D}(\xi) = npq$.

Доказательство. Пусть η – число успехов в одном испытании Бернули. Тогда:

$$\eta \sim B(p,\, 1) \sim \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$$

В таком случае $\mathbb{D}(\eta) = \mathbb{E}(\eta^2) - (\mathbb{E}(\eta))^2 = p - p^2 = pq$. Тогда по теореме 1.4:

$$\mathbb{D}(\xi) = \sum_{i=1}^{n} \mathbb{D}(\xi_i) = pq \cdot n = npq$$

■

1.1.3 Геометрическое распределение

Определение 9. Случайная величина ξ распределена геометрически, если она моделирует схему испытаний до первого успеха с вероятностью p.

$$Geom(p) \sim \begin{pmatrix} 1 & 2 & \cdots & n & \cdots \\ p & qp & \cdots & q^{n-1}p & \cdots \end{pmatrix}$$

Лемма 1.6. Математическое ожидание геометрически распределенной случайной величины ξ может быть вычислено, как $\mathbb{E}(\xi) = \dfrac{1}{p}$.

Доказательство.

$$\begin{aligned}
&\mathbb{E}(\xi) = p + 2qp + 2q^2p + \ldots + kq^{k-1}p + \ldots = \\
=&(p + qp + q^2p + \ldots + q^{k-1}p + \ldots) + (qp + 2q^2p + \ldots + (k-1)q^{k-1}p + \ldots) = \\
=&\frac{p}{1-q} + q(p + 2pq + \ldots + (k-1)q^{k-2}p + \ldots)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}(\xi) = 1 + q\mathbb{E}(\xi) \\
&\mathbb{E}(\xi)(1-q) = 1 \\
&\mathbb{E}(\xi) = \frac{1}{p}
\end{aligned}$$

■

Лемма 1.7. Дисперсия геометрически распределенной случайной величины ξ может быть вычислена, как $\mathbb{D}(\xi) = \dfrac{q}{p^2}$.

Доказательство.

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

$$\begin{aligned}
&\mathbb{E}(\xi^2) = p + 2qp + 9q^2p + \ldots + k^2q^{k-1}p + \ldots = \\
=&p + qp + 3qp + 4q^2p + 5q^2p + \ldots = \\
=&(qp + 4q^2p + \ldots) + (p + 3qp + 5q^2p + \ldots)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}(\xi^2) = q\mathbb{E}(\xi^2) + \mathbb{E}(2\xi - 1) = \\
=&(1-p)\mathbb{E}(\xi^2) + \frac{2}{p} - 1 = \frac{2-p}{p^2}
\end{aligned}$$

$$\mathbb{D}(\xi) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

■

1.1.4 Гипергеометрическое распределение

Определение 10. Случайная величина ξ распределена гипергеометрически, если она моделирует выбор n элементов из множества мощности N с K помеченными и является числом помеченных в выборке.

$$\xi \sim HG(N,\, K,\, n)$$

$$\mathbb{P}(\xi = k) = \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n}$$

Утверждение 1.8. Математическое ожидание гипергеометрически распределенной случайной величины ξ может быть вычислено, как $\mathbb{E}(\xi) = \dfrac{n \cdot K}{N}$.

Доказательство.

$$\xi = \mathbb{I}(A_1) + \mathbb{I}(A_2) + \ldots + \mathbb{I}(A_n), \text{ где } A_i = \{i\text{-ый элемент выборки помечен}\}$$
$$\mathbb{I}(A_i) \sim \begin{pmatrix} 0 & 1 \\ 1 - \frac{K}{N} & \frac{K}{N} \end{pmatrix}$$
$$\mathbb{E}(\xi) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{I}(A_i)) = n \cdot \frac{K}{N}$$

■

1.1.5 Распределение Паскаля

Определение 11. Случайная величина ξ распределена по Паскалю, если она моделирует испытания до первых k успехов.

Определение 12.

$$\xi \sim NB(p,\, k), \text{ если } \xi = \sum_{i=1}^{k} \eta_i : \forall i \in \{1,\, 2,\, \ldots,\, k\} : \eta_i \sim Geom(p)$$

$$\mathbb{P}(\xi = n) = C_{n-1}^{k-1} p^k q^{n-k}$$

Утверждение 1.9. Математическое ожидание случайной величины ξ, распределенной по Паскалю, может быть вычислено, как $\mathbb{E}(\xi) = \dfrac{k}{p}$.

Доказательство. Поскольку математическое ожидание линейно:

$$\mathbb{E}(\xi) = \sum_{i=1}^{k} \mathbb{E}(\xi_i) = \frac{1}{p} \cdot k = \frac{k}{p}$$

■

2 Ковариация

Определение 13. Пусть ξ и η – случайные величины, тогда ковариацией называется:

$$cov(\xi;\, \eta) = \mathbb{E}((\xi - \mathbb{E}(\xi))(\eta - \mathbb{E}(\eta)))$$

Теорема 2.1. Для $cov(\xi;\, \eta)$ выполняются свойства:

1. $cov(\xi;\, \xi) \geq 0$
2. $cov(\xi;\, \eta) = cov(\eta;\, \xi)$
3. $cov(\lambda\xi;\, \eta) = \lambda \cdot cov(\xi;\, \eta)$
4. $cov(\xi_1 + \xi_2;\, \eta) = cov(\xi_1;\, \eta) + cov(\xi_2;\, \eta)$
5. $cov(\xi;\, \eta) \leq \mathbb{D}(\xi) \cdot \mathbb{D}(\eta)$

Теорема 2.2.

$$cov(\xi;\, \eta) = \mathbb{E}(\xi \cdot \eta) - \mathbb{E}(\xi) \cdot \mathbb{E}(\eta)$$

Доказательство.

$$
\begin{aligned}
&\mathbb{E}((\xi - \mathbb{E}(\xi))(\eta - \mathbb{E}(\eta))) = \\
=&\mathbb{E}(\xi \cdot \eta - \xi\mathbb{E}(\eta) - \eta\mathbb{E}(\xi) + \mathbb{E}(\xi) \cdot \mathbb{E}(\eta)) = \\
=&\mathbb{E}(\xi \cdot \eta) - \mathbb{E}(\xi\mathbb{E}(\eta)) - \mathbb{E}(\eta\mathbb{E}(\xi)) + \mathbb{E}(\xi) \cdot \mathbb{E}(\eta) = \\
=&\mathbb{E}(\xi \cdot \eta) - \mathbb{E}(\xi) \cdot \mathbb{E}(\eta)
\end{aligned}
$$

∎

Теорема 2.3.

$$\mathbb{D}(\xi + \eta) = \mathbb{D}(\xi) + \mathbb{D}(\eta) + 2 \cdot cov(\xi;\, \eta)$$

3 Корреляция

Определение 14. Пусть ξ и η – случайные величины: $\mathbb{D}(\xi) \neq 0$, $\mathbb{D}(\eta) \neq 0$, $cov(\xi;\, \eta)$ определена корректно. Тогда коэффициентом корреляции ξ и η называется:

$$corr(\xi;\, \eta) = r_{\xi\eta} = \frac{cov(\xi;\, \eta)}{\sigma(\xi) \cdot \sigma(\eta)}$$

Свойства:

1. $|r_{\xi\eta}| \leq 1$
2. $|r_{\xi\eta}| = 1 \Longleftrightarrow \exists\, k \neq 0,\, b :\ \eta = k\xi + b$ (почти наверное).

4 Мера Жордана

Определение 15. A измеримо по Жордану, если $\mu^j(A) = \mu_j(A)$, где $\mu^j(A) = \inf\{\mu(\delta) :\ A \subset \delta\}$, $\mu_j(A) = \sup\{\mu(\delta) :\ \delta \subset A\}$.

Определение 16. Пусть $A \subset \Omega$, тогда $\mathbb{P}(x \in A) = \dfrac{\mu(A)}{\mu(\Omega)}$.

5 Распределение Пуассона

Теорема 5.1 (Теорема Пуассона)**.** Пусть $n \to \infty$, $p \to 0$, $np \to \lambda$, $\lambda = \text{const}$, тогда если ξ – количество успехов в серии испытаний Бернулли, то она распределена по Пуассону:

$$\xi \sim P(\lambda) :\ \mathbb{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Доказательство.

$$
\begin{aligned}
&P(\xi = k) = C_n^k p^k q^{n-k} \to \\
\to&\frac{n!}{(n-k)! \cdot k!} \cdot p^k q^{n-k} \to \\
\to&\frac{p^k}{k! \cdot q^k} \cdot \frac{n!}{(n-k)!} \cdot q^n \to \\
\to&\frac{p^k q^n}{k! \cdot q^k} \cdot n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1) \to \\
\to&\frac{p^k q^n n^k}{k! \cdot q^k} \cdot 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \ldots \cdot \left(1 - \frac{k-1}{n}\right) \to
\end{aligned}
$$

$$
\begin{aligned}
\to &\frac{\lambda^k \cdot q^n}{k! \cdot q^k} \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \ldots \cdot \left(1 - \frac{k-1}{n}\right) \to \\
\to &\frac{\lambda^k}{k!} q^n
\end{aligned}
$$

$$
\ln q^n = n \cdot \ln(1-p) \to -np \to -\lambda \Longrightarrow \frac{\lambda^k}{k!} q^n = \frac{\lambda^k}{k!} e^{-\lambda}
$$

■

Теорема 5.2.

$$
\lim_{n\to\infty} \sum_{i=0}^{n} \frac{x^i}{i!} = e^x, \; x \in \mathbb{R}
$$

Теорема 5.3. Пусть $\xi \sim P(\lambda)$. Тогда $\mathbb{E}(\xi) = \mathbb{D}(\xi) = \lambda$.

Доказательство.

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

■

Лемма 5.4. Пусть $\xi \sim P(\lambda_\xi)$, $\eta \sim P(\lambda_\eta)$, ξ и η независимы. Тогда $(\xi + \eta) \sim P(\lambda_\xi + \lambda_\eta)$.

Доказательство.

$$
\begin{aligned}
&\mathbb{P}(\xi + \eta = n) = \\
=&\sum_{i=0}^{n} \mathbb{P}(\xi = i) \cdot \mathbb{P}(\eta = n - i) = \\
=&\sum_{i=0}^{n} \frac{\lambda_\xi^i}{i!} \cdot e^{-\lambda_\xi} \cdot \frac{\lambda_\eta^{n-i}}{(n-i)!} \cdot e^{-\lambda_\eta} = \\
=&e^{-(\lambda_\xi + \lambda_\eta)} \sum_{i=0}^{n} \frac{\lambda_\xi^i \cdot \lambda_\eta^{n-i}}{i! \cdot (n-i)!} \cdot \frac{n!}{n!} = \\
=&\frac{e^{-(\lambda_\xi + \lambda_\eta)}}{n!} \sum_{i=0}^{n} C_n^i \lambda_\xi^i \lambda_\eta^{n-i} = \\
=&e^{-(\lambda_\xi + \lambda_\eta)} \cdot \frac{(\lambda_\xi + \lambda_\eta)^n}{n!}
\end{aligned}
$$

■

6 Ветвящиеся процессы

6.1 Цепи Маркова

Определение 17. Последовательность случайных величин $\xi_0, \xi_1, \ldots, \xi_n$ называется Цепью Маркова, если

$$
\forall n, i_0, i_1, \ldots, i_n : \; \mathbb{P}(\xi_{n-1} = x_{i_{n-1}}, \ldots, \xi_0 = x_{i_0})
$$

верно, что:

$$
\mathbb{P}(\xi_n = x_{i_n} \,|\, \xi_{n-1} = x_{i_{n-1}}, \ldots, \xi_0 = x_{i_0}) = \mathbb{P}(\xi_n = x_{i_n} \,|\, \xi_{n-1} = x_{i_{n-1}})
$$

Определение 18. Цепь Маркова называется однородной, если:

$$\forall i,\, j:\ \mathbb{P}(\xi_n = x_j \,|\, \xi_{n-1} = x_i) = p_{i,j} \text{ не зависит от } n.$$

Определение 19. Матрица $A = (a_{i,j})$ называется стохастической, если:

$$\forall i,\, j:\ a_{i,j} \in [0;\, 1],\ \sum_i (a_{i,j}) = 1$$

Определение 20. Матрица $\pi = (p_{i,j})$ называется матрицей переходных вероятностей.

Теорема 6.1. Пусть $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, \ldots, p_n^{(0)})$ и $p^{(k)} = (p_1^{(k)}, p_2^{(k)}, \ldots, p_n^{(k)})$ – начальное распределение и распределение на k-ом шаге соответственно вероятностей Марковской цепи, где $p_i^{(k)} = \mathbb{P}(\xi_k = x_i)$. Тогда:

$$p^{(k)} = p^{(0)} \cdot \pi^k$$

6.1.1 Классификация состояний Марковских цепей

Определение 21. Состояние x_j достижимо из x_i, если:

$$\exists\, k:\ P_{ij}^k = \mathbb{P}(\xi_{m+k} = x_j \,|\, \xi_m = x_i) > 0$$

Определение 22. Состояния называются сообщающимися, если они достижимы друг для друга.

Определение 23. Состояние x_i называется несущественным, если существует такое состояние x_j, что x_j достижимо из x_i, но x_i недостижимо из x_j.

Определение 24. Состояние x_i называется существенным, если существует такое состояние x_j, что x_j достижимо из x_i и x_i достижимо из x_j.

Определение 25. Марковская цепь, все состояния которой составляют один класс сообщающихся состояний, называется неразложимой.

Определение 26. Состояние x_i называется возвратным, если вероятность возвращения в это состояние равна 1.

Определение 27. Состояние x_i называется невозвратным, если вероятность возвращения в это состояние не равна 1.

Определение 28. Возвратное состояние x_i называется возвратным положительным, если среднее время возвращения в него конечно.

Определение 29. Возвратное состояние x_i называется возвратным нулевым, если среднее время возвращения в него бесконечно.

Определение 30. Состояние x_i называется периодическим, если НОД$\{k:\ P_{ii}^{(k)} > 0\} = d > 1$, где d – период состояния.

6.1.2 Эргодичность

Определение 31. Марковская цепь называется эргодической, если:

$$\forall i,\, j:\ \exists \lim_{k\to\infty} P_{ij}^{(k)} = p_{ij} > 0,\ \sum_j p_j = 1$$

Теорема 6.2 (Критерий эргодичности)**.** Марковская цепь эргодична, если:

$$\exists\, k:\ \forall i,\, j:\ P_{ij}^{(k)} > 0$$

6.2 Процесс Гальтона-Ватсона

Определение 32. Пусть $p_0, p_1, \ldots, p_m : p_m \geq 0;\ p_0 + p_1 + \ldots + p_m = 1$ – начальное распределение. Пусть для $i \geq 2$ определено:

$$p_i^{*k} = \sum_{i_1+i_2+\ldots+i_k=i} p_{i_1} \cdot p_{i_2} \cdot \ldots \cdot p_{i_k}$$

Процесс Гальтона-Ватсона есть марковская цепь $Z(n)$, $n \in \mathbb{N}_0$ с начальным распределением $P_0(k) = \mathbb{P}(Z(0) = k)$ и переходными вероятностями:

$$P_{ij} = \mathbb{P}(Z(n+1) = j \mid Z(n) = i) = \begin{cases} p_j^{*i}, \text{ если } i \geq 1,\ j \geq 0 \\ \delta_{0j}, \text{ если } i \geq 0,\ j \geq 0 \end{cases}$$

Если не оговорено иного, $P_0(1) = \mathbb{P}(Z(0) = 1) = 1;\ P_0(k) = 0,\ k \neq 1.$

Пример (Деление клетки 1)**.** Рассмотрим популяцию частиц. Пусть после каждой единицы времени частица либо умирает, либо делится на двое. При этом пусть в начале мы имели только одну частицу, то есть $Z(0) = 1$. Рассмотрим случайную величину $\xi_i^{(n)}$ – число потомков i-й частицы n-го поколения:

$$Z(n+1) = \xi_1^{(n)} + \xi_2^{(n)} + \ldots + \xi_{Z(n)}^{(n)},\ n \geq 0$$

Заметим, что можно также фиксировать не число частиц в момент времени n, а число частиц первого поколения:

$$Z(n+1) = Z_1(n) + Z_2(n) + \ldots + Z_{Z(1)}(n)$$

Таким образом для независимых $\xi_i^{(n)}$ верно:

$$\xi_1^{(n)} + \xi_2^{(n)} + \ldots + \xi_{Z(n)}^{(n)} \stackrel{\text{def}}{=} Z(n+1) = Z_1(n) + Z_2(n) + \ldots + Z_{Z(1)}(n)$$

Далее смотреть в примере 7.2.

6.3 Тождество Вальда

Теорема 6.3. Пусть $\xi_0, \xi_1, \ldots$ – независимые одинаково распределенные случайные величины. Пусть τ – случайный момент времени, не зависящий от (ξ_i). Пусть $S_n = \xi_0 + \xi_1 + \ldots + \xi_n$. Тогда:

$$\mathbb{E}(S_\tau) = \mathbb{E}(\xi) \cdot \mathbb{E}(\tau)$$

Доказательство.

$$\sum_{k=1}^{\infty} \mathbb{P}(\tau \geq k) = \sum_{k=1}^{\infty} \sum_{n=1}^{k} \mathbb{P}(\tau = n) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{P}(\tau = n) = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(\tau = n) = \mathbb{E}(\tau)$$

$$\mathbb{E}(S_\tau) = \sum_{n=1}^{\infty} \mathbb{E}(S_\tau;\ \tau = n) = \sum_{n=1}^{\infty} \mathbb{E}(\xi_1 + \xi_2 \ldots + \xi_n;\ \tau = n) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{E}(\xi_k;\ \tau = n) =$$

$$= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \mathbb{E}(\xi_k;\ \tau = n) = \sum_{k=1}^{\infty} \mathbb{E}(\xi_k;\ \tau \geq k) = \sum_{k=i}^{\infty} \mathbb{E}(\xi_k) \cdot \mathbb{E}(\tau \geq k) = \mathbb{E}(\xi) \cdot \mathbb{E}(\tau)$$

■

7 Производящие функции

Определение 33. Производящей функцией произвольной последовательности (a_n) называется выражение вида:

$$a_0 + a_1 z + a_2 z^2 + \ldots = \sum_{i=0}^{\infty} a_i z^i$$

7.1 Операции с производящими функциями

Определение 34. Суммой производящих функций $A(z) = a_0 + a_1 z + a_2 z^2 + \ldots$ и $B(z) = b_0 + b_1 z + b_2 z^2 + \ldots$ называется производящая функция:

$$A(z) + B(z) = (a_0 + b_0) + (a_1 + b_1)z + (a_2 + b_2)z^2 + \ldots$$

Определение 35. Произведением производящих функций $A(z) = a_0 + a_1 z + a_2 z^2 + \ldots$ и $B(z) = b_0 + b_1 z + b_2 z^2 + \ldots$ называется производящая функция:

$$A(z) \cdot B(z) = a_0 b_0 + (a_0 b_1 + a_1 b_0)z + (a_0 b_2 + a_1 b_1 + a_2 b_0)z^2 + \ldots$$

Определение 36. Пусть $A(z) = a_0 + a_1 z + a_2 z^2 + \ldots$; $B(t) = b_0 + b_1 t + b_2 t^2 + \ldots$; $b_0 = 0$ – производящие функции. Подстановкой производящей функции B в производящую функцию A будет называться производящая функция:

$$A(B(t)) = a_0 + a_1 b_1 t + (a_1 b_2 + a_2 b_1^2)t^2 + (a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3)t^3 + \ldots$$

Теорема 7.1. Пусть $B(t) = b_0 + b_1 t + b_2 t^2 + \ldots$; $b_0 = 0$; $b_1 \neq 0$ – производящая функция. Тогда существуют единственные такие функции $A(z) = a_0 + a_1 z + a_2 z^2 + \ldots$; $a_0 = 0$ и $C(u) = c_0 + c_1 u + c_2 u^2 = \ldots$; $c_0 = 0$, что $A(B(t)) = t$ и $B(C(u)) = u$. Функция A называется левой обратной, а функция C – правой обратной к функции B.

Доказательство. Рассмотрим левую обратную функцию:

$$A(B(t)) = a_1 b_1 t + (a_1 b_2 + a_2 b_1^2)t^2 + (a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3)t^3 + \ldots = t$$

Чтобы равенство выполнялось, коэффициент при t должен равняться 1, а коэффициенты при t^n, $n \geq 2$ должны равняться 0. Отсюда $a_1 b_1 = 1 \implies a_1 = \dfrac{1}{b_1}$. Пусть аналогично определены коэффициенты $a_1, a_2, \ldots, a_n$. Тогда коэффициент a_{n+1} будет определяться из условия, что многочлен $a_{n+1} b_1^{n+1} + \ldots$ от $a_1, a_2, \ldots, a_n, a_{n+1}$ и $b_1, b_2, \ldots, b_n, b_{n+1}$, являющийся коэффициентом при t^{n+1}, будет равен нулю. Поскольку $b_1 \neq 0$ по условию, получаем уравнение от a_{n+1} с единственным корнем. То есть мы однозначно можем задать такие коэффициенты $a_1, a_2, \ldots$, чтобы $A(B(t)) = t$.

Доказательство для правой обратной функции аналогично. ■

Определение 37. Производящая функция называется рациональной, если ее можно представить в виде $\dfrac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ – многочлены.

7.2 Производящие функции вероятности

Определение 38. Производящей функцией φ_ξ случайной величины ξ называется производящая функция последовательности $(\mathbb{P}(\xi = n))_{n=0}^{\infty}$:

$$\varphi_\xi(z) = \sum_{n=0}^{\infty} z^n \mathbb{P}(\xi = n) = \mathbb{E}(z^\xi)$$

Причем $\varphi_\xi(1) = 1$ как сумма вероятностей. Рассмотрим производную данной функции:

$$\varphi'_\xi(z) = \sum_{n=0}^{\infty} n \cdot z^{n-1}\mathbb{P}(\xi = n) \qquad\qquad \varphi'_\xi(1) = \sum_{n=0}^{\infty} n \cdot \mathbb{P}(\xi = n) = \mathbb{E}(\xi)$$

Определение 39. Итерацией производящей функции случайной величины порядка n называется композиция, строящаяся рекуррентно:

$$\varphi_0(z) = z \qquad\qquad \varphi_1(z) = \varphi(z) \qquad\qquad \varphi_{n+1}(z) = \varphi_n(\varphi(z))$$

Определение 40. Пусть $\xi_i,\ i \in \{1,\ 2, \ldots,\ n\}$ – независимые случайные величины, тогда:

$$\varphi_{\xi_1+\xi_2+\ldots+\xi_n}(z) = \varphi_{\xi_1} \cdot \varphi_{\xi_2} \cdot \ldots \cdot \varphi_{\xi_n} = \varphi_n(\varphi_{\xi_1}(z))$$

При этом ξ_1 может быть заменена на любую из ξ_i, так как они одинаково распределены.

Пример (Деление клетки 2)**.** Продолжим рассмотрение примера 6.2.

$$F(n; z) \overset{\text{def}}{=} \mathbb{E}(z^{Z_{(n)}}) = \sum_{k=0}^{\infty} z^k \mathbb{P}(Z(n) = k)$$

Рассмотрим $F(n+1; z)$ при учёте, что $Z(0) = 1$ и используя свойства итераций:

$$\begin{aligned} F(n+1; z) &= \mathbb{E}\left(z^{Z_{(n+1)}}\right) = \mathbb{E}\left(\mathbb{E}\left(z^{Z_{(n+1)}} | Z(n)\right)\right) = \\ &= \mathbb{E}\left(\mathbb{E}\left(z^{\xi_1^{(n)}+\xi_2^{(n)}+\ldots+\xi_{Z(n)}^{(n)}} | Z(n)\right)\right) = \mathbb{E}\left(\prod_{k=1}^{Z_{(n)}} \mathbb{E}\left(z^{\xi_k^{(n)}}\right)\right) = \mathbb{E}\left(\mathbb{E}\left(z^{\xi}\right)^{Z(n)}\right) = \\ &= F(n; \varphi(z)) = F(n-1; \varphi_2(z)) = F(0; \varphi_{n+1}(z)) = \varphi_{n+1}(z) \end{aligned}$$

Тогда вероятность вырождения к моменту n может быть вычислена как итерация:

$$\varphi_n(0) = F(n; 0) = \sum_{k=0}^{\infty} 0^k \mathbb{P}(Z(n) = k) = \mathbb{P}(Z(n) = 0)$$

7.3 Классификация процессов Гальтона-Ватсона

Определение 41. Пусть $A \overset{\text{def}}{=} \mathbb{E}(\xi) = \mathbb{E}(Z(1)|Z(0) = 1) < \infty$. Тогда процесс Гальтона-Ватсона называется:

Докритическим, если $A < 1$ Критическим, если $A = 1$ Надкритическим, если $A > 1$

Лемма 7.2. $\mathbb{E}(Z(n)) = A^n$. При этом если $\mathbb{D}(\xi) < \infty$, то:

$$\mathbb{D}(Z(n)) = \begin{cases} \mathbb{D}\left(\dfrac{A^{n-1}(A^n - 1)}{A - 1}\right), & \text{если } A \neq 1 \\ \sigma(n^2), & \text{если } A = 1 \end{cases}$$

Доказательство.

$$\begin{aligned} \mathbb{E}(Z(n)) &= \mathbb{E}\left(\xi_1^{(n-1)} + \xi_2^{(n-1)} + \ldots + \xi_{Z(n-1)}^{(n-1)}\right) = \mathbb{E}(\xi) \cdot \mathbb{E}(Z(n-1)) = \\ &= A \cdot \mathbb{E}(Z(n-1)) = A^2 \cdot \mathbb{E}(Z(n-2)) = \ldots = A^n \cdot \mathbb{E}(Z(0)) = A^n \end{aligned}$$

■

Теорема 7.3. Вероятность вырождения ветвящегося процесса равна наименьшему неотрицательному корню p уравнения $z = \varphi(z)$.

Доказательство.
$$\varphi_n(z) = \mathbb{P}(Z(n) = 0) \le \mathbb{P}(Z(n+1) = 0) = \varphi_{n+1}(z)$$
В силу монотонности и возрастания последовательности
$$P(n) = \mathbb{P}(Z(n) = 0) = \varphi_n(0)$$
выполняется теорема Вейерштрасса, по которой она имеет предел r при $n \to \infty$. Заметим, что
$$P(1) = \varphi_1(0) = \varphi_1(\varphi_0(0)) \le p = \varphi(p)$$
Пусть $\varphi_{n-1}(0) = P(n-1) < p$, тогда по индукции получаем:
$$P(n) = \varphi_n(0) = \varphi(\varphi_{n-1}(0)) = \varphi(P(n-1)) \le \varphi(p) = p$$
Отсюда в силу непрерывности φ имеем:
$$r = \lim_{n\to\infty} P(n) = \lim_{n\to\infty} \varphi(P(n-1)) = \varphi\left(\lim_{n\to\infty} P(n-1)\right) = \varphi(r)$$
Таким образом, мы также получаем, что докритические процессы вырождаются с вероятностью $p = 1$, а надкритические процессы вырождаются с вероятностью $p < 1$, которая является минимальным неотрицательным корнем уравнения $z = \varphi(z)$. ■

Пример. Рассмотрим бинарное деление клетки с производящей функцией $\varphi(z) = 1 - p + pz^2$. Решим уравнение $z = 1 - p + pz^2$, получив корни $z = 1$ и $z = \dfrac{1-p}{p}$. То есть при $p > 0{,}5$ вероятность вырождения равна $\mathbb{P} = \dfrac{1-p}{p}$, а при $p \le 0{,}5$ имеем $\mathbb{P} = 1$.

8 Математическая статистика

Определение 42. Функция распределения случайной величины ξ суть $F_\xi(x) = \mathbb{P}(\xi \le x)$.

Определение 43. Квантилем уровня α называется такой x_α, что он является решением уравнения $F_\xi(x) = \alpha$.

Определение 44. Выборкой объема $n;\ n > 1$ называется случайный вектор $Z_n = (\xi_1, \dots, \xi_n)$, где $\xi_i,\ i \in \{1, \dots, n\}$ называются элементами выборки и являются независимыми случайными величинами с одной функцией распределения $F_\xi(x)$. Тогда выборка Z_n соответствует функции распределения $F_\xi(x)$.

Определение 45. Реализацией выборки называется неслучайный вектор $z_n = (x_1, \dots, x_n)$, компонентами которого являются реализации соответствующих элементов выборки $\xi_i,\ i \in \{1, \dots, n\}$.

Следствие 8.0.1. Из определений 44 и 45 следует, что реализацию выборки z_n можно также рассматривать как последовательность $x_1, \dots, x_n$ из n реализаций одной и той же случайной величины ξ, полученных в серии из n независимых одинаковых опытов, проводимых в одинаковых условиях. Поэтому можно говорить, что выборка Z_n порождена наблюдаемой случайной величиной ξ, имеющей распределение $F_\xi(x) = F(x)$.

Определение 46. Параметром распределения θ случайной величины ξ называется любая числовая характеристика этой случайной величины или любая константа, явно входящая в выражение для функции распределения.

Определение 47. Случайная величина $T = \varphi(Z_n)$, где $\varphi(Z_n)$ – произвольная функция определенная на выборочном пространстве, называется статистикой.

Определение 48. Статистической гипотезой H называется любое предположение относительно вида распределения, параметров распределения или свойств закона распределения наблюдаемой в эксперименте случайной величины X.

Определение 49. Любое предположение относительно параметров распределения случайной величины X называют параметрической гипотезой.

Определение 50. Проверяемая гипотеза называется основной (или нулевой) и обозначается H_0. Гипотеза, конкурирующая с H_0, называется альтернативной и обозначается H_1.

Определение 51. Статистическим критерием проверки гипотезы H_0 называется правило, в соответствии с которым по реализации $t = \varphi(z_n)$ статистики T гипотеза H_0 принимается или отвергается.

Определение 52. Критической областью $\overline{G}$ статистического критерия называется область реализаций t статистики T, при которых гипотеза H_0 отвергается.

Определение 53. Доверительной областью G статистического критерия называется область реализаций t статистики T, при которых гипотеза H_0 принимается.

Определение 54. Ошибкой 1-го рода называется событие, состоящее в том, что гипотеза H_0 отвергается, когда она верна.

Определение 55. Ошибкой 2-го рода называется событие, состоящее в том, что принимается гипотеза H_0, когда верна гипотеза H_1.

Определение 56. Уровнем значимости статистического критерия называется вероятность ошибки 1-го рода:

$$\alpha = \mathbb{P}(T \in \overline{G} | H_0 - \text{верна})$$

8.1 Нормальное распределение

Определение 57. Случайная величина ξ распределена нормально с параметрами μ и σ, если её функция плотности вероятности имеет вид:

$$F_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Теорема 8.1. Пусть $\xi \sim \mathcal{N}(\mu;\ \sigma)$. Тогда $\mathbb{E}(\xi) = \mu$ и $\mathbb{D}(\xi) = \sigma^2$.

Определение 58. Случайна величина ξ распределена стандартно нормально, если $\xi \sim \mathcal{N}(0;\ 1)$.

Определение 59. Функция плотности стандартно нормально распределенной случайной величины ξ обозначается:

$$\varphi_\xi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

Теорема 8.2. Функцией распределения стандартно нормально распределенной случайной величины ξ является функция Лапласа:

$$\Phi_\xi(x) = \int_{-\infty}^{x} \varphi_\xi(y)\, dy$$

Теорема 8.3 (Правило трёх сигм)**.** Для случайной величины ξ, имеющей стандартное нормальное распределение:

$$\mathbb{P}(|\xi - \mu| < 3\sigma) \approx 0{,}9973$$

Теорема 8.4 (Центральная предельная теорема)**.** Пусть $\xi_1, \ldots, \xi_n$ – последовательность независимых случайных величин, а S_n – их сумма, тогда:

$$\frac{S_n - n \cdot \mathbb{E}(\xi)}{\sigma(\xi) \cdot \sqrt{n}} \xrightarrow{n \to \infty} \mathcal{N}(0;\, 1)$$

Пусть $\hat{\xi}$ – центрированная и нормированная случайная величина, тогда:

$$\mathbb{P}(a < \hat{\xi} < b) = \mathbb{P}\left(a < \frac{S_n - n \cdot \mathbb{E}(\xi)}{\sigma(\xi) \cdot \sqrt{n}} < b\right) \approx \Phi(b) - \Phi(a)$$

Теорема 8.5 (Муавр-Лаплас)**.** Пусть $\xi \sim B(p;\, n)$, $n \geq 30$, тогда можно сказать, что $\xi \sim \mathcal{N}(np;\, npq)$, а значит:

$$\mathbb{P}(a < \xi < b) = \mathbb{P}\left(\frac{a - np}{npq} < \frac{\xi - np}{npq} < \frac{b - np}{npq}\right) = \Phi\left(\frac{b - np}{npq}\right) - \Phi\left(\frac{a - np}{npq}\right)$$

8.2 Оценка выборки

Определение 60. Точечной (выборочной) оценкой неизвестного параметра распределения θ называется произвольная статистика $\hat{\theta}(Z_n)$, построенная по выборке Z_n.

Определение 61. Оценка $\hat{\theta}$ называется несмещённой оценкой параметра θ, если математическое ожидание оценки равно оцениваемому параметру, то есть $\mathbb{E}(\hat{\theta}) = \theta$.

Определение 62. Оценка $\hat{\theta}$ называется состоятельной оценкой параметра θ, если она сходится по вероятности к оцениваемому параметру, то есть:

$$\mathbb{P}(|\hat{\theta} - \theta| < \varepsilon) = 1$$

Теорема 8.6. Размер выборки, необходимой для оценки, с точностью δ и уровнем доверия $1-\alpha$ можно выразить как:

$$n = \left(\frac{z_{1-\frac{\alpha}{2}}}{\delta}\right)^2 pq$$

Доказательство. Пусть $\hat{p}$ – вероятность наличия у объекта выборки какого-то признака, тогда должно выполняться условие:

$$\mathbb{P}(|\hat{p} - p| < \delta) = 1 - \alpha$$

Поскольку частота является несмещённой и состоятельной оценкой вероятности, при проведении n опытов, в X из которых произойдёт нужное нам событие, вероятность можно будет представить, как $\hat{p} = \frac{X}{n}$. При этом в случае выбора с возвращением, X будет иметь биномиальное распределение с параметрами n и p, где n – количество проведенных экспериментов, а p – вероятность появления признака. Тогда по центральной предельной теореме:

$$\frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{D}(X)}} = \frac{X - np}{\sqrt{npq}} \xrightarrow{n \to \infty} \mathcal{N}(0;\, 1)$$

В таком случае можно считать, что $\hat{p} \sim \mathcal{N}(0;\, 1)$, а значит:

$$\mathbb{E}(\hat{p}) = \mathbb{E}\left(\frac{X}{n}\right) = \frac{np}{n} = p \qquad\qquad \mathbb{D}(\hat{p}) = \mathbb{D}\left(\frac{X}{n}\right) = \frac{npq}{n^2} = \frac{pq}{n}$$

Зная распределение $\hat{p}$, мы можем установить, когда выполняется равенство:

$$\mathbb{P}(|\hat{p}-p| < \delta) = \mathbb{P}(p-\delta < \hat{p} < p+\delta) = \Phi\left(\frac{p+\delta-p}{\sqrt{\frac{pq}{n}}}\right) - \Phi\left(\frac{p-\delta-p}{\sqrt{\frac{pq}{n}}}\right) =$$

$$= \Phi\left(\frac{\delta}{\sqrt{\frac{pq}{n}}}\right) - \left(1 - \Phi\left(\frac{\delta}{\sqrt{\frac{pq}{n}}}\right)\right) = 2\Phi\left(\frac{\delta}{\sqrt{\frac{pq}{n}}}\right) - 1 = 1-\alpha \Longrightarrow \Phi\left(\frac{\delta}{\sqrt{\frac{pq}{n}}}\right) = 1 - \frac{\alpha}{2}$$

Получили, что аргумент функции Лапласа является квантилем уровня $1 - \dfrac{\alpha}{2}$ стандартного нормального распределения, тогда остается выразить n:

$$n = \left(\frac{z_{1-\frac{\alpha}{2}}}{\delta}\right)^2 pq$$

Здесь p определяется либо из предварительного опроса, либо подбирается такое её значение, при котором $\mathbb{D}(\hat{p})$ максимальное, то есть $p = 0{,}5$. ■