Забавные факты по теории вероятностей

Источник: Чернова Н.И., Теория вероятностей

Авторы заметок: Хоружий Кирилл

От: *8 июля 2021 г.*

Содержание

Контрольная работа №1 по теории вероятностей

Автор работы: Хоружий Кирилл

От: *8 июля 2021 г.*

1 Контрольная работа №1

T1

Известно, что ξ нормально распределена, и $\mathrm{E}\,\xi = -11$, а $\mathrm{E}\left((2\xi+3)(\xi-2)\right) = 253$, найти $\mathrm{P}(\xi \in (-15, 0.16))$.
Итак, $\mathrm{P}(\xi)$ вида

$$\mathrm{P}(\xi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-a)^2}{2\sigma^2}\right).$$

Матожидание можем найти, как интеграл вида

$$\mathrm{E}\,\xi = \int_{-\infty}^{+\infty} \frac{\xi}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-a)^2}{2\sigma^2}\right) d\xi = \Big/ \xi - a = t \Big/ = \frac{a}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{t^2}{2\sigma^2}\,dt\right) = a.$$

Аналогчино находим второй момент:

$$\mathrm{E}\,\xi^2 = \int_{-\infty}^{+\infty} \frac{\xi^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-a)^2}{2\sigma^2}\right) d\xi = \Big/ \xi - a = t \Big/ = \int_{-\infty}^{+\infty} \frac{t^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt + \int_{-\infty}^{+\infty} \frac{a^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt.$$

Первый интеграл легко сводится к Гамма-функции, находим, что

$$\int_{-\infty}^{+\infty} \frac{t^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt = \Gamma\left(\frac{3}{2}\right) \cdot \frac{(2\sigma^2)^{3/2}}{\sqrt{2\pi}} = \sigma^2,$$

второй интеграл просто сводится к a^2, итого

$$\mathrm{E}\,\xi^2 = \sigma^2 + a^2.$$

Так приходим к системе вида

$$\begin{cases} 2\,\mathrm{E}(\xi^2) - \mathrm{E}(\xi) - 6 = 253 \\ \mathrm{E}(\xi) = a \end{cases} \quad \Rightarrow \quad \begin{cases} \sigma^2 = 3, \\ a = -11. \end{cases}$$

Искомая веростность $\mathrm{P}(\xi \in (-15, 0.16))$ в таком случае равна

$$\int \mathbf{N}(\xi,\, \sigma,\ a)\, d\xi = \frac{1}{2}\mathrm{Erf}\left(\frac{\xi - a}{\sqrt{2}\sigma}\right), \quad \Rightarrow \quad \mathrm{P}(\xi \in (-15, 0.16)) = \frac{1}{2}\left(\mathrm{Erf}\left(\frac{0.16-11}{\sqrt{6}}\right) - \mathrm{Erf}\left(\frac{-15+11}{\sqrt{6}}\right)\right) \approx 0.99$$

T2

Для функции вида

$$f(x) = \begin{cases} C/(x-3), & -5 < x < 2 \\ 0, & x \leqslant -5 \ || \ 2 \leqslant x \end{cases}$$

являющейся функцией распределения некоторой случайной величины ζ найдём C и характеристики $\mathrm{E}(\zeta)$, $\mathrm{D}(\zeta)$.
Для начала найдём C:

$$1 = \int_{-\infty}^{+\infty} f(x)\, dx = -\int_{-5}^{2} \frac{C}{3-x} = C\ln(3-x)\Big|_{-5}^{2} = -C\ln(8), \quad \Rightarrow \quad C = -\frac{1}{\ln 8}.$$

Теперь найдём характеристики ζ:

$$\mathrm{E}(\zeta) = -\int_{-5}^{2} x\frac{C}{3-x} = \Big/ -\frac{x}{3-x} = \frac{3}{x-3} + 1 \Big/ = \frac{7 - 3\ln(8)}{-\ln 8} \approx -0.37$$

$$\mathrm{E}(\zeta^2) = -\int_{-5}^{2} x^2\frac{C}{3-x} = \Big/ -\frac{x^2}{3-x} = x + \frac{9}{x-3} + 3 \Big/ = \frac{18\ln 2 - 7}{\ln 4},$$

тогда дисперсия ζ:

$$\mathrm{D}(\zeta) = \mathrm{E}(\zeta^2) - \mathrm{E}^2(\zeta) = \frac{7(27\ln(2) - 14)}{18\ln^2(2)} \approx 3.82.$$

T3

Введем для каждого места величину ξ_i, равную 1 в случае нечётного числа и 0 иначе. Число, наверное, подразумевает, что на первом месте не может стоять ноль, но на всякий случай пока обозначим вероятность быть первой цифре нечетной за γ, остальных местах равновероятны значения 0 и 1.

Вероятность существования хотя бы одной нечётной цифры найдём через вероятность их отсутсвия:

$$\mathrm{P}(\exists a_i \in \mathrm{Odd}) = 1 - \left(\frac{1}{2}\right)^2 \cdot (1-\gamma).$$

Матожидание же величины $\xi = \sum_{i=1}^{8} \xi_i$ легко найти, в силу независимости ξ:

$$\mathrm{E}(\xi) = \sum_{i=1}^{8} \mathrm{E}(\xi_i) = 7\,\mathrm{E}(\xi_i) + \gamma.$$

Если на первом месте может стоять 0, то $\gamma = 0.5$ и, соответственно,

$$\mathrm{P}(\exists a_i \in \mathrm{Odd}) = \frac{255}{256} \approx 1 - 3.9 \cdot 10^{-3}, \qquad \mathrm{E}(\xi) = 4.$$

Если же 0 стоять на первом месте не может, то $\gamma = 5/9$ и, соответственно,

$$\mathrm{P}(\exists a_i \in \mathrm{Odd}) = 1 - \frac{1}{128}\frac{5}{9} \approx 1 - 4.3 \cdot 10^{-3}, \qquad \mathrm{E}(\xi) = \frac{73}{18} \approx 4.06.$$

T4

Найдём производящую функция для биномиального распределения, вида

$$\mathrm{P}(\xi = k) = C_n^k p^k q^{n-k},$$

что соответствует количеству успехов в схеме Бернулли, где вероятность успеха p.

Коэффициенты в биноме Ньютона выглядят очень похоже на P, так что заметим, что производящая функция вида

$$P(s) = (q + ps)^n,$$

нам подходит. Найдём матожидание и дисперсию, как

$$\mathrm{E}(\xi) = P'_s(s=1), \qquad \mathrm{D}(\xi) = P''(1) + P'(1) - \mathrm{E}^2(\xi).$$

Производные $P(s)$:

$$P'(s) = np(q+ps)^{n-1}, \quad P'(1) = np, \qquad P''(s) = n(n-1)p^2(q+ps)^{n-1}, \quad P''(1) = n(n-1)p^2,$$

тогда искомые величины:

$$\mathrm{E}(\xi) = np, \qquad \mathrm{D}(\xi) = np(1-p) = npq.$$

T5

Известно, что $\mathrm{E}\,x = 6$, $\mathrm{E}\,y = 19$, $\mathrm{D}\,x = 7$, $\mathrm{D}\,y = 12$, тогда матожидание и дисперсия для $z = 3x - 2y$ равны

$$\mathrm{E}\,z = \mathrm{E}(3x) - \mathrm{E}(2y) = 3\,\mathrm{E}(x) - 2\,\mathrm{E}(y) = -20, \qquad \mathrm{D}\,z = \mathrm{D}(3x) + \mathrm{D}(2y) = 9\,\mathrm{D}(x) + 4\,\mathrm{D}(y) = 111.$$

T6

Для поиска коэффициента корреляции сначала найдём дисперсию и матожидание количества людей, выходящих на 8 этаже, (ξ) и количества людей, выходящих на 8 этаже или выше (η).

Представим величину ξ как сумму четырёх других $\xi = \sum_{i=1}^{4} \xi_i$, где ξ – вероятность выйти на 8 этаже для каждого из четырех людей:

ξ	1	0
P	1/12	11/12

В силу независимости ξ_i верно, что

$$\mathrm{E}(\xi) = \mathrm{E}\left(\sum_{i=1}^{4} \xi_i\right) = \sum_{i=1}^{4} \mathrm{E}(\xi_i) = \frac{4}{12}, \qquad \mathrm{D}(\xi) = \mathrm{D}\left(\sum_{i=1}^{4} \xi_i\right) = \sum_{i=1}^{4} \mathrm{D}(\xi_i) = 4\left(\frac{1}{12} - \left(\frac{1}{12}\right)^2\right) = \frac{11}{36}.$$

Аналогично найдём характеристики η, представив через сумму независимых величин $\eta = \sum_{i=1}^4 \eta_i$, где η_i – вероятность для каждого человека выйти на этаже, выше восьмого

η	1	0
P	5/12	7/12

Тогда, аналогично, в силу незаивимости η_i, находим

$$\mathrm{E}(\eta) = \sum_{i=1}^{4} \mathrm{E}(\eta_i) = 4 \cdot \frac{5}{12} = \frac{5}{3}, \qquad \mathrm{D}(\eta) = \sum_{i=1}^{4} \mathrm{D}(\eta_i) = 4 \cdot \left(\frac{5}{12} - \frac{25}{144} \right) = \frac{35}{36}.$$

Теперь найдём матожидание $\mathrm{E}(\xi\eta)$, построив таблицу $\mathrm{P}(\xi, \eta)$, где ξ и η принимает значения от 0 до 4. Заметим, что таблица будет верхнетреугольной: если на 8 этаже вышло n людей, то $\eta \geqslant n$. Сформировать вероятность $\mathrm{P}(\xi = \xi_0, \eta = \eta_0)$ можно, выбирая ξ_0 людей из 4 – оставшихся на 8 этаже, выбирая $\eta - \xi$ людей из $4 - \xi$ – оказавшихся на 9 этаже и выше, где вероятность оказаться ниже 8 – (7/12), и вероятность быть на 9 и выше – (5/12), итого находим

$$\mathrm{P}(\xi = \xi_0, \eta = \eta_0) = \binom{4}{\xi_0} \left(\frac{1}{12}\right)^{\xi_0} \binom{4-\xi_0}{\eta_0 - \xi_0} \left(\frac{7}{12}\right)^{4-\eta_0} \left(\frac{5}{12}\right)^{\eta_0 - \xi_0},$$

а искомое матожидание тогда будет равно

$$\mathrm{E}(\xi\eta) = \sum_{\xi_0=0}^{4} \sum_{\eta_0=0}^{4} \xi_0 \eta_0 \, \mathrm{P}(\xi =_{\xi_0}, \eta = \eta_0) = \frac{3}{4},$$

что нетрудно получить прямым вычислением. Конечно, судя по простоте ответа, его можно было получить и более простым путём, но зато мы уверены в результатах.

Наконец, корреляция ξ и η, равна

$$\rho(\xi, \eta) = \frac{\mathrm{cov}(\xi, \eta)}{\sqrt{\mathrm{D}\,\xi}\sqrt{\mathrm{D}\,\eta}} = \frac{\mathrm{E}(\xi\eta) - \mathrm{E}(\xi)\,\mathrm{E}(\eta)}{\sqrt{\mathrm{D}\,\xi}\sqrt{\mathrm{D}\,\eta}} = \sqrt{\frac{7}{55}} \approx 0.36.$$

T7

Так как доска небольшая, то, имея калькулятор, ничего в принципе не мешает просто посчитать количество доступных путей:

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	3	6	10	15	21	28	36	45	55	66	78	91	105	120
1	4	10	20	35	56	84	120	165	220	286	364	455	560	680
1	5	15	35	70	126	210	330	495	715	1001	1365	1820	2380	3060
1	6	21	56	126	252	462	792	1287	2002	3003	4368	6188	8568	11628
1	7	28	84	210	462	924	1716	3003	5005	8008	12376	18564	27132	38760
1	8	36	120	330	792	1716	3432	6435	11440	19448	31824	50388	77520	116280
1	9	45	165	495	1287	3003	6435	12870	24310	43758	75582	125970	203490	319770

Таблица 1: Заполненная количеством доступных путей доска к задаче №T7

Расчёт происходит из предположения о том, что у количество путей $N[i, j]$ равно

$$N[i, j] = N[i-1, j] + N[i, j-1],$$

а левый столбей и верхняя строка «заполнены» единицами – существует единственный способ добраться до этой клетки.

Если нас интересует движение такое, что последние три клетки были сделаны по короткой стороне доски, то вероятность такого маршрута:

$$\mathrm{P}_0 = \frac{11628}{319770} = \frac{2}{55} \approx 3.64 \cdot 10^{-2}.$$

Вообще можно заметить в числах биномиальные коэффициенты – действительно, достигая[1] $[i, j]$ клетки, мы делаем i шагов вправо и j вниз, то есть необходимо в $i+j$ элементах выбрать i элементов (или j), тогда искомая

[1] Формулы удобнее выглядят, когда i и j нумеруются с 0.

вероятность

$$P_0 = \binom{14+8}{8} \Big/ \binom{14+5}{5} = \frac{2}{55},$$

что сходится с прямым вычислением.

T8

Известно следующее совместное распределение:

$\xi\backslash\eta$	-2	0	2
-1	α	β	$2/13$
1	$3/13$	$2/13$	$1/13$

, $\qquad \alpha+\beta = \dfrac{5}{13}.$

где также известно, что $7\,\mathrm{D}(\xi) = 19\,\mathrm{D}(\eta)$.

Для начала найдём первые и вторые моменты для ξ и η:

$$\mathrm{E}(\xi) = -2\cdot\left(\alpha + \frac{3}{13}\right) + 2\cdot\frac{3}{13} = -2\alpha,$$

$$\mathrm{E}(\eta) = -1\cdot\left(\alpha+\beta+\frac{2}{13}\right) + 1\cdot\frac{6}{13} = -\frac{1}{13},$$

$$\mathrm{E}(\xi^2) = 4\cdot\left(\frac{6}{13}+\alpha\right) = \frac{24}{13} + 4\alpha,$$

$$\mathrm{E}(\eta^2) = 1.$$

Теперь можем перейти к квадратному уравнению

$$19\cdot\left(1-\frac{1}{13^2}\right) = 7\left(\frac{24}{13}+4\alpha-4\alpha^2\right), \quad \Rightarrow \quad 13^2(x^2-x)+36=0, \quad \Rightarrow \quad x_1 = \frac{4}{13}, \quad x_2 = \frac{9}{13}.$$

Но, так как $\alpha+\beta = 5/13$, а также $\operatorname{sign}\alpha = \operatorname{sign}\beta = 1$, то $\alpha < 5/13$, а значит искомая величина

$$\alpha = \frac{4}{13}.$$

2 Контрольная работа №2

Первая задача

Вероятность выпадения решки равна $p_0 = 0.42$. Монетка подброшена 1000 раз, и решка выпала 360 раз. Сколько раз необходимо подрбросить такую же монетку, чтобы доля выпавших решек отличалась от p_0 менее, чем в первые 1000 бросков с вероятностью $p_1 = 0.95$.

Thr 2.1 (ЦПТ Ляпунова)**.** *Пусть ξ_1, ξ_2, … – независимые и одинакоово распределенные случайные величины с конечной и ненулевой дисперсией: $0 < \mathrm{D}\,\xi_1 < \infty$. Тогда имеет место слабая сходимость*

$$\frac{S_n - n\,\mathrm{E}\,\xi_1}{\sqrt{n\,\mathrm{D}\,\xi_1}} \underset{n\to\infty}{\to} \mathrm{N}_{0,1}, \qquad S_n = \xi_1 + \ldots + \xi_n.$$

последовательности центрированных и нормированных сумм случайных величин к стандартному нормальному распределению.

Точнее S_n стремится к N_{a,σ^2}, где в пределах данной задачи верно, что $A = n\,\mathrm{E}\,\xi_1 = 0.42n$, а $\sigma = \sqrt{n\,\mathrm{D}\,\xi_1} = \sqrt{n0.42(1-0.42)} = \sqrt{npq}$.

По условиям задачи требуется попадание в интервал $[a,b] = [0.36n,\, 0.48n]$. Тогда

$$\int_a^b \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-A)^2}{2\sigma^2}\right) = \Phi\left(\sqrt{n}\frac{A-a}{\sqrt{2pq}}\right) = 0.95, \quad \Rightarrow \quad n = \frac{2X^2pq}{(A-a)^2} = 260,$$

где $X = 1.38$ можно найти по таблице.

Вторая задача

Известно, что ковариационная матрица случайного вектора $(X,\, Y,\, Z)^{\mathrm{T}}$ равна

$$M = \begin{pmatrix} 2 & -1 & \lambda \\ -1 & 2 & 1 \\ \lambda & 1 & 3 \end{pmatrix}.$$

Хочется найти все возможные значения λ, а также λ, соответсвующий минимальной вариации величины $\xi = X + \lambda Y - 2Z$.

Для начала поймём возможные значения λ: матрица неотрицательно определена, а значит, по критерию Сильвестра:

$$\det M = 7 - 2\lambda - 2\lambda^2 > 0, \qquad \Rightarrow \Big/ \lambda_{1,2} = \frac{1}{2}\left(-1 \pm \sqrt{15}\right) \Big/ \Rightarrow \qquad \lambda \in \left[-\frac{1}{2} - \frac{\sqrt{15}}{2};\; -\frac{1}{2} + \frac{\sqrt{15}}{2}\right].$$

Теперь можем найти оптимальное значение λ для $\boldsymbol{\xi} = X + \lambda Y - 2Z = (1,\, \lambda,\, -2)^{\mathrm{T}}$ в базисе (X, Y, Z):

$$\boldsymbol{\xi}^{\mathrm{T}} M \xi = 2\lambda^2 - 10\lambda + 14 = 2\left(\lambda - \frac{5}{2}\right)^2 + \frac{3}{2}, \qquad \Rightarrow \qquad \lambda = \frac{5}{2}.$$

Однако можно заметить, что верхняя граница $(-1 + \sqrt{15})/2 \approx 1.44 < 2.5$, следовательно минимум достигается на правой границе $\lambda_{\mathrm{opt}} = \frac{1}{2}\left(-1 + \sqrt{15}\right)$.

Четвертая задача

Известно, что плотность распределения переменной Y дана

$$f_Y(y) = C \exp\left(-y^2 + 4y - 10\right), \qquad y \in \mathbb{R}.$$

Найдём константу C, а также матожидание и дисперсию Y.

Заметим, что $f_Y(y) = C \exp\left(-(y-2)^2 - 6\right) = \frac{C}{e^6} \exp\left(-(y-2)^2\right)$, – нормальное распределение с $\mathrm{E}\, Y = a = 2$. Осталось найти

$$\int_{-\infty}^{+\infty} \frac{C}{e^6} \exp\left(-(y-2)^2\right)\, dy = \sqrt{\pi} C e^{-6} = 1, \qquad \Rightarrow \qquad C = \frac{e^6}{\sqrt{\pi}}.$$

Итого, распределение перепишется в виде

$$f_\xi(y) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{2}}} \exp\left(-\frac{(y-2)^2}{2(\frac{1}{\sqrt{2}})^2}\right),$$

собственно $\mathrm{D}\, Y = \sigma^2 = 1/2$.

Пятая задача

Случайные величины $X_1, \ldots, X_{100}$ независимы и одинаково распределены, в частности с $N(0, \sigma^2)$, $\sigma = 2$. Найдём распределение вектора $(Y,\, Z)^{\mathrm{T}}$, где $Y = X_{61} + X_{62} + \ldots + X_{100}$, и $Z = X_1 + X_2 + \ldots + X_{80}$.

Lem 2.2. *Если две случайные величины $\xi \in N_{a_1, \sigma_1^2}$, $\eta \in N_{a_2, \sigma_2^2}$ независимы, то $\xi + \eta \in N_{a_1 + a_2, \sigma_1^2 + \sigma_2^2}$.*

Lem 2.3. *Пусть есть две величины $\xi_1 \in N_{a_1, \sigma_1^2}$ и $\xi_2 \in N_{a_2, \sigma_2^2}$ с ковариациоей σ_{12}, тогда $\det \Sigma = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$, где $\rho = \sigma_{12}/(\sigma_1 \sigma_2)$ – коэффицент корреляции. Плотность двумерного нормального распределения в этом случае:*

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}^2 \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1 - a_1)^2}{\sigma_1^2} - \rho \frac{2(x_1 - a_1)(x_2 - a_2)}{\sigma_1 \sigma_2} + \frac{(x_2 - a_2)^2}{\sigma_2^2}\right]\right).$$

И это замечательно, ведь по условию X_i и X_j независимы при $i \neq j$, а тогда Y и Z распределены нормально. Получается, остается найти ковариацию $\rho(Y, Z)$, и мы найдём искомое распределение.

Для начала, $\mathrm{D}\, Y = 40\sigma^2$, $\mathrm{D}\, Z = 80\sigma^2$, также по условию $a_1 = a_2 = 0$. Ковариация Y и Z по линейности:

$$\operatorname{cov}\left(\textstyle\sum_{i=61}^{100} X_i,\, \sum_{j=1}^{80} X_j\right) = \sum_{i=61}^{100} \sum_{j=1}^{80} \operatorname{cov}(X_i, X_j) = 20\sigma^2, \qquad \Rightarrow \qquad \rho = \frac{20}{\sqrt{40 \cdot 80}} = \frac{1}{2\sqrt{2}} = 2^{-3/2}.$$

где учтено, что

$$\operatorname{cov}(X_i,\, X_i) = \mathrm{E}\, X_i^2 - \mathrm{E}^2\, X_i = \mathrm{D}\, X_i = \sigma^2.$$

Итого, искомое распределение:

$$f(Y, Z) = \frac{1}{2\pi\sigma^2 \cdot 20\sqrt{7}} \exp\left(-\frac{1}{140\sigma^2}\left[2Y^2 - YZ + Z^2\right]\right).$$