Отрицательно биномиальное распределение.

1. Обозначение семейства распределений.

$$\mathcal{NBIN} \tag{1}$$

2. Параметрическое пространство.

$$\Theta = \{\theta = (\nu, \tau)^T,\ \nu > 0,\ 0 < \tau < 1\} \tag{2}$$

3. Обозначение и область значений случайной величины.

$$NBIN(\nu, \tau) \in \mathsf{N}_0. \tag{3}$$

4. Функция распределения и плотность распределения.

4.1. Функция распределения и связанные с ней характеристики.

$$\mathcal{L}(NBIN(\nu, \tau)) = \mathcal{NBIN}(r; \nu, \tau),\ r \in \mathsf{N}_0, \tag{4}$$

$$\mathcal{NBIN}(r; \nu, \tau) = \sum_{j=0}^{r} nbin(j; \nu, \tau),\ r \geq 0,\ r - \text{целое}. \tag{5}$$

4.1.1. Если обозначить

$$\overline{\mathcal{NBIN}}(r; \nu, \tau) = 1 - \mathcal{NBIN}(r; \nu, \tau) = \sum_{j=r+1}^{\infty} nbin(j, \nu, \tau), \tag{6}$$

то получим представление функции надежности $\mathcal{NBIN}$ – распределений через функции распределения семейства бета-распределений и функции надежности биномиального семейства распределений:

$$\begin{aligned}\overline{\mathcal{NBIN}}(r-1; \nu, \tau) &= \mathcal{BETA}(1-\tau; r, \nu) = \\ &= 1 - \mathcal{BIN}(r-1; \nu + r - 1, 1 - \tau).\end{aligned} \tag{7}$$

4.1.2.

$$\mathcal{NBIN} = (r; \nu, \tau) = \mathcal{BETA}(\tau; \nu, r+1) \tag{8}$$

4.1.3. Аппроксимации для функции отрицательно биномиального распределения проводятся на основе существующей связи между семействами распределений $\mathcal{NBIN}$, $\mathcal{BIN}$ и $\mathcal{BETA}$ (см. 4.1.1, 4.1.2, а также справочник [1]).

4.2. Плотность распределения и связанные с ней характеристики.

$$\mathcal{L}'(NBIN(\nu, \tau)) = nbin(r; \nu, \tau), \tag{9}$$

где

$$nbin(r; \nu, \tau) = \frac{\Gamma(\nu + r)}{\Gamma(\nu) r!} \tau^{\nu} (1-\tau)^r, \quad r = 0, 1, 2 \ldots. \tag{10}$$

Если ν – целое, то

$$nbin(r; \nu, \tau) = C^r_{\nu+r-1} \tau^{\nu} (1-\tau)^r, \quad r = 0, 1, 2, \ldots. \tag{11}$$

т.е.

4.2.1. **Рекуррентное соотношение.**

$$\frac{nbin(r+1;\nu;\tau)}{nbin(r;\nu,\tau)} = \frac{(\nu+r)}{(r+1)}(1-\tau), \quad r = 0,1,\ldots \tag{12}$$

используется для последовательного вычисления вероятностей $nbin(r;\nu,\tau)$.

4.2.2. **Модальное значение.**

$$r_{mod} = \begin{cases} \left[(\nu-1)\,\frac{1-\tau}{\tau}\right], & \text{если } (\nu-1)\frac{1-\tau}{\tau} \text{ не целое число} \\ \left.\begin{matrix} \left[(\nu-1)\frac{1-\tau}{\tau}\right] \\ \frac{\nu(1-\tau)-1}{\tau} \end{matrix}\right\} & \text{две моды, если} (\nu-1)\frac{1-\tau}{\tau} \ - \text{ целое число} \\ 0, & \text{если } \nu\frac{1-\tau}{\tau} < \frac{1}{\tau}. \end{cases} \tag{13}$$

4.2.3. При фиксированных r и ν

$$nbin(r;\nu,\tau) \downarrow \tau; \tag{14}$$

при фиксированных r и τ

$$nbin(r;\nu,\tau) \uparrow \nu. \tag{15}$$

5. Интерпретация и области применения.

Это распределение иногда называют еще распределением *Пойа*. Если ν –целое, то семейство распределений $\mathcal{NBIN}$ известно также под названием семейства распределений *Паскаля*. В этом случае величина $\nu + NBIN(\nu,\tau)$ интерпретируется как число испытаний Бернулли с вероятностью успеха в отдельном испытании равном τ до появления ν-го успеха, или $NBIN(\nu,\tau)$ – число осуществления неудач до появления ν-го успеха. Если $\nu = 1$, то имеем дело с семейством геометрических распределений. В страховой сфере (и в частности, в автостраховании) семейство $\mathcal{NBIN}$ используется для моделирования числа страховых случаев.

6. Стохастические представления и тождества.

6.1. Если $NBIN(\nu_1,\tau)$ и $NBIN(\nu_2,\tau)$ независимы, то

$$NBIN(\nu_1,\tau) + NBIN(\nu_2,\tau) \stackrel{d}{=} NBIN(\nu_1+\nu_2,\tau). \tag{16}$$

6.1.1. Если ν – целое положительное число, а $Z_1,\ldots Z_\nu$ – н.о.р. случайные величины,

$$Z_1 \stackrel{d}{=} GEOM(\tau), \tag{17}$$

то

$$NBIN(\nu,\tau) \stackrel{d}{=} \sum_{i=1}^{\nu} Z_i \tag{18}$$

6.2. **Модель смешанного распределения.** Пусть η и λ – случайные величины,

$$\eta/\lambda \ \stackrel{d}{=}\ POIS(\lambda t), \tag{19}$$

$$\lambda \ \stackrel{d}{=}\ G(\alpha,\sigma). \tag{20}$$

Тогда

$$\eta \stackrel{d}{=} NBIN(\nu,\tau), \tag{21}$$

где

$$\nu \ = \ \alpha, \tag{22}$$

$$\tau \ = \ \frac{1}{1+\sigma t}. \tag{23}$$

6.3. **Модель составного распределения.** Пусть $y_1, \ldots, y_\nu$ – н.о.р. случайные величины,

$$y_1 \stackrel{d}{=} LOG(\varkappa), \quad 0 < \varkappa < 1, \tag{24}$$

независящие от случайной величины ν, где

$$\nu = POIS(\lambda). \tag{25}$$

Тогда

$$\sum_{i=1}^{\nu} y_1 \stackrel{d}{=} NBIN\left(-\frac{\lambda}{\ln(1-\varkappa)}, 1-\varkappa\right). \tag{26}$$

6.4. Если $\nu \to \infty$, то

$$\frac{NBIN(\nu,\tau) - \nu\frac{1-\tau}{\tau}}{\sqrt{\nu\frac{1-\tau}{\tau^2}}} \stackrel{d}{=} N(0,1) + 0_d(1) \tag{27}$$

6.5. Если $\nu \to \infty$, $\frac{1-\tau}{\tau} \to 0$, а $\nu\frac{1-\tau}{\tau} \to \lambda < \infty$, то

$$NBIN(\nu,\tau) = POIS(\lambda) + 0_p(1). \tag{28}$$

7. Характеристические преобразования распределений.

7.1. Характеристическая функция.

$$\chi(t) = \left(\frac{\tau}{1-(1-\tau)\exp\{it\}}\right)^{\nu}. \tag{29}$$

7.2. Производящая функция моментов.

$$\mu(t) = \left(\frac{\tau}{1-(1-\tau)e^t}\right)^{\nu}, \quad t < \ln\frac{1}{1-\tau}. \tag{30}$$

7.3. Производящая функция.

$$g(t) = \left(\frac{\tau}{1-(1-\tau)z}\right)^{\nu}, \quad |z| < \frac{1}{1-\tau}. \tag{31}$$

7.4. Преобразование Лапласа.

$$\lambda(t) = \left(\frac{\tau}{1-(1-\tau)e^{-t}}\right)^{\nu}, \quad t \geq 0. \tag{32}$$

7.5. Другие характеристические преобразования.

7.5.1. Производящая функция семиинвариантов.

$$k(t) = \nu\ln\tau - \nu\ln(1-(1-\tau)\exp t). \tag{33}$$

8. Моментные характеристики случайной величины.

Моментные характеристики отрицательно биномиального распределения легко получить из соответствующих характеристик биномиального распределения, осуществляя подстановку

$$\begin{cases} n = -\nu \\ \theta = 1 - \frac{1}{\tau}. \end{cases} \tag{34}$$

8.1. Начальные моменты.

$$a_1 = \nu\frac{1-\tau}{\tau}, \tag{35}$$

$$a_2 = \frac{\nu(1-\tau)}{\tau^2}\{1+\nu(1-\tau)\}, \tag{36}$$

$$a_3 = \frac{\nu(1-\tau)}{\tau^3}\{1+(1+3\nu)(1-\tau)+\nu^2(1-\tau)^2\}, \tag{37}$$

$$\begin{aligned} a_4 &= \frac{\nu(1-\tau)}{\tau^4}\{1+(4+7\nu)(1-\tau)+ \\ &+ (1+4\nu+6\nu^2)(1-\tau)^2+\nu^3(1-\tau)^3\}. \end{aligned} \tag{38}$$

8.2. Центральные моменты.

$$m_2 = \frac{\nu(1-\tau)}{\tau^2}, \tag{39}$$

$$m_3 = \frac{\nu(1-\tau)}{\tau^3}(2-\tau), \tag{40}$$

$$m_4 = \frac{\nu(1-\tau)}{\tau^4}\{1+(4+3\nu)(1-\tau)+(1-\tau)^2\}. \tag{41}$$

8.3. Факториальные моменты.

$$f_1 = \nu\frac{1-\tau}{\tau}, \tag{42}$$

$$f_2 = \nu(\nu+1)\left(\frac{1-\tau}{\tau}\right)^2, \tag{43}$$

$$f_3 = \nu(\nu+1)(\nu+2)\left(\frac{1-\tau}{\tau}\right)^3. \tag{44}$$

$$f_4 = \nu(\nu+1)(\nu+2)(\nu+3)\left(\frac{1-\tau}{\tau}\right)^4, \tag{45}$$

$$\ldots \quad \ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$f_k = \nu(\nu+1)\cdots(\nu+k-1)\left(\frac{1-\tau}{\tau}\right)^k, \quad k = 1, 2 \ldots. \tag{46}$$

8.4. Семиинварианты.

$$k_1 = \nu\frac{1-\tau}{\tau}, \tag{47}$$

$$k_2 = \frac{\nu(1-\tau)}{\tau^2}, \tag{48}$$

$$k_3 = \frac{\nu(1-\tau)}{\tau^3}(2-\tau), \tag{49}$$

$$k_4 = \frac{\nu(1-\tau)}{\tau^4}\{1+4(1-\tau)+(1-\tau)^2\}, \tag{50}$$

$$k_{r+1(\gamma,0)} = (1-\tau)\frac{\partial k_r(\nu,\tau)}{\partial(1-\tau)}. \tag{51}$$

8.5. Коэффициент вариации.

$$\gamma_0 = \sqrt{\nu(1-\tau)}. \tag{52}$$

8.6. Коэффициент ассиметрии.

$$\gamma_1 = \frac{2-\tau}{\sqrt{\nu(1-\tau)}}. \tag{53}$$

8.7. Коэффициент эксцесса.

$$\gamma_2 = \frac{\tau^2+6(1-\tau)}{\nu(1-\tau)}. \tag{54}$$

8.8. Другие моментные характеристики.

8.8.1. Среднее уклонение.

$$b_1 = E|NBIN(\nu,\tau) - ENBIN(\nu,\tau)| = \frac{2m(\nu+m-1)!\tau^{\nu-1}(1-\tau)^m}{m!(\nu-1)!}, \tag{55}$$

где $m = \left[\frac{\nu(1-\tau)}{\tau_1}+1\right] = [a_1+1]$.

8.8.2. Индекс рассеяния.

$$\nu_2 = \frac{m_2}{a_1} = \frac{1}{\tau}. \tag{56}$$

9. Информационные функции.

9.1. Информационная функция Фишера.

9.1.1. Пусть ν – известная величина, т.е. единственным неизвестным параметром является τ. Тогда

$$I^0(\tau) = \frac{\nu}{\tau^2(1-\tau)}. \tag{57}$$

9.1.2. Если оба параметра и ν, и τ не известны, тогда $I^0(\tau,\nu)$ имеет весьма сложный вид

10. Аналитические свойства распределений.

10.1. Семейство $\mathcal{NBIN}$ является безгранично делимым, поскольку для любого n, $n = 1, 2, \ldots$,

$$\left(\chi\left(t; NBIN\left(\frac{\nu}{n},\tau\right)\right)\right)^n = \chi(t; NBIN(\nu,\tau)). \tag{58}$$

10.2. О связи "хвостов"распределения отрицательно-биномиальных и биномиальных распределений для целых ν

$$\mathbf{P}\{NBIN(\nu,\tau)\geq m\}=\mathbf{P}\{BIN(m+\nu-1,\tau)\geq m\}. \tag{59}$$

10.2.1.

$$\begin{aligned}\mathbf{P}\{NBIN(\nu,\tau)\geq m\} &= \sum_{k=m}^{\nu+m-1} b(k;\nu+m-1,\tau)= \\ &= \frac{(\nu+m-1)!}{(\nu-1)!(m-1)!}\int\limits_0^{\tau} v^{m-1}(1-v)^{\nu-1}\alpha v.\end{aligned} \tag{60}$$

10.3. Семейство $\mathcal{NBIN}$ имеет убывающую интенсивность отказа при $\nu<1$, возрастающую при $\nu>1$ и постоянную (свойство отсутствия памяти) при $\nu=1$, т.е. при геометрическом распределении.

10.4. Если $NBIN(\nu_1,\tau)$ и $NBIN(\nu_2,\tau)$ независимы, то

$$NBIN(\nu_1,\tau)+NBIN(\nu_2,\tau)\stackrel{d}{=}NBIN(\nu_1+\nu_2,\tau). \tag{61}$$

10.5. Пусть $\nu\to\infty$, $\tau\to 1$, $\nu\frac{1-\tau}{r}\to\theta$.
Тогда

$$NBIN(\nu,\tau)\stackrel{d}{=}POIS(\theta)+0_d(1). \tag{62}$$

10.6. Пусть $\tau\to 1$, $\nu=const$.
Тогда $NBIN(\nu,\tau)\stackrel{d}{=}0_d(1)$.

10.7. **Нормализующие преобразования.**

Принимая обозначения Sh и Sh^{-1} для прямой и обратной функций гиперболического синуса получим:

10.7.1. При $\nu\to\infty$ (см. [1])

$$\sqrt{\nu}sh^{-1}\sqrt{\frac{NBIN(\nu,\tau)}{\nu}}\stackrel{d}{=}N(0,1)+0_d(1). \tag{63}$$

10.7.2. **Преобразование Anscomb (1948).**

$$\sqrt{\nu-0.5}sh^{-1}\sqrt{\frac{NBIN(\nu,\tau)+3/8}{\nu-3/4}}\stackrel{d}{=}N(0,1)+0_d(1). \tag{64}$$

11. Моделирование случайной величины.

12. Статистические выводы.

12.1. Достаточные статистики.

12.1.1. Параметр ν – известная величина.
Пусть $y=(y_1,\ldots,y_n)$, $y_i\in\mathsf{N}$, y_i – н.о.р. случайные величины, $i=\overline{1,n}$,

$$y_1\stackrel{d}{=}NBIN(\nu,\tau). \tag{65}$$

Тогда $S=\sum\limits_{i=1}^{n}y_i$ – полная достаточная статистика,

$$S\stackrel{d}{=}NBIN(n\nu,\tau). \tag{66}$$

12.1.2. Оба параметра не известны. Достаточной статистикой является вариационный ряд $(y_{(1)}, \dots y_{(n)})$.

12.2. Точечное оценивание.

12.2.1. Параметр ν – известная величина.
ОМП имеет вид

$$\check{\tau}_n = \frac{n\nu}{S + n\nu}, \tag{67}$$

а

$$\hat{\tau}_n^0 = \frac{n\nu - 1}{S + n\nu - 1} \tag{68}$$

несмещенная оценка с минимальной дисперсией.
Смещением ОМП является

$$b(\tau) = E\{\check{\tau}_n - \hat{\tau}_n^0\} = E\left\{\frac{S}{(S + n\nu)(S + n\nu - 1)}\right\} > 0. \tag{69}$$

Список литературы

[1] Джонсон Н.Л., Коц С., Кемп А. *Одномерные дискретные распределения.* БИНОМ, Лаборатория знаний, М., 2010.