32. Дискретни разпределения. Задачи, в които възникват. Моменти – математическо очакване и дисперсия.

Случайните величини са числови функции, определени върху множеството от елементарни събития Ω, но тяхното определение силно зависи от това кои са случайните събития Ω.
Пълна група от събития наричаме множеството от всички събития (непресичащи се), които могат да се „случат" в резултат от даден опит. При това обединението от събитията в пълната група съвпада с Ω.
Нека е зададена пълната група събития (H_1, H_2, …, H_n). Ще казваме, че е определена ***проста случайна величина*** ξ, ако $\xi(\omega)=x_i$ за всяко $\omega\in H_i$, i=1, 2, …, n, където x_i е реално число. Тоест случайната величина е реална функция, съпоставяща на събитията от пълна група събития реални числа.
Ще казваме, че функцията ξ(ω), определена на Ω със стойности в R, е ***случайна величина***, ако за всяко $x\in R^1$ множеството $\{\omega: \xi(\omega) < x\} \in \mathfrak{A}$ (булевата σ-алгебра).
Реалната функция P, определена върху елементите на булевата σ-алгебра $\mathfrak{A}$, се нарича ***вероятност***, ако удовлетворява условията:

1. *Неотрицателност*: $P(A) \geq 0, \forall A \in \mathfrak{A}$;
2. *Нормираност:* $P(\Omega) = 1$;
3. *Адитивност:* $P(A_1 + A_2 + \cdots) = P(A_1) + P(A_2) + \cdots$.

Ще наричаме ***функция на разпределение*** на случайната величина ξ функцията F(x) = P(ω: ξ(ω) < x). Тоест функцията на разпределение е вероятността случайната величина да приеме стойност, по-малка от дадена.
Функцията F(x) е монотонно ненамаляваща и непрекъсната отляво. Освен това F(-∞)=0 и F(∞)=1.
Случайна величина, която приема само стойностите x_1, x_2, x_3, … с вероятности съответно p_1, p_2, p_3, ... се нарича ***дискретна случайна величина***.
За стойностите p_1, p_2, … е изпълнено, че $\Sigma p_i=1$ и $1\geq p_i\geq 0$. За дискретните случайни величини функцията на разпределение F(x) има само скокове в точките x_i и навсякъде другаде е константа. В точката x_i скокът ѝ е равен точно на числото p_i.
Ако случайната величина е такава, че за всяко x съществува производна на функцията на разпределение f(x)=F'(x), то ще я наричаме ***непрекъсната случайна величина***. Производната наричаме ***плътност***.
Случайна величина, приемаща за стойности натуралните числа, наричаме ***целочислена случайна величина***.
Нека ξ е дефинирана във вероятностното пространство $(\Omega, \mathfrak{F}, P)$, а $F(x) = P(\xi < x), x \in R_1$ е нейната функция на разпределение. Средна стойност (***математическо очакване***) на ξ – означава се с $\mathbb{E}\xi$, се дефинира с равенството:

$$\mathbb{E}\xi = \int_{-\infty}^{\infty} x dF(x)$$

при условие, че този интеграл (в смисъл на Лебег-Стилтес) е абсолютно сходящ, тоест когато $\mathbb{E}\{|\xi|\} < \infty$.
Математическо очакване на проста случайна величина ξ, приемаща стойности $x_1, x_2, \dots, x_n$ върху събитията от пълната група, определяме като $\sum_{k=1}^{n} x_k P(H_k)$.
Дисперсията на случайната величина ξ се определя като числото $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2$, като тя може и да е безкрайна. За непрекъснати и дискретни разпределения, тя се смята съответно по формулите:

32. Дискретни разпределения. Задачи, в които възникват. Моменти – математическо очакване и дисперсия.

$$\mathbb{D}\xi = \int (x-\mathbb{E}\xi)^2 f(x)dx\,,\ \ \mathbb{D}\xi = \sum_i (x-\mathbb{E}\xi)^2 p_i$$

Числата $a_k = \mathbb{E}\{(\xi - c)^k\}$ се наричат ***момент от k-ти ред*** на случайната величина ξ относно константата *c*. При $c = 0$ тези моменти се наричат ***начални***, а при $c = \mathbb{E}\xi$ – ***централни***.

Пораждащата функция на целочислена случайна величина ξ се задава с формулата:

$$p(s) = \mathbb{E}s^{\xi}$$

Пораждащата функция е удобна, защото винаги съществува при достатъчно малко *s*, например при $s \leq 1$.

Свойства:

- $p(1) = \mathbb{E}1^{\xi} = \mathbb{E}1 = 1$
- $p(0) = \mathbb{E}0^{\xi} = P(\xi = 0)$
- $p'(1) = \left(\mathbb{E}1^{\xi}\right)' = \mathbb{E}\left(1^{\xi}\right)' = \mathbb{E}\xi$, когато съществува
- $p''(1) = \mathbb{E}\xi(\xi-1) = \mathbb{E}\xi^2 - \mathbb{E}\xi$, когато съществува
- Когато $\xi \perp \eta, p_{\xi+\eta}(s) = p_{\xi}(s)p_{\eta}(s)$.

Редица от независими, еднакво разпределени случайни величини $\{\xi_i$, i=1, 2, …}, всяка от които приема две стойности: 1 и 0 с вероятности съответно p и q=1-p, наричаме ***схема на Бернули***.

Разглеждаме сумата η_n на n случайни величини от схемата на Бернули. Това е целочислена случайна величина, приемаща стойности от 0 до n. Тази случайна величина ще я разглеждаме като брой успехи от n опита с постоянна вероятност p за успех във всеки опит. Разпределението на такава случайна величина наричаме ***биномно***. Биномните вероятности се пресмятат по формулата:

$$b(n,k,p) = \binom{n}{k} p^k q^{n-k}$$

Тогава: $\mathbb{E}\eta_n = \sum_{i=1}^{n} \mathbb{E}\xi_i = n\mathbb{E}\xi_1 = n(1.p + 0.q) = np$ и $\mathbb{D}\eta_n = \sum_{i=1}^{n} \mathbb{D}\xi_i = n\mathbb{D}\xi_1 = n(\mathbb{E}\xi_1^2 - (\mathbb{E}\xi_1)^2) = n(p-p^2) = npq$.

В задачите, където искаме да намерим вероятността първият успех в серия от опити, във всеки от които вероятността за неуспех да е *p*, да постигнем на *m*-тия опит, ще използваме геометричното разпределение.

Казваме, че целочислената случайна величина ξ има ***геометрично разпределение***, ако

$$P(\xi = m) = p^m q, m = 0, 1, 2, \dots$$

Тогава математическото очакване и дисперсията на геометричното разпределение ще бъдат съответно

$$\mathbb{E}\xi = \sum_{k=0}^{\infty} kp^k q = q\sum_{k=0}^{\infty} kp^k = qp\sum_{k=0}^{\infty} kp^{k-1} = qp\frac{d}{dp}\left(\frac{1}{1-p}\right) = \frac{p}{q}$$

$$\begin{aligned}\mathbb{D}\xi &= \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - \mathbb{E}\xi + \mathbb{E}\xi - (\mathbb{E}\xi)^2 = \mathbb{E}\xi(\xi-1) + \mathbb{E}\xi - (\mathbb{E}\xi)^2 \\ &= q\sum_{k=0}^{\infty} k(k-1)p^k + \frac{p}{q} - \left(\frac{p}{q}\right)^2 = qp^2\frac{d^2}{dp^2}\left(\frac{1}{1-p}\right) + \frac{p}{q} - \left(\frac{p}{q}\right)^2 = 2\left(\frac{p}{q}\right)^2 \\ &+ \frac{p}{q} - \left(\frac{p}{q}\right)^2 = \left(\frac{p}{q}\right)^2 + \frac{p}{q}\end{aligned}$$

Казваме, че целочислената случайна величина има ***хипергеометрично разпределение***, ако

$$P(\xi = m) = \frac{\binom{n}{m}\binom{N-n}{M-m}}{\binom{N}{n}}$$

32. Дискретни разпределения. Задачи, в които възникват. Моменти – математическо очакване и дисперсия.

Да разгледаме следната задача. Нека е дадена партида, съдържаща N изделия, от които M са дефектни. Правим случайна извадка от $n < N$ изделия. Пита се каква е вероятността точно m от тях са дефектни. Оказва се, че разпределението на случайната величина брой дефектни е хипергеометричното, тъй като броят на всички възможни извадки (равновероятни) са $\binom{N}{n}$. „Благоприятните” са тези, които съдържат точно m дефектни детайла, могат да се получат чрез комбиниране на извадка от m дефектни и извадка от

$$\mathbb{E}\xi = \sum_{k=0}^{\infty} k \frac{\binom{n}{k}\binom{N-n}{M-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k=0}^{\infty} k \binom{M}{k}\binom{N-M}{n-k} = \frac{1}{\binom{N}{n}}.M\binom{N-1}{n-1}$$
$$= \frac{n!\,(N-n)!}{N!}.M\frac{(N-1)!}{(n-1)!\,(N-1-n+1)!} = \frac{n.M}{N}$$

(Обяснение:

$$\frac{1}{\binom{N}{n}}\sum_{k=0}^{\infty} k\binom{M}{k}\binom{N-M}{n-k} = \frac{M\binom{N-1}{n-1}}{\binom{N}{n}}\sum_{k=0}^{\infty}\frac{k\binom{M}{k}\binom{N-M}{n-k}}{M\binom{N-1}{n-1}} = \frac{M\binom{N-1}{n-1}}{\binom{N}{n}}\sum_{k=0}^{\infty}\frac{\binom{N-n}{M-k}\binom{n-1}{k-1}}{\binom{N-1}{M-1}},$$ но

$\sum_{k=0}^{\infty}\frac{\binom{N-n}{M-k}\binom{n-1}{k-1}}{\binom{N-1}{M-1}}$ е сума от вероятности – вероятност да изтеглим $n-1$ предмета, от които точно $k-1$ да са дефектни, като теглим от $N-1$, от които $M-1$ са дефектни. Но тъй като разглеждаме всички случаи за k, тоест това е сума от вероятностите на всички възможности, тоест е 1.)

$$\mathbb{D}\xi = \mathbb{E}\xi(\xi-1) + \mathbb{E}\xi - (\mathbb{E}\xi)^2 = \sum_{k=0}^{\infty} k(k-1)\frac{\binom{n}{k}\binom{N-n}{M-k}}{\binom{N}{n}} + \frac{nM}{N} - \frac{n^2M^2}{N^2}$$
$$= n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-1}{N-n}$$

Така ако отбележим вероятността при единичен избор да попаднем на дефектен елемент с $p = \frac{M}{N}$ за математическото очакване на хипергеометричното разпределение ще получим np, а за дисперсията му – $npq\frac{N-1}{N-n}$.

Поасоновото разпределение се определя като граница на биномни разпределения, когато $n \to \infty$, така че $np \to \lambda > 0$. Случайната величина може да приема всякакви целочислени стойности:

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

То е подходящо за моделиране на броя на случайни редки събития като брой радиоактивни разпадания на единица време.

Пораждащата функция на Поасоновото разпределение е:

$$p(s) = \mathbb{E}s^{\eta} = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k s^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda s)^k}{k!} = e^{-\lambda}.e^{\lambda s} = e^{\lambda(s-1)}$$

От нея смятаме математическото очакване и дисперсията на Поасоновото разпределение:

$$\mathbb{E}\xi = p'(1) = e^{\lambda(s-1)}.ln\left(e^{\lambda}\right) = e^{0}.\lambda = \lambda$$

$$\mathbb{E}\xi^2 - \mathbb{E}\xi = p''(1) = \left(e^{\lambda(s-1)}.ln\left(e^{\lambda}\right)\right)' = \lambda^2 e^{\lambda(s-1)} = \lambda^2 \Rightarrow \mathbb{E}\xi^2 = \lambda^2 + \lambda$$

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