

Тема 15

<https://github.com/v--/se2018>

Случайни величини с дискретни разпределения - дискретно равномерно, биномно, геометрично, поасоново разпределения. Задачи, в които възникват.

Янис Василев

Оригинал: 20 юни 2019

Ревизия: 91d5bc9 от 09 юни 2021

За всеки случай проверете дали няма по-нова ревизия

1. Теория

Теорията е представена представена с минимални препратки към теорията на мярката и е базирана частично на изложението в Боровков, *Теория вероятностей* и Димитров и Янев, *Вероятности и статистика*. За пълнота съм включил доказателства на основните свойства на пораждаци, пораждаци моментите и характеристична функции. За да се запази простотата на изложението, не са включени понятия като функция на разпределение и функция на вероятностите.

1.1. Анотация

Изложената анотацията е взета от *Конспект за ДИ за спец. статистика*.

1. Дефиниция за дискретно разпределение на случайна величина.
2. Неотрицателност и нормираност на вероятностите на дискретна случайна величина.
3. Дефиниция за моментите на дискретна случайна величина.
4. Дефиниция и свойства (без доказателства) на пораждаща/пораждаща моментите/характеристична функция (по избор).
5. Дефиниция, коректност, мотивиращ пример, пораждаща/пораждаща моментите/характеристична функция, очакване и дисперсия на две избрани от комисията дискретни разпределения.

1.2. Основни дефиниции и теореми

Определение 1. (Реална) случайна величина над вероятностното пространство (Ω, \mathcal{F}, P) наричаме всяка измерима функция $\xi : \Omega \rightarrow \mathbb{R}$.

Условието за измеримост на ξ може да се запише така: за всяко борелово множество $B \in \mathfrak{B}(\mathbb{R})$ имаме

$$\xi^{-1}(B) = \{\omega \in \Omega \mid \xi(\omega) \in B\} \in \mathcal{F}.$$

Разпределение на ξ наричаме мярката

$$P_\xi(A) := P(\xi \in A).$$

Две случайни величини ξ и η наричаме **независими**, ако за всички $A, B \in \mathcal{F}$ е изпълнено

$$P(\xi \in A, \eta \in B) = P_\xi(A) P_\eta(B).$$

Случайната величина ξ наричаме **дискретна** и казваме, че ξ има **дискретно разпределение**, ако тя приема крайно или изброимо много стойности x_1, x_2, \dots с вероятности съответно

$$p_k = P(\xi = x_k), k = 1, 2, \dots$$

Множеството от стойности на ξ ще бележим с $\text{img}(\xi) := \{x_1, x_2, \dots\}$.

Теорема 2. Редицата p_1, p_2, \dots са е редица от вероятности на някоя дискретна случайна величина тогава и само тогава, когато са изпълнени

$$\begin{aligned} p_k &\geq 0, k = 1, 2, \dots && \text{(неотрицателност)} \\ \sum_k p_k &= 1 && \text{(нормираност)} \end{aligned}$$

Теорема 2 ни дава обосновка да задаваме дискретни случайни величини изцяло чрез редици от вероятности, т.е. без изрично да задаваме вероятностни пространства. По тази причина понякога разпределение на дискретна случайна величина наричаме не съответната вероятностна мярка, а редицата от вероятностите.

Доказателство на теорема 2.

Доказателство на достатъчност. Нека ξ е дискретна случайна величина над вероятностното пространство (Ω, \mathcal{F}, P) със стойности x_1, x_2, \dots и вероятности p_1, p_2, \dots

Тъй като $P : \mathcal{F} \rightarrow [0, 1]$ е вероятностна мярка, директно получаваме

$$p_k = P(\xi = x_k) \in [0, 1]$$

за всяко $k = 1, 2, \dots$ и

$$\begin{aligned}
 \sum_k p_k &= \sum_k P(\xi = x_k) = \\
 &= \sum_k P(\{\omega \in \Omega \mid \xi(\omega) = x_k\}) = \\
 &= P\left(\bigcup_k \{\omega \in \Omega \mid \xi(\omega) = x_k\}\right) = \\
 &= P(\{\omega \in \Omega \mid \xi(\omega) \in \text{img}(\xi)\}) = \\
 &= P(\xi \in \text{img}(\xi)) = \\
 &= 1.
 \end{aligned}$$

Доказателство на необходимост. Нека редицата p_1, p_2, \dots удовлетворява условията **неотрицателност** и **нормираност**. Предполагаме, че това са вероятности за някакви елементарни събития $\Omega = \{\omega_1, \omega_2, \dots\}$.

Дефинираме случайната величина $\xi : \Omega \rightarrow \mathbb{R}$, $\xi(\omega_k) := p_k$, която е измерима над $(\Omega, \mathcal{P}(\Omega))$, тъй като праобразът $\xi^{-1}(A)$ на всяко борелово множество $A \in \mathfrak{B}(\mathbb{R})$ е подмножество на Ω .

Сега дефинираме $P(A) := \sum_{\omega \in A} \xi(\omega)$. Тази сума е добре дефинирана, тъй като редът $\sum_k p_k$ е абсолютно сходящ според **неотрицателност** и следователно всеки подред също е абсолютно сходящ.

P е вероятностна мярка над $(\Omega, \mathcal{P}(\Omega))$, тъй като

1. $P(\emptyset) = 0$,
2. $P(A) = \sum_{\omega \in A} \xi(\omega) \geq 0$ за произволно събитие $A \subseteq \Omega$ според **неотрицателност**,
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{\omega \in \bigcup_{i=1}^{\infty} A_i} \xi(\omega) = \sum_{i=1}^{\infty} \sum_{\omega \in A_i} \xi(\omega) = \sum_{i=1}^{\infty} P(A_i)$,
4. $P(A) \leq 1$ за произволно събитие $A \subseteq \Omega$ според доказаната адитивност и **нормираност**.

Така построихме дискретна случайна величина ξ над пространството $(\Omega, \mathcal{P}(\Omega), P)$ с вероятности p_1, p_2, \dots □

Твърдение 3. *Дискретните случайни величини над едно вероятностно пространство (Ω, \mathcal{F}, P) образуват линейно пространство относно операциите събиране и умножение с число.*

Доказателство. Ще докажем само затвореността относно операциите, тъй като останалите аксиоми за линейно пространство са изпълнени по тривиални причини.

Сумата $\xi + \eta$ на две дискретни случайни величини приема стойности $x + y$ за $x \in \text{img}(\xi)$ и $y \in \text{img}(\eta)$, които са не повече от изброимо много, следователно $\xi + \eta$ също е дискретна случайна величина.

Произведението $c\xi$ за $c \in \mathbb{R}$ приема или само една стойност при $c = 0$, или същият брой стойности като ξ , следователно $c\xi$ също е дискретна случайна величина. \square

1.3. Очакване и моменти

До края на темата ще считаме, че работим над вероятностното пространство (Ω, \mathcal{F}, P) .

Определение 4. Нека ξ е дискретна случайна величина със стойности x_1, x_2, \dots . Дефинираме **очакване** на ξ чрез

$$E(\xi) := \sum_k x_k P(\xi = x_k) = \sum_{\omega \in \Omega} \xi(\omega)P(\{\omega\}).$$

Казваме, че ξ има (крайно) очакване, ако горният ред е абсолютно сходящ.

Случайни величини с очакване нула наричаме **центрирани**.

Очакване от константа $x \in \mathbb{R}$ дефинираме да бъде самата константа x .

Забележка 5. Формално, полагаме $E(x) := E(I_x)$, където I_x е индикатор за x , т.е.

$$P(I_x = y) = \delta_{xy} = \begin{cases} 1, & x = y, \\ 0, & x \neq y. \end{cases}$$

Оттук директно следва $E(x) = E(I_x) = x$.

Твърдение 6. За независими дискретни случайни величини ξ и η с крайно очакване е изпълнено

$$E(\xi\eta) = E(\xi)E(\eta).$$

Доказателство. Нека x_1, x_2, \dots и y_1, y_2, \dots са стойностите съответно на ξ и η . Тогава стойностите на ξ и η имат вида $x_i y_j$ и

$$\begin{aligned} E(\xi)E(\eta) &= \left(\sum_i x_i P(\xi = x_i) \right) \left(\sum_j y_j P(\eta = y_j) \right) = \\ &= \sum_{k=1}^{\infty} \sum_{m=1}^k x_m y_{k-m} P(\xi = x_m) P(\eta = y_{k-m}) = \\ &= \sum_{k=1}^{\infty} \sum_{m=1}^k x_m y_{k-m} P(\xi = x_m, \eta = y_{k-m}) = \\ &= \sum_{i,j} x_i y_j P(\xi = x_i, \eta = y_j) = \\ &= \sum_{i,j} x_i y_j P(\xi\eta = x_i y_j) = \\ &= E(\xi\eta). \end{aligned}$$

\square

Твърдение 7. Очакването е линеен функционал над подпространството от дискретни случайни величини над (Ω, \mathcal{F}, P) , за които очакването съществува.

Доказателство. Директно проверяваме дефиницията за линеен функционал:

1. За две дискретни случайни ξ и η величини с крайно очакване имаме

$$\begin{aligned} E(\xi + \eta) &= \sum_{\omega \in \Omega} (\xi + \eta)(\omega)P(\{\omega\}) = \\ &= \sum_{\omega \in \Omega} [\xi(\omega) + \eta(\omega)]P(\{\omega\}) = \\ &= \sum_{\omega \in \Omega} \xi(\omega)P(\{\omega\}) + \sum_{\omega \in \Omega} \eta(\omega)P(\{\omega\}) = \\ &= E(\xi) + E(\eta). \end{aligned}$$

2. За дискретна случайна величина ξ със стойности x_1, x_2, \dots и крайно очакване и константа $c \in \mathbb{R}$ по твърдение 8 имаме

$$E(c\xi) = \sum_k (cx_k)P(\xi = x_k) = c \sum_k x_k P(\xi = x_k) = c E(\xi).$$

Видяхме и това, че дискретните случайни величини с крайно очакване са затворени относно събиране и умножение с число, с което доказваме, че те образуват линейно подпространство на всички дискретни случайни величини. \square

Твърдение 8. Нека ξ е дискретна случайна величина със стойности x_1, x_2, \dots и крайно очакване. Нека $\psi : \mathbb{R} \rightarrow \mathbb{R}$ е произволна функция. Тогава $\psi(\xi)$ е дискретна случайна величина и е изпълнено

$$E(\psi(\xi)) = \sum_k \psi(x_k)P(\xi = x_k).$$

Доказателство. Образът на крайно или измеримо множество е най-много измеримо, затова $\psi(\xi)$ приема не повече от изброимо много различни стойности и следователно е дискретна случайна величина. Нека стойностите на $\psi(\xi)$ са y_1, y_2, \dots . Преобразите $\psi^{-1}(y_i)$ на $y_i, i = 1, 2, \dots$ са крайни или изброими, непресичащи се и $\text{img}(\xi) = \bigcup_i \psi^{-1}(y_i)$. Имаме

$$\begin{aligned} E(\psi(\xi)) &= \sum_i y_i P(\psi(\xi) = y_i) = \\ &= \sum_i y_i \sum_{x \in \psi^{-1}(y_i)} P(\xi = x) = \\ &= \sum_i \sum_{x \in \psi^{-1}(y_i)} \psi(x)P(\xi = x) = \\ &= \sum_k \psi(x_k)P(\xi = x_k). \end{aligned}$$

\square

Доказаните в твърдения 6 to 8 свойства на очакването значително опростяват работата с него.

Определение 9. Ковариация на случайните величини ξ и η наричаме

$$\text{cov}(\xi, \eta) := E((\xi - E\xi)(\eta - E\eta)).$$

Дисперсия или вариация на случайната величина ξ наричаме

$$\begin{aligned} \text{var}(\xi) &:= \text{cov}(\xi, \xi) = E((\xi - E\xi)^2) = \\ &= E(\xi^2 - 2\xi E\xi + E(\xi)^2) = \\ &= E(\xi^2) - 2E(\xi)^2 + E(\xi)^2 = \\ &= E(\xi^2) - E(\xi)^2. \end{aligned}$$

Корелация на ξ и η наричаме

$$\rho(\xi, \eta) := \frac{\text{cov}(\xi, \eta)}{\sqrt{\text{var}(\xi)\text{var}(\eta)}}.$$

От неравенството на Коши-Буняковски-Шварц следва, че $|\rho(\xi, \eta)| \leq 1$.

Числото $E(\xi^n)$ наричаме n -ти **момент** на ξ , а $E((\xi - E\xi)^n)$ наричаме n -ти **централен момент** на ξ .

Очакването всъщност е просто първият момент, а дисперсията - вторият централен момент. Коренът на дисперсията се нарича **стандартно отклонение** и често се бележи със σ_ξ .

Две случайни величини се наричат **некорелирани** или **ортогонални**, ако ковариацията им е 0, защото ковариацията играе ролята на скалярно произведение в пространството L^2 от (всички, не непременно дискретни) случайни величини с краен втори момент.

Случайни величини със стандартно отклонение единица наричаме **нормирани**, тъй като стандартно отклонение играе ролята на норма в L^2 .

Твърдение 10. Ако две случайни величини са независими и имат крайно очакване, те са ортогонални.

Доказателство. Нека ξ и η са независими и имат крайни очаквания съответно μ и ν . Тогава

$$\text{cov}(\xi, \eta) = E((\xi - \mu)(\eta - \nu)) = E(\xi - \mu)E(\eta - \nu) = (\mu - \mu)(\nu - \nu) = 0 \cdot 0 = 0.$$

□

Твърдение 11. Ако $E(\xi^n)$ съществува, съществуват и моментите от по-нисък ред.

Доказателство. Първо да забележим, че за $y \in (0, 1)$ имаме $P(\xi = x_k)^y < P(\xi = x_k)$. Тъй като вероятностната мярка е нормирана, можем да приложим неравенството на Йенсен и да получим:

$$\begin{aligned} E(|\xi|^{n-1}) &\leq E(|\xi|^{n-1})^{\frac{n}{n-1}} = \\ &= \left(\sum_k |x_k|^{n-1} P(\xi = x_k) \right)^{\frac{n}{n-1}} \leq \\ &\leq \sum_k \left(|x_k|^{n-1} P(\xi = x_k) \right)^{\frac{n}{n-1}} < \\ &< \sum_k \left(|x_k|^{n-1} \right)^{\frac{n}{n-1}} P(\xi = x_k) = \\ &= \sum_k |x_k|^n P(\xi = x_k) = \\ &= E(|\xi|^n). \end{aligned}$$

□

1.4. Пораждащи, пораждащи моментите и характеристични функции

Определение 12. Пораждаща функция на ξ наричаме

$$\text{PGF}_\xi(z) := E(z^\xi).$$

Пораждаща моментите функция на ξ наричаме

$$\text{MGF}_\xi(t) := E(e^{t\xi}).$$

Характеристична функция на ξ наричаме

$$\varphi_\xi(t) := E(e^{it\xi}).$$

Изпълнено е $\varphi_\xi(t) = \text{MGF}_\xi(it) = \text{PGF}_\xi(e^{it})$.

Забележка 13. Не сме дефинирали очакване от комплексна случайна величина, но теоретичната обосновка идва от формулите на Ойлер:

$$\varphi_\xi(t) = E(e^{it\xi}) = E(\cos(t\xi) + i \sin(t\xi)) = E(\cos(t\xi)) + i E(\sin(t\xi)).$$

Забележка 14. Дефинициите за моменти и функции от очакването се пренасят без изменение за случайни величини, които не са дискретни. Пораждащите функции, обаче, се полезни само за случайни величини, които приемат неотрицателни цели стойности.

Теорема 15 (Свойства на пораждащите функции). Нека ξ и η са независими целочислени случайни величини със стойности $0, 1, \dots$ (възможно е $P(\xi = k) > 0$ само за краен брой k).

За пораждащата функция PGF_ξ са изпълнени следните свойства

1. $\text{PGF}_\xi(z)$ е аналитична функция поне в $-1 < z < 1$.
2. $P(\xi = m) = \frac{\text{PGF}_\xi^{(m)}(0)}{m!}$, където $\text{PGF}_\xi^{(m)}$ е m -тата производна на PGF_ξ .
3. Ако пораждащите функции на ξ , η и $\xi + \eta$ съществуват в точка $z \in \mathbb{R}$, имаме

$$\text{PGF}_{\xi+\eta}(z) = \text{PGF}_\xi(z) \text{PGF}_\eta(z)$$

4. Ако пораждащите функции на две целочислени дискретни случайни величини съвпадат, самите случайни величини съвпадат.

Доказателство.

1. $\text{PGF}_\xi(z)$ се дефинира чрез степенен ред. За да докажем, че тя е аналитична в някоя област, е достатъчно да покажем сходимост на степенния ред в тази област. Използваме това, че за произволно събитие A имаме $0 \leq P(A) \leq 1$, и оценяваме редът отгоре

$$\begin{aligned} |\text{PGF}_\xi(z)| &= |E(z^\xi)| = \\ &= \left| \sum_{k=0}^{\infty} z^k P(\xi = k) \right| \leq \\ &\leq \sum_{k=0}^{\infty} |z^k P(\xi = k)| = \\ &= \sum_{k=0}^{\infty} |z|^k P(\xi = k) \leq \\ &\leq \sum_{k=0}^{\infty} |z|^k. \end{aligned}$$

Последният ред е сходящ при $|z| < 1$.

2. Разглеждаме степенния ред на PGF_ξ :

$$\begin{aligned} \text{PGF}_\xi(z) &= \sum_{k=0}^{\infty} z^k P(\xi = k) = \\ &= \sum_{k=0}^{m-1} z^k P(\xi = k) + z^m P(\xi = m) + \sum_{k=m+1}^{\infty} z^k P(\xi = k). \end{aligned}$$

Аналитичността на PGF_ξ ни позволява да диференцираме очакването почленно. След m -кратно диференциране получаваме

$$\begin{aligned} \text{PGF}_\xi^{(m)}(z) &= \\ &= \sum_{k=0}^{m-1} k! \cdot 0 \text{P}(\xi = k) + m! \text{P}(\xi = m) + \sum_{k=m+1}^{\infty} (k-m) \cdots (k-1) k z^{k-m} \text{P}(\xi = k). \end{aligned}$$

Следователно $\text{PGF}_\xi^{(m)}(0) = m! \text{P}(\xi = m)$.

3. Тъй като ξ и η са независими, за произволно $z \in \mathbb{R}$ величините z^ξ и z^η са независими и

$$\text{PGF}_{\xi+\eta}(z) = \text{E}(z^{\xi+\eta}) = \text{E}(z^\xi z^\eta) = \text{E}(z^\xi) \text{E}(z^\eta) = \text{PGF}_\xi(z) \text{PGF}_\eta(z).$$

4. Директно следва от изразяването на стойностите на ξ и η чрез производните на пораждащите ги функции. □

Теорема 16 (Свойства на пораждащите моментите функции). *Нека ξ и η са независими дискретни случайни величини.*

За пораждащите моментите функции MGF_ξ и MGF_η са изпълнени следните свойства

1. *В общия случай пораждащата моментите функция съществува само в 0. Ако тя съществува в околност на 0, то тя е гладка в тази околност, съществуват всички моменти и е изпълнено $\text{E}(\xi^n) = \text{MGF}_\xi^{(n)}(0)$ за $n = 1, 2, \dots$*
2. *Ако пораждащите моментите функции на ξ , η и $\xi + \eta$ съществуват в точка $t \in \mathbb{R}$, имаме*

$$\text{MGF}_{\xi+\eta}(t) = \text{MGF}_\xi(t) \text{MGF}_\eta(t).$$

3. *Ако MGF_ξ и MGF_η имат обща дефиниционна област, различна от $\{0\}$, в която те съвпадат, то ξ и η също съвпадат.*

Доказателство. Нека стойностите на ξ са x_1, x_2, \dots

1. Пораждащата моментите функция винаги съществува в 0, тъй като $\text{MGF}_\xi(0) = \text{E}(e^{0\xi}) = \text{E}(1) = 1$.

Нека MGF_ξ съществува в околност U на 0. Без ограничение на общността ще считаме, че U е ограничена. Полагаме $\tau := \min(-\inf U, \sup U)$. Тогава сумата $\text{MGF}_\xi(-\tau) + \text{MGF}_\xi(\tau)$ е крайна. Развиваме тази сума в ред на Тейлър:

$$\text{MGF}_\xi(-\tau) + \text{MGF}_\xi(\tau) = \text{E}(e^{-\tau\xi} + e^{\tau\xi}) = \text{E}\left(2 \sum_{k=0}^{\infty} \frac{\xi^{2k}}{(2k)!} \tau^{2k}\right) = 2 \sum_{k=0}^{\infty} \frac{\text{E}(\xi^{2k})}{(2k)!} \tau^{2k}.$$

Внасянето на очакването е възможно, защото всички членове на реда са неотрицателни. От $E(\xi^{2m}) = E(|\xi^{2m}|)$ се вижда, че всички четни моменти съществуват. Според твърдение 11 съществуват и всички нечетни моменти. При фиксирано $n = 1, 2, \dots$ за някаква околност на 0, съдържаща се в U , за някакво $\alpha \in [0, 1]$ теоремата на Тейлър, приложена към експоненциалната функция, ни дава

$$\begin{aligned} \text{MGF}_\xi(t) &= E(e^{t\xi}) = \\ &= E\left(\sum_{k=0}^n \frac{\xi^k}{k!} t^k + \frac{(\alpha\xi)^{(n+1)} e^{\alpha t\xi}}{(n+1)!} t^{n+1}\right) = \\ &= \sum_{k=0}^n \frac{E(\xi^k)}{k!} t^k + \frac{\alpha^{(n+1)} E(h(t))}{(n+1)!} t^{n+1}, \end{aligned}$$

където $h(t) := \xi^{(n+1)} e^{\alpha t\xi} \xrightarrow{t \rightarrow 0} 0$ поточково.

Получихме полином на Тейлър за функцията $\text{MGF}_\xi(t)$. Тогава $\text{MGF}_\xi(t)$ има n -та производна, при това $\text{MGF}_\xi^{(n)}(0) = E(\xi^n)$.

Следователно $\text{MGF}_\xi(t)$ има n -та производна, при това

$$\begin{aligned} \text{MGF}_\xi^{(n)}(t) &= \sum_{k=0}^{n-1} \frac{E(\xi^k)}{k!} k! \cdot 0 + \frac{E(\xi^n)}{n!} n! + \frac{\alpha^{(n+1)} E(\xi^{(n+1)} e^{\alpha t\xi})}{(n+1)!} (n+1)! t = \\ &= E(\xi^n) + \alpha^{(n+1)} E(\xi^{(n+1)} e^{\alpha t\xi}) t. \end{aligned}$$

В частност, $\text{MGF}_\xi^{(n)}(0) = E(\xi^n)$.

2. Ако пораждащите моментите функции съществуват в $t \in \mathbb{R}$, тъй като ξ и η са независими, случайните величини $e^{t\xi}$ и $e^{t\eta}$ също са независими и

$$\text{MGF}_{\xi+\eta}(t) = E(e^{t(\xi+\eta)}) = E(e^{t\xi} e^{t\eta}) = E(e^{t\xi}) E(e^{t\eta}) = \text{MGF}_\xi(t) \text{MGF}_\eta(t).$$

3. Нека z_1, z_2, \dots е обединение на стойностите на ξ и η . Ако функциите MGF_ξ и MGF_η съвпадат в областта си на дефиниция U , за $t \in U$ имаме

$$\begin{aligned} \text{MGF}_\xi(t) - \text{MGF}_\eta(t) &= 0 \\ \sum_k e^{tz_k} P(\xi = z_k) - \sum_k e^{tz_k} P(\eta = z_k) &= 0 \\ \sum_k e^{tz_k} (P(\xi = z_k) - P(\eta = z_k)) &= 0. \end{aligned}$$

Последното равенство е изпълнено за всяко $t \in U$ точно когато $P(\xi = z) = P(\eta = z)$ за всяко $z \in \text{img}(\xi) \cup \text{img}(\eta)$. Следователно ξ и η приемат едни и същи стойности $\text{img}(\xi) = \text{img}(\eta)$ с една и съща вероятност и тъй като и двете са дискретни, те съвпадат.

□

Теорема 17 (Свойства на характеристичните функции). Нека ξ и η са независими дискретни случайни величини.

За характеристичните функции φ_ξ и φ_η са изпълнени следните свойства

1. φ_ξ съществува и е равномерно непрекъсната навсякъде върху реалната права.
2. Ако ξ^n има краен n -ти момент, е изпълнено $E(\xi^n) = i^{-n} \varphi_\xi^{(n)}(0)$.

3. За всяко $t \in \mathbb{R}$ имаме

$$\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t).$$

4. Ако φ_ξ и φ_η съвпадат, то ξ и η също съвпадат.

Доказателство. Нека стойностите на ξ са x_1, x_2, \dots

1. За да докажем, че φ_ξ е дефинирана навсякъде в \mathbb{R} , оценяваме отгоре абсолютната стойност на φ_ξ за $t \in \mathbb{R}$:

$$\begin{aligned} |\varphi_\xi(t)| &= |E(e^{it\xi})| = \\ &= \left| \sum_k e^{itx_k} P(\xi = x_k) \right| \leq \\ &\leq \sum_k |e^{itx_k}| P(\xi = x_k) = \\ &= \sum_k P(\xi = x_k) = 1. \end{aligned}$$

За да докажем и равномерната непрекъснатост в \mathbb{R} , първо оценяваме отгоре израза

$$\begin{aligned} |\varphi_\xi(t+h) - \varphi_\xi(t)| &= |E(e^{i(t+h)\xi}) - E(e^{it\xi})| = \\ &= |E(e^{it\xi}(e^{ih\xi} - 1))| = \\ &= \left| \sum_k e^{itx_k}(e^{ihx_k} - 1) P(\xi = x_k) \right| \leq \\ &\leq \sum_k |e^{itx_k}| |e^{ihx_k} - 1| P(\xi = x_k) = \\ &= \sum_k |e^{ihx_k} - 1| P(\xi = x_k). \end{aligned}$$

Ще използваме, че за всяко $z \in \mathbb{C}$ неравенството на Йенсен ни дава

$$|e^{iz} - 1| = \left| i \int_0^z e^{it} dt \right| \leq \int_0^{|z|} |e^{it}| dt = \int_0^{|z|} dt = |z|.$$

Фиксираме $\varepsilon > 0$. Избираме реална константа $c_\varepsilon > 0$, такава че $P(|\xi| > c_\varepsilon) < \frac{\varepsilon}{3}$. Въвеждаме две множества от индекси: $A := \{k = 1, 2, \dots \mid |x_k| \leq c_\varepsilon\}$ и $B = \mathbb{Z}^{>0} \setminus A$.

За $k \in A$ имаме

$$\sum_{k \in A} |e^{ihx_k} - 1| P(\xi = x_k) \leq \sum_{k \in A} |hx_k| P(\xi = x_k) \leq c_\varepsilon |h| \sum_{k \in A} P(\xi = x_k) \leq c_\varepsilon |h|.$$

За $k \in B$ имаме

$$\sum_{k \in B} |e^{ihx_k} - 1| P(\xi = x_k) \leq \sum_{k \in B} (|e^{ihx_k}| + 1) P(\xi = x_k) = 2 \sum_{k \in B} P(\xi = x_k) < \frac{2\varepsilon}{3}.$$

За целия ред тогава получаваме

$$\sum_k |e^{ihx_k} - 1| P(\xi = x_k) < c_\varepsilon |h| + 2\varepsilon.$$

Полагаме $\delta = \frac{\varepsilon}{3c_\varepsilon}$.

Тогава за $|h| < \delta$ имаме

$$|\varphi_\xi(t+h) - \varphi_\xi(t)| < c_\varepsilon |h| + \frac{2\varepsilon}{3} < \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon.$$

Числото δ зависи само от ε , следователно $\varphi_\xi(t)$ е равномерно непрекъснатата върху цялата реална права.

2. Нека съществува моментът $E\xi^n$. Тогава съществуват и моментите от по-нисък ред и в някаква околност на 0 за някакво $\alpha \in [0, 1]$ теоремата на Тейлър, приложена към експоненциалната функция, ни дава

$$\begin{aligned} \varphi_\xi(t) &= E(e^{it\xi}) = \\ &= E\left(\sum_{k=0}^n \frac{i^k \xi^k}{k!} t^k + \frac{(\alpha i \xi)^{(n+1)} e^{\alpha i t \xi}}{(n+1)!} t^{n+1}\right) = \\ &= \sum_{k=0}^n \frac{i^k E(\xi^k)}{k!} t^k + \frac{(\alpha i)^{(n+1)} E(h(t))}{(n+1)!} t^{n+1}, \end{aligned}$$

където $h(t) := \xi^{(n+1)} e^{\alpha i t \xi} \xrightarrow{t \rightarrow 0} 0$ поточно.

Получихме полином на Тейлър за функцията $\varphi_\xi(t)$. Тогава $\varphi_\xi(t)$ има n -та производна, при това $\varphi_\xi^{(n)}(0) = i^n E(\xi^n)$.

3. Нека z_1, z_2, \dots е обединение на стойностите на ξ и η . Ако функциите φ_ξ и φ_η съвпадат, имаме

$$\begin{aligned} \varphi_\xi(t) - \varphi_\eta(t) &= 0 \\ \sum_k e^{itz_k} P(\xi = z_k) - \sum_k e^{itz_k} P(\eta = z_k) &= 0 \\ \sum_k e^{itz_k} (P(\xi = z_k) - P(\eta = z_k)) &= 0. \end{aligned}$$

Последното равенство е изпълнено за всяко $t \in \mathbb{R}$ точно когато $P(\xi = z) = P(\eta = z)$ за всяко $z \in \text{img}(\xi) \cup \text{img}(\eta)$. Следователно ξ и η приемат едни и същи стойности $\text{img}(\xi) = \text{img}(\eta)$ с една и съща вероятност и тъй като и двете са дискретни, те съвпадат.

□

1.5. Често срещани дискретни разпределения

1.5.1. Дискретно равномерно разпределение

Определение 18. Казваме, че случайната величина ξ има **дискретно равномерно разпределение** над множеството N с размер n и пишем $\xi \in \mathbf{Uniform}(N)$, ако за всяко $k \in N$ е изпълнено

$$P(\xi = k) = \frac{1}{n}.$$

Дефиницията е коректна, тъй като $\frac{1}{n} > 0$ и $\sum_{k \in N} P(\xi = k) = \sum_{k \in N} \frac{1}{n} = \frac{n}{n} = 1$.

За простота обикновено считаме, че N е множеството на целите числа между две положителни цели числа a и b , т.е. $N = \{a, a + 1, \dots, b - 1, b\}$. В такъв случай бележим $\xi \in \mathbf{Uniform}(a, b)$.

Равномерното дискретно разпределение моделира експерименти с краен брой изходи, за които предполагаме, че са равновероятни. Такива са например разни игри - боята на случайно избрана игрална карта или броят точки, паднали се при хвърляне на зар. Друг пример идва от втората световна война, където статистици са се опитали да оценят броя n произведени от Германия танкове, предполагайки, че серийните им номера са равномерно разпределени между 1 и n .

За пораждащата функция, очакването и дисперсията на $\xi \in \mathbf{Uniform}(a, b)$ имаме

$$\begin{aligned} \text{PGF}_\xi(z) &= E(z^\xi) = \\ &= \frac{1}{b - a + 1} \sum_{k=a}^b z^k = \\ &= \frac{1}{b - a + 1} \left(\sum_{k=1}^b z^k - \sum_{k=1}^{a-1} z^k \right) = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{b-a+1} \left(\frac{z^{b+1}-1}{z-1} - \frac{z^a-1}{z-1} \right) = \\
&= \boxed{\frac{z^{b+1}-z^a}{(b-a+1)(z-1)}},
\end{aligned}$$

$$\begin{aligned}
E(\xi) &= \frac{1}{b-a+1} \sum_{k=a}^b k = \\
&= \frac{1}{b-a+1} \left(\sum_{k=1}^b k - \sum_{k=1}^{a-1} k \right) = \\
&= \frac{1}{b-a+1} \left(\frac{b(b+1)}{2} - \frac{(a-1)a}{2} \right) = \\
&= \frac{b^2+b-a^2+a}{2(b-a+1)} = \\
&= \frac{(b-a)(b+a) + (b+a)}{2(b-a+1)} = \\
&= \boxed{\frac{a+b}{2}},
\end{aligned}$$

$$\begin{aligned}
E(\xi^2) &= \frac{1}{b-a+1} \sum_{k=a}^b k^2 = \\
&= \frac{1}{b-a+1} \left(\sum_{k=1}^b k^2 - \sum_{k=1}^{a-1} k^2 \right) = \\
&= \frac{1}{b-a+1} \left(\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)a(2a-1)}{6} \right) = \\
&= \frac{b(2b^2+3b+1) - b(2a^2-3a+1)}{6(b-a+1)} = \frac{2(b^3-a^3) + 3(b^2+a^2) + (b-a)}{6(b-a+1)} = \\
&= \frac{2[(b-a)^3 + 3b^2a - 3ba^2] + 3[(b-a)^2 + 2ba] + (b-a)}{6(b-a+1)} = \\
&= \frac{(b-a)[2(b-a)^2 + 3(b-a) + 1] + 6ba(b-a+1)}{6(b-a+1)} = \\
&= \frac{(b-a)[2(b^2-2ba+a^2+2b-2a+1) - (b-a+1)] + 6ba(b-a+1)}{6(b-a+1)} = \\
&= \frac{(b-a)[2(b-a+1)^2 - (b-a+1)] + 6ba(b-a+1)}{6(b-a+1)} = \\
&= \frac{(b-a)[2(b-a+1) - 1] + 6ba}{6} = \\
&= \boxed{\frac{2b^2+2ba+2a^2+b-a}{6}},
\end{aligned}$$

$$\begin{aligned}
\text{var}(\xi) &= E(\xi^2) - E(\xi)^2 = \\
&= \frac{2b^2 + 2ba + 2a^2 + b - a}{6} - \left(\frac{b+a}{2}\right)^2 = \\
&= \frac{2(2b^2 + 2ba + 2a^2 + b - a) - 3b^2 + 2ba + a^2}{12} = \\
&= \frac{b^2 - 2ba + a^2 + 2b - 2a}{12} = \\
&= \boxed{\frac{(b-a+1)^2 - 1}{12}}.
\end{aligned}$$

1.5.2. Биномно разпределение

Определение 19. Казваме, че случайната величина ξ има **биномно разпределение** с обем $n = 1, 2, \dots$ и вероятност $p \in (0, 1)$ и пишем $\xi \in \mathbf{Binom}(n, p)$, ако за всяко $k = 0, 1, \dots, n$ е изпълнено

$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Дефиницията е коректна, тъй като $P(\xi = k) > 0$ за $k = 0, \dots, n$ и от биномната теорема имаме

$$\sum_{k=0}^n P(\xi = k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

Частният случай $\mathbf{Binom}(1, p)$ наричаме **разпределение на Бернули** и бележим с $\mathbf{Bernoulli}(p)$.

Биномното разпределение моделира брой успехи сред общо n независими опита с еднаква вероятност p . Тога го прави изключително приложимо - от моделиране на количеството ръце с поне един туз при n раздавания на игрални карти до моделиране на броя освободени невротрансмитери в химичните синапси на мозъка за единица време.

За пораждащата моментите функция, очакването и дисперсията на случайната величина $\xi \in \mathbf{Binom}(n, p)$ имаме

$$\begin{aligned}
\text{MGF}_\xi(t) &= E(e^{t\xi}) = \\
&= \sum_{k=0}^n \binom{n}{k} e^{tk} p^k (1-p)^{n-k} = \\
&= \sum_{k=0}^n \binom{n}{k} (e^t p)^k (1-p)^{n-k} = \\
&= \boxed{(e^t p + (1-p))^n},
\end{aligned}$$

$$\begin{aligned}
\text{MGF}'_{\xi}(t) &= \boxed{np e^t (e^t p + (1-p))^{n-1}}, \\
\text{MGF}''_{\xi}(t) &= np \left[e^t \cdot (e^t p + (1-p))^{n-1} + e^t \cdot (n-1)p (e^t p + (1-p))^{n-2} \right] = \\
&= np e^t (e^t p + (1-p))^{n-2} (e^t p + (1-p) + (n-1)p) = \\
&= \boxed{np e^t (e^t p + (1-p))^{n-2} (np + e^t p - 2p + 1)}, \\
E(\xi) &= \text{MGF}'_{\xi}(0) = \boxed{np}, \\
E(\xi^2) &= \text{MGF}''_{\xi}(0) = \boxed{np(np - p + 1)} \\
\text{var}(\xi) &= E(\xi^2) - E(\xi)^2 = np(np - p + 1) - (np)^2 = \boxed{np(1-p)}.
\end{aligned}$$

Както става ясно от вида на пораждащата моментите функция, ако

$$\xi_1, \dots, \xi_n \in \mathbf{Bernoulli}(p),$$

то сумата им има разпределение $\mathbf{Binom}(n, p)$. Нещо повече, сумата на случайни величини с разпределения $\mathbf{Binom}(n, p)$ и $\mathbf{Binom}(m, p)$ има разпределение $\mathbf{Binom}(n + m, p)$.

Биномното разпределение е изиграло важна историческа роля, тъй като за него първо са формулирани известни гранични теореми.

Теорема 20 (Бернули, предшественик на законите за големите числа). *Нека са дадени независими и еднакво $\mathbf{Binom}(n, p)$ -разпределени случайни величини ξ_1, ξ_2, \dots . Тогава*

$$\frac{1}{n} \sum_{k=1}^n \xi_k \rightarrow I_p$$

по вероятност.

Теорема 21 (Моавър-Лаплас, предшественик на централните гранични теореми). *Ако $\xi \in \mathbf{Binom}(n, p)$, тогава*

$$\frac{\xi - E(\xi)}{\sigma_{\xi}} = \frac{\xi - np}{\sqrt{np(1-p)}}$$

клони по вероятност към стандартно нормално разпределение при $n \rightarrow \infty$.

Вж. също теорема 25.

1.5.3. Геометрично разпределение

Определение 22. Казваме, че случайната величина ξ има **геометрични разпределение** с вероятност $p \in (0, 1)$ и пишем $\xi \in \mathbf{Geom}(p)$, ако за всяко $k = 1, 2, \dots$ е изпълнено

$$P(\xi = k) = (1-p)^{k-1} p.$$

Дефиницията е коректна, тъй като $P(\xi = k) > 0$ за $k = 1, 2, \dots$ и

$$\sum_{k=1}^{\infty} P(\xi = k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=0}^{\infty} (1-p)^k = p \frac{1}{1-(1-p)} = 1.$$

Понякога се ползва алтернативна дефиниция, при която $\eta = \xi - 1$ и за $k = 0, 1, \dots$

$$P(\eta = k) = (1-p)^k p.$$

Геометричното разпределение моделира номера на първия успешен опит сред безброй независими опити с еднаква вероятност p или, при алтернативната дефиниция, броя неуспехи до появата на първия успех. То също е изключително приложимо, например за намирането на очаквания брой изтеглени игрални карти преди първия туз или за оценяване на необходимото количество допълнително произведени изделия, ако целим да произведем дадено количество изделия без дефект при вероятност за дефект p . Предимство на геометричното разпределение за тези модели е свойството липса на памет, описано в теорема 23.

За пораждащата моментите функция, очакването и дисперсията на $\xi \in \mathbf{Geom}(p)$ имаме

$$\begin{aligned} \text{MGF}_{\xi}(t) &= E(e^{t\xi}) = e^t p \sum_{k=1}^{\infty} ((1-p)e^t)^{k-1} = e^t p \sum_{k=0}^{\infty} ((1-p)e^t)^k = \\ &= \boxed{\frac{pe^t}{1-(1-p)e^t}, (1-p)e^t < 1 \iff t < -\ln(1-p)}, \\ \text{MGF}'_{\xi}(t) &= \frac{pe^t \cdot (1-(1-p)e^t) - pe^t \cdot (-(1-p))e^t}{(1-(1-p)e^t)^2} = \\ &= \boxed{\frac{pe^t}{(1-(1-p)e^t)^2}, t < -\ln(1-p)}, \\ \text{MGF}''_{\xi}(t) &= \frac{pe^t \cdot (1-(1-p)e^t)^2 - pe^t \cdot 2(1-(1-p)e^t)(-(1-p))e^t}{(1-(1-p)e^t)^4} = \\ &= \frac{pe^t(1-(1-p)e^t)(1-(1-p)e^t + 2(1-p)e^t)}{(1-(1-p)e^t)^4} = \\ &= \boxed{\frac{pe^t(1+e^t-pe^t)}{(1-(1-p)e^t)^3}, t < -\ln(1-p)} \\ E(\xi) &= \frac{p}{p^2} = \boxed{\frac{1}{p}}, \\ E(\xi^2) &= \frac{p(2-p)}{p^3} = \boxed{\frac{2-p}{p^2}}, \end{aligned}$$

$$\text{var}(\xi) = E(\xi^2) - E(\xi)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \boxed{\frac{1-p}{p^2}}.$$

Теорема 23 (Липса на памет). За $\xi \in \mathbf{Geom}(p)$ и положителни цели числа n и m е изпълнено

$$P(\xi = n + m \mid \xi > n) = P(\xi = m)$$

Доказателство.

$$\begin{aligned} P(\xi = n + m \mid \xi > n) &= \frac{P(\xi = n + m, \xi > n)}{P(\xi > n)} = \\ &= \frac{P(\xi = n + m)}{P(\xi > n)} = \\ &= \frac{(1-p)^{n+m-1} p}{\sum_{k=n+1}^{\infty} (1-p)^{k-1} p} = \\ &= \frac{(1-p)^{n+m-1} p}{(1-p)^n p \sum_{k=0}^{\infty} (1-p)^k} = \\ &= \frac{(1-p)^{m-1}}{\frac{1}{1-p}} = \\ &= (1-p)^{m-1} p = \\ &= P(\xi = m). \end{aligned}$$

□

Вж. също задача 1.

1.5.4. Поасоново разпределение

Определение 24. Казваме, че случайната величина ξ има **поасоново разпределение** със степен $\lambda > 0$ и пишем $\xi \in \mathbf{Poisson}(p)$, ако за всяко $k = 0, 1, \dots$ е изпълнено

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

Дефиницията е коректна, тъй като $P(\xi = k) > 0$ за $k = 0, 1, \dots$ имаме

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

Ако дадено явление се случва със средна честота λ , поасоновото разпределение моделира количеството независими явления за единица време. Това го прави подходящо за широк спектър от задачи - от моделиране на броя посетители на

магазин в натоварен час до моделиране на броя скокове в цените на финансов дериват.

За пораждащата моментите функция, очакването и дисперсията на случайната величина $\xi \in \text{Poisson}(p)$ имаме

$$\text{MGF}_{\xi}(t) = E(e^{t\xi}) = \sum_{k=0}^{\infty} \frac{e^{tk} e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!} = e^{-\lambda} e^{e^t \lambda} = \boxed{e^{\lambda(e^t - 1)}},$$

$$\text{MGF}'_{\xi}(t) = \lambda e^t e^{\lambda(e^t - 1)} = \boxed{\lambda e^{t + \lambda(e^t - 1)}},$$

$$\text{MGF}''_{\xi}(t) = \lambda e^t \cdot e^{\lambda(e^t - 1)} + \lambda e^t \cdot \lambda e^t e^{\lambda(e^t - 1)} = \boxed{\lambda e^{t + \lambda(e^t - 1)}(1 + \lambda e^t)},$$

$$E(\xi) = \boxed{\lambda},$$

$$E(\xi^2) = \boxed{\lambda(1 + \lambda)},$$

$$\text{var}(\xi) = E(\xi^2) - E(\xi)^2 = \lambda(1 + \lambda) - \lambda^2 = \boxed{\lambda}.$$

Теорема 25 (Поасон). Нека ξ_1, ξ_2, \dots са независими и $\xi_k \in \text{Binom}(k, p_k), k = 1, 2, \dots$. Ако $k p_k \xrightarrow[k \rightarrow \infty]{} \lambda$, тогава граница по вероятност $\lim_{k \rightarrow \infty} \xi_k$ съществува и има разпределение **Poisson**(λ).

Както става ясно от видът на пораждащата моментите функция, ако

$$\xi_m \in \text{Poisson}(\lambda_m), m = 1, \dots, n,$$

то сумата им има разпределение **Poisson**($\sum_{m=1}^n \lambda_m$).

Вж. също задача 2.

2. Задачи

В конспекта не е посочен списък с възможни задачи, затова съм включил разни задачи, дадени на държавен изпит.

Задача 1 (*Задачи за ДИ за спец. статистика*). Дадена е редица от независими опити с два изхода: успех и неуспех. Вероятността за успех е $p \in (0, 1)$. Опитите се провеждат до първа поява на успех. Случайната величина X е равна на броя на неуспехите. Да се намери:

- Разпределението на случайната величина X
- Пораждащата функция на X
- Математическото очакване и дисперсията на X
- За $k \geq 0$ да се намери $P(k \geq 0)$

Решение.

- а) Това е именно алтернативната формулировка на геометричното разпределение, която споменахме в параграф 1.5.3, т.е. $P(X = k) = (1 - p)^k p$.

Ако вече разполагаме с информацията за геометричното разпределение от параграф 1.5.3 и ξ има разпределение **Geom**(p) спрямо оригиналната дефиниция, тогава $X = \xi - 1$ и директно получаваме

$$\begin{aligned} \text{PGF}_X(k) &= \text{PGF}_\xi(k - 1), \\ E_X(k) &= E_\xi(k) - 1 = \frac{1}{p} - 1 = \frac{1 - p}{p}, \\ \text{var}_X(k) &= \text{var}_\xi(k) = \frac{1 - p}{p^2}. \end{aligned}$$

Можем да изведем търсените характеристики и директно

- б) Пресмятаме пораждащата функция:

$$\begin{aligned} \text{PGF}_X(z) &= \sum_{k=0}^{\infty} ((1 - p)z)^k p = \\ &= p \sum_{k=0}^{\infty} ((1 - p)z)^k = \\ &= \boxed{\frac{p}{1 - (1 - p)z}, |1 - p|z < 1 \iff |z| < \frac{1}{1 - p}}, \end{aligned}$$

при което $\text{MGF}_X(t) = \text{PGF}_X(e^t) = \boxed{\frac{p}{1 - (1 - p)e^t}, t < -\ln(1 - p)}$.

- в) Пресмятаме очакването и дисперсията, използвайки намерената пораждаща моментите функция:

$$\begin{aligned} \text{MGF}'_X(t) &= \boxed{\frac{p(1 - p)e^t}{(1 - (1 - p)e^t)^2}, t < -\ln(1 - p)} \\ \text{MGF}''_X(t) &= \frac{p(1 - p)e^t \cdot (1 - (1 - p)e^t)^2 - p(1 - p)e^t \cdot 2(1 - (1 - p)e^t)(p - 1)e^t}{(1 - (1 - p)e^t)^4} = \\ &= \frac{p(1 - p)e^t(1 - (1 - p)e^t + 2(1 - p)e^t)}{(1 - (1 - p)e^t)^3} = \\ &= \boxed{\frac{p(1 - p)e^t(1 + (1 - p)e^t)}{(1 - (1 - p)e^t)^3}, t < -\ln(1 - p)}. \end{aligned}$$

$$E(X) = \text{MGF}'_X(0) = \frac{1 - p}{p},$$

$$E(X^2) = \text{MGF}_X''(0) = \frac{p(1-p)(2-p)}{p^3} = \frac{(1-p)(2-p)}{p^2},$$

$$\text{var}(X) = E(X^2) - E(X)^2 = \frac{(1-p)(2-p)}{p^2} - \frac{(1-p)^2}{p^2} = \boxed{\frac{1-p}{p^2}}.$$

г) За $m = 0, 1, \dots$ имаме

$$P(X \geq m) = \sum_{k=m}^{\infty} (1-p)^k p = (1-p)^m p \sum_{k=0}^{\infty} (1-p)^k = \frac{(1-p)^m p}{1-(1-p)} = \boxed{(1-p)^m}.$$

□

Задача 2 (*Задачи за ДИ за спец. статистика*). Нека X е случайна величина с разпределение на Поасон с параметър $\lambda > 0$:

$$X \in \mathbf{Poisson}(\lambda), \quad P(X = x) = e^{-\lambda} \frac{x^\lambda}{x!}, x = 0, 1, 2, \dots$$

- а) Намерете математическото очакване EX и дисперсията $\text{var} X$ на случайната величина X .
- б) Нека случайните величини X_1 и X_2 са независими и с разпределение на Поасон с параметри съответно $\lambda_1 > 0$ и $\lambda_2 > 0$:

$$X_1 \in \mathbf{Poisson}(\lambda_1), \quad X_2 \in \mathbf{Poisson}(\lambda_2).$$

Намерете разпределението на сумата им

$$S_2 = X_1 + X_2.$$

- в) Нека X_1, X_2, \dots, X_n са независими наблюдения над $\mathbf{Poisson}(\lambda)$ случайна величина X и нека сумата им е $S_n = X_1 + X_2 + \dots + X_n$.

Намерете значимостта (p -value) на статистиката S_n при проверката на основна и алтернативни хипотези:

$$\begin{cases} H_0 : \lambda = 3 \\ H_1 : \lambda = 6 \end{cases}$$

при $n = 4$ и $S_4 = 15$.

Забележка 26. За тази задача е дадена таблица за разпределението на Поасон за $x = 0, 1, \dots, 24$ и $\lambda = 3, 6, \dots, 21, 24$.

Решение.

- а) В параграф 1.5.4 сме намерили, че $E(X) = \text{var}(X) = \lambda$.
- б) Според теорема 17 $\text{MGF}_{X_1+X_2}$ определя напълно разпределението на $X_1 + X_2$. Използваме пораждащата моментите функция от параграф 1.5.4, за да намерим

$$\text{MGF}_{X_1+X_2} = \text{MGF}_{X_1} \cdot \text{MGF}_{X_2} = e^{\lambda_1(e^t-1)} \cdot e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)},$$

което е пораждаща моментите функция на **Poisson**($\lambda_1 + \lambda_2$) разпределение.

- в) От предната подточка знаем, че $S_n \in \text{Poisson}(n\lambda)$. Тогава значимостта на наблюдението $S_4 = 15$ при десен едностранен тест с нулева хипотеза $\lambda = 3$ е

$$P(S_4 \geq 15 \mid 4\lambda = 12) = 1 - P(S_4 \leq 14 \mid 4\lambda = 12) \approx 1 - 0.772 = 0.228.$$

При ниво на значимост $\alpha = 0.05$ бихме имали основание да отхвърлим нулевата хипотеза.

□

3. Литература

Боровков, Александр Александрович. *Теория вероятностей*. Рус. 3-е изд. Едиториал УРСС, 1999. ISBN: 5901006666.

Димитров, Боян и Николай Янев. *Вероятности и статистика*. 2-е изд. Софтех, 2007.

Задачи за ДИ за спец. статистика. Поправителен изпит. 2014. URL: <https://intranet.fmi.uni-sofia.bg/index.php/s/v5CMDRQNMxjY00> (дата на посещ. 20.06.2019).

Задачи за ДИ за спец. статистика. Поправителен изпит. 2016. URL: <https://intranet.fmi.uni-sofia.bg/index.php/s/ictJ7xGqzZvgX40> (дата на посещ. 20.06.2019).

Конспект за ДИ за спец. статистика. 2018. URL: <https://intranet.fmi.uni-sofia.bg/index.php/s/KOTdUnmqbrnd0sX> (дата на посещ. 24.03.2019).