Случайни величини с непрекъснати разпределения. Задачи, в които възникват. Моменти – математическо очакване и дисперсия.

П.Матеев

София, 2012

Дефиниция на непрекъснато разпределение на случайна величина. Свойства на вероятностната плътност (неотрицателност и нормираност). За всяко от разпределенията - равномерно разпределение, експоненциално разпределение, гама разпределение, бета разпределение, нормално разпределение - да се опишат моделите, водещи до съответните разпределения. Пресмятане на моментите на разпределенията до втори включително. Описват се връзките между разпределенията – експоненциално и гама, нормално и гама, гама и бета. При изчисленията може да се използват и характеристични функции, но това не е задължително.

Съдържание

1 Непрекъснато разпределение

Непрекъснато разпределение на случайна величина определяме с функцията *вероятностна плътност*. Нека си представим, че случайната величина X има възможност да приема стойности все по-"гъсто" върху реалната права.

Например нека X приема стойности $x_k = \frac{k}{2^n}, k = 1, 2, \dots, 2^n$, с равни вероятности $\frac{1}{2^n}$. Очевидно, че за големи стойности на n вероятността X да принадлежи на подинтервал $(a, b]$ на интервала $(0, 1]$ зависи само от ширината на интервала и разликата на вероятността за два интервала с равни ширини е не повече от $\frac{1}{2^n}$.

1.1 Вероятностна плътност

Удобен инструмент за работа с разпределения на случайни величини, чиито възможни стойности са гъсто разположени върху реалната права е функцията *вероятностна плътност*. Ще предложим нестроги евристични разсъждения. Нека апроксимираме сумата от вероятностите в малък интервал Δx, съдържащ x и с ширина Δ, с помощта на функция

$$f(x) = \frac{P(X \in \Delta x)}{\Delta} \quad \text{или} \quad P(X \in \Delta x) = f(x).\Delta$$

Естествени свойства на такава функция (ако съществува) са

- $f(x) \geq 0$ за всяко x от реалната права;
- $f(x)$ е интегруема функция върху реалната права $(-\infty, +\infty)$:

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

Първото следва от това, че вероятността винаги е неотрицателна и $\Delta > 0$.

Нека разбием правата на изброимо много интервали Δx (непресичащи се и покриващи правата) и с еднакви ширини Δ. От определянето на $f(x)$ следва равенството

$$\sum_{\Delta x} P(X \in \Delta x) = \sum_{\Delta x} f(x).\Delta.$$

Безкрайната сума отляво е сходяща и е равна на $P(\Omega) = 1$, а сумата в дясно е приближение на определения интеграл от $f(x)$ върху интервала $(-\infty, +\infty)$.

Разпределението на всяка случайна величина X (в зададено вероятностно пространство $(\Omega, \mathcal{A}, P)$) се определя с функцията ѝ на разпределение $F_X(x) = P(X \leq x)$.

Случайната величина X е с *непрекъснато разпределение* ако функцията ѝ на разпределение $F_X(x)$ е непрекъсната и *абсолютно непрекъсната*, ако съществува функция $f(x)$ такава, че

$$F_X(x) = \int_{-\infty}^{x} f(u)du.$$

Функцията $f(x)$ наричаме *плътност на разпределението* на X и ако диференцираме двете страни на равенството по x, я получаваме в явен вид

$$f(x) = \frac{dF(x)}{dx},$$

за онези стойности x, за които $f(x)$ е непрекъсната и производната на $F(x)$ (поне отдясно) съществува.

Всяка функция $f(x)$, която удовлетворява естествените условия по-горе (неотрицателна и интегруема върху правата с интеграл равен на 1) определя функция

$$F(x) = \int_{-\infty}^{x} f(u)du,$$

която може да служи за функция на разпределение на някоя случайна величина.

Така определеното разпределение може да служи за приближение на разпределението на дискретна случайна величина.

Плътността определя еднозначно функцията на разпределение, но една функция на разпределение може да се определя от различни плътности $f(x)$ и $g(x)$ стига стойностите на аргумента, за които двете функции различават да не са много (няколко, може и изброимо много, но не цял интервал).

1.2 Независимост на непрекъснати разпределения

Две случайни величини (X, Y) със съвместна функция на разпределение $F_{XY}(x, y) = P(\{X \leq x\} \cap \{Y \leq y\})$ са със съвместно непрекъснато разпределение, ако съществува съвместната плътност $f(u, v)$ и е в сила представянето

$$F_{XY}(x, y) = \int_{-\infty}^{x} \left[\int_{-\infty}^{x} f(u, v)dv \right] du.$$

Двете маргинални разпределения също са непрекъснати и плътностите им се получават с интегриране върху правата:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, v)dv \quad \text{и} \quad f_Y(y) = \int_{-\infty}^{+\infty} f(u, y)du$$

Условието за независимост на две случайни величини (X, Y) изразено чрез съвместнота и маргиналните функции на разпределение

$$F_{X,Y}(x, y) = F_X(x).F_Y(y)$$

при непрекъснатите разпределения е еквивалентно на

$$f(x, v) = f_X(x).f_Y(y).$$

1.3 Моменти

Математическото очакване на случайна величина X с непрекъснато разпределение и вероятностна плътност $f(x)$ е също граница на безкрайни суми на стойностите ѝ с тегла съответни вероятности или интеграл върху реалната права

$$\mathbb{E}\,X = \int_{-\infty}^{+\infty} x.f(x)dx,$$

който е добре определен, ако интегралът е абсолютно сходящ.

Забележка: Ако е разходящ върху $[0, +\infty)$ и сходящ върху $(-\infty, 0]$ можем да приемем че очакването е $+\infty$ и обратно, очакването е $-\infty$, ако интегралът е сходящ върху $[0, +\infty)$ и разходящ върху $(-\infty, 0]$. Ако интегралите върху двата интервала са разходящи не е възможно да се определи математическото очакване.

Свързаните с математическото очакване на случайна величина X - моменти $\mathbb{E}\,X^k$, дисперсията $\mathbb{V}ar\,x = \mathbb{E}\,(X - \mathbb{E}\,X)^2$ и централни моменти от по-висок ред, пораждащата моментите функция $M(t) = \mathbb{E}\,e^{tX}$ и комплексния и аналог характеристичната функция $\varphi(t) = \mathbb{E}\,e^{-itX}$ се определят също като съответни интегали и съществуват, когато са сходящи (абсолютно!).

2 Някои непрекъснати разпределения

Ще разгледаме основните непрекъснати разпределения, като следваме схемата:

- задачи, в които възникват;
- функция на разпределение (плътност) и параметри;
- моменти – математическо очакване и дисперсия (функция пораждаща моментите).

2.1 Равномерно разпределение

Задачи и приложение. Случайна величина с равномерно разпределение е модел на наблюдения, за които е известно единствено, че са от даден интервал $[a, b]$. Стандартното равномерно разпределение е върху интервала $[0, 1]$. Означаваме с $X \sim U(0, 1)$, че случайната величина X е равномерно разпределена върху единичния интервал. Случайна величина с равномерно разпределение в интервал $[a, b]$ ще получим с очевидна линейна трансформация $Y = a + X.(b - a)$, предполагаме $a < b$, и означаваме $Y \sim U(a, b)$).

Стандартно равномерно разпределение се получава и в резултат на следната трансформация: ако случайната величина X има непрекъсната функция на разпределение $F(x)$, то $Y = F(X) \sim U(0, 1)$. Този факт позволява да се генерират редици от случайни величини с дадено разпределение (с дадена непекъснатата функция на разпределение) с помощта на генератор на случайни (или псевдо-случайни) редици от равномерно разпределени случайни величини. Генератор на псевдо-случайно равномерно разпределени величини в единичния интервал е задължителен атрибут на почти всяка среда за програмиране/библиотека.

Функция на разпределение на стандартно равномерно разпределение е диагоналът на единичния квадрат, с хоризонтално продължение извън интервала $[0,1]$:

$$F(x) = \begin{cases} 0 & x \in (-\infty, 0] \\ x & x \in [0,1] \\ 1 & x \in (1, \infty) \end{cases}$$

Плътността на равномерно разпределена в $[0,1]$ случайна величина е константа 1 в интервала $[0,1]$ и 0 извън него.

Плътността на разпределение на $Y \sim U(a,b)$ е равна на $\frac{1}{b-a}$ в интервала $[a,b]$ и 0 извън него, а функцията на разпределение е диагоналът на правоъгълника над интервала $[a,b]$, съединяващ точките $(a,0)$ и $(b,1)$:

$$F(x) = \begin{cases} 0 & x \in (-\infty, a] \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x \in (b, \infty) \end{cases}$$

Моменти. Математическото очакване на случайната величина X със стандартно равномерно разпределение получаваме директно

$$\mathbb{E}\,X = \int_0^1 x dx = \left.\frac{x^2}{2}\right|_{x=0}^1 = \frac{1}{2} - 0 = \frac{1}{2}.$$

Аналогично получаваме, че вторият момент е

$$\mathbb{E}\,X^2 = \int_0^1 x^2.dx = \left.\frac{x^3}{3}\right|_{x=0}^1 = \frac{1}{3} - 0 = \frac{1}{3}.$$

и за дисперсията на X

$$\mathbb{V}ar\,X = \mathbb{E}\,X^2 - (\mathbb{E}\,X)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Случайна величина $Y \sim U(a,b)$ получихме с трансформацията $Y = a + X.(b-a)$. Математическото ѝ очакване е

$$\mathbb{E}\,Y = \mathbb{E}\,(a + X.(b-a)) = a + (b-a)\mathbb{E}\,X = a + (b-a)\frac{1}{2} = \frac{a+b}{2},$$

и

$$\mathbb{V}ar\,Y = \mathbb{V}ar\,(a + X.(b-a)) = (b-a)^2\mathbb{V}ar\,X = \frac{(b-a)^2}{12}$$

е диспрерсията.

2.2 Експоненциално разпределение

Разпределението е еднопараметрично и е върху $(0, +\infty)$. Че случайната величина X е с експоненциално разпределение с параметър $\lambda > 0$ означаваме с $X \sim \mathrm{Exp}(\lambda)$.

Задачи и приложение. Случайна величина с експоненциално (показателно) разпределение е модел за случайно време T на настъпване на очаквано (или не) събитие. Евристични съображения за експоненциалната зависимост на вероятността на очакваното събитие от времето на чакане са следствие от вероятността за *ненастъпване* на събитие в Поасоновото разпределение с параметър λ. Вероятността за ненастъпване или точно 0 събития за единица време е $P(X=0)=e^{-\lambda}$. Ако означим с T момента на настъпване на очакваното събитие, то събитието $\{X=0\}$ може да се запише и като $\{T>1\}$. При t последователни и независими опита (при същото средно λ) ненастъпване на събитието може да се запише като $\{T>t\}$ и вероятността му е:

$$P(\{T>t\})=\left(e^{-\lambda}\right)^t=e^{-\lambda t}.$$

Същественото допускане тук е в независимостта на събитията "ненастъпване" в съседни непресичащи се интервали, което се изразява с равенството:

$$P(\{T>t+s\})=P(\{T>t\}).P(\{T>s\}).$$

Записано във вида

$$\frac{P(\{T>t+s\})}{P(\{T>t\})}=P(\{T>t+s\}|\{T>t\})=P(\{T>s\})$$

означава, че вероятността за ненастъпване в нов интервал с дължина s при условие, че вече е изчакан интервал с дължина t е равна на безусловната вероятност за ненанъстъпване в интервала с дължина s.

Последното равенство може да се запише като функционално уравнение

$$g(t+s)=g(t).g(s),$$

с известно общо решение $g(x)=e^{ax}$. За случая $g(x)=1-F(x)$ можем да определим и стойността на параметъра $a=-\lambda$.

Функция на разпределение на експоненциално разпределение с параметър $\lambda>0$ определяме:

$$F(x)=\begin{cases} 1-e^{-\lambda.x} & x\geq 0\\ 0 & x\leq 0\end{cases}$$

Плътността на разпределението при $x>0$ получаваме директно с диференциране на $F(x)$:

$$f(x)=\frac{F(x)}{dx}=\left((1-e^{-\lambda.x}\right)'_x=0-(-\lambda.e^{-\lambda.x})=\lambda.e^{-\lambda.x}.$$

За отрицателни стойности на x плътността е нула: $f(x)=0$.

Моменти. Ще определим математическото очакване и дисперсията като използваме пораждащата моментите функция $M_X(t)=\mathbb{E}\,e^{tX}$. За случайна величина с експоненциално разпределение $X\sim \mathrm{Exp}(\lambda)$ имаме

$$M_X(t)=\mathbb{E}\,e^{tX}=\int_0^\infty e^{tx}.\lambda.e^{-\lambda.x}dx=\frac{\lambda}{\lambda-t}\underbrace{\int_0^\infty(\lambda-t).e^{-(\lambda-t).x}dx}_{=1}=\frac{\lambda}{\lambda-t}.$$

Първата производна на $M_X(t)$ е

$$M'_X(t) = \left(\frac{\lambda}{\lambda - t}\right)' = \frac{\lambda}{(\lambda - t)^2}$$

и получаваме математическото очакване на X

$$\mathbb{E}\, X = M'_X(0) = \frac{\lambda}{(\lambda - 0)^2} = \frac{1}{\lambda}.$$

Аналогично от втората производна на $M_X(t)$

$$M''_X(t) = \left(\frac{\lambda}{(\lambda - t)^2}\right)' = \frac{2.\lambda}{(\lambda - t)^3}$$

получаваме

$$\mathbb{E}\,(X^2) = M'_X(0) = \frac{2.\lambda}{(\lambda - 0)^3} = \frac{2}{\lambda^2}.$$

за втория момент на X. Дисперсията получавме като използваме връзката на дисперсията с първите два момента:

$$\mathbb{V}ar\, X = \mathbb{E}\,(X^2) - (\mathbb{E}\, X)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Окончателно $\mathbb{E}\, X = \frac{1}{\lambda}$ и $\mathbb{V}ar\, X = \frac{1}{\lambda^2}$ или за експоненциалното разпределение математическото очакване и стандартното отклонение съвпадат.

2.3 Гама разпределение

Разпределението е двупараметрично и е върху $(0, +\infty)$. Означаваме $Y \sim \Gamma(k, \beta)$.

Задачи и приложение. Гама разпределението подобно на експоненциалното е модел на времето на очакване нещо да се случи. Нека k независими и еднакво разпределени случайни величини $Y_1, \ldots, Y_k$ имат експоненциално разпределение с параметър $\lambda > 0$. Тяхната сума има разпределение, което наричаме "гама" с параметри k и β като $\beta = \frac{1}{\lambda}$. Параметърът k определен така е цяло число, но формално плътността на разпределението може да приема и всяка положителна стойност на k и го означаваме с α.

Функция на разпределение. Ще потърсим плътността, тъй като функцията на разпределение на случайна величина $Y \sim \Gamma(k, \beta)$ не се изразява в явен вид. Плътността на разпределение на сумата на две независими случайни величини с експоненциално разпределение определяме като композиция на плътностите им. Нека Y и Z са независими и с разпределение $\text{Exp}(\lambda)$. Сумата им $X = Y + Z$ има плътност

$$\begin{aligned} f_X(x) &= \int_0^{+\infty} f_Y(y).f_Z(x-y)dy = \\ &= \int_0^x \lambda.e^{-\lambda.y}.\underbrace{\lambda.e^{-\lambda.(x-y)}}_{f_Z(x-y),\ y\leq x} dy + \int_x^{+\infty} \lambda.e^{-\lambda.y}.\underbrace{f_Z(x-y)}_{=0,\ y\leq x} dy = \end{aligned}$$

$$
\begin{aligned}
&= \int_0^x \lambda^2.e^{-\lambda.(y+x-y)}dy = \lambda^2.e^{-\lambda.x}\underbrace{\int_0^x dy}_{=x-0} = \\
&= \lambda^2.x.e^{-\lambda.x}.
\end{aligned}
$$

Приемаме хипотезата, която ще докажем по индукция, че плътността на разпределението на случайната величина $V \sim \Gamma(k, \lambda)$ е

$$
f_V(x) = \frac{\lambda^k}{(k-1)!}.x^{k-1}.e^{-\lambda.x}.
$$

Ще покажем, че хипотезата е изпълнена и за $k+1$ или плътността на $U = V + Y \sim \Gamma(k+1, \lambda)$ ще има същия вид, ако $Y \sim \mathrm{Exp}(\lambda)$ и не зависи от V.

$$
\begin{aligned}
f_U(x) &= \int_0^{+\infty} f_V(y).f_Y(x-y)dy = \\
&= \int_0^x \frac{\lambda^k}{(k-1)!}.y^{k-1}.e^{-\lambda.y}.\underbrace{\lambda.e^{-\lambda.(x-y)}}_{f_Y(x-y),\ y\leq x} dy = \\
&= \frac{\lambda^{k+1}}{(k-1)!}.e^{-\lambda.x}.\int_0^x y^{k-1}.dy = \\
&= \frac{\lambda^{k+1}}{(k-1)!}.e^{-\lambda.x}.\int_0^x d\frac{y^k}{k} = \\
&= \frac{\lambda^{k+1}}{k!}.x^k.e^{-\lambda.x}.
\end{aligned}
$$

Полученият вид на плътността съвпада с точност до множител с подинтегралната функция на известната гама-функция определяна с равенството

$$
\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}.e^{-x}dx.
$$

Интегралът отдясно е сходящ и функцията е определена за всяка положителна стойност на аргумента α. Непосредствено се проверява, че $\Gamma(1) = 1$ и основното равенство $\Gamma(\alpha) = (\alpha-1).\Gamma(\alpha-1)$ с интегриране по части

$$
\begin{aligned}
\Gamma(\alpha) &= \int_0^{+\infty} x^{\alpha-1}.e^{-x}dx = -\int_0^{+\infty} x^{\alpha-1}d(e^{-x}) = \\
&= \underbrace{x^{\alpha-1}d(e^{-x}|_0^{+\infty}}_{=0} + \int_0^{+\infty} e^{-x}d(x^{\alpha-1}) = \\
&= (\alpha-1).\int_0^{+\infty} x^{\alpha-2}.e^{-x}dx = (\alpha-1).\Gamma(\alpha-1).
\end{aligned}
$$

Очевидно, че за цели стойности k на аргумента $\Gamma(k) = (k-1)!$, а при $k = 1$ аргумент в полза на полагането $0! = 1$.

Окончателно определяме плътността на случайна величина $X \sim \Gamma(\alpha, \beta)$ с равенството

$$
f_X(x) = \frac{\beta^{-\alpha}}{\Gamma(\alpha)}.x^{\alpha-1}.e^{-x/\beta},
$$

като сме заместили λ с параметър β^{-1}.

Моменти. Пораждащата моментите функция на $X \sim \Gamma(\alpha, \beta)$ има прост явен вид

$$\begin{aligned}
M(t) = \mathbb{E}\, e^{tX} &= \frac{\beta^{-\alpha}}{\Gamma(\alpha)} . \int_0^{+\infty} e^{tx} x^{\alpha-1} . e^{-x/\beta}\, dx = \\
&= \frac{\beta^{-\alpha}}{\Gamma(\alpha)} . \int_0^{+\infty} x^{\alpha-1} . e^{-x.(\frac{1}{\beta}-t)}\, dx = \\
&= \frac{\beta^{-\alpha}}{\Gamma(\alpha)} . \int_0^{+\infty} x^{\alpha-1} . e^{-x./(\frac{\beta}{1-\beta.t})}\, dx = \\
&= \frac{1}{(1-\beta.t)^{\alpha}} . \underbrace{\frac{\left(\frac{\beta}{1-\beta.t}\right)^{-\alpha}}{\Gamma(\alpha)} . \int_0^{+\infty} x^{\alpha-1} . e^{-x./(\frac{\beta}{1-\beta.t})}\, dx}_{=1} = \\
&= \frac{1}{(1-\beta.t)^{\alpha}} .
\end{aligned}$$

От вида на пораждащата моментите функция на гама-разпределението следват две забележителни свойства

- ако $Y_1 \sim \Gamma(\alpha_1, \beta)$ и $Y_2 \sim \Gamma(\alpha_2, \beta)$ са независими, то
$$Y_1 + Y_2 \sim \Gamma(\alpha_1 + \alpha_2, \beta);$$
- ако $Y \sim \Gamma(1, \beta)$, то $Y \sim \mathrm{Exp}(\frac{1}{\beta})$.

Ще покажем, че $\mathbb{E}\, X = \alpha.\beta$ и $\mathbb{V}ar\, X = \alpha.\beta^2$.
Диференцираме два пъти функцията $M(t)$:

$$\begin{aligned}
M'(t) &= \left(\frac{1}{(1-\beta.t)^{\alpha}}\right)' = (-\alpha)\frac{1}{(1-\beta.t)^{\alpha+1}}(-\beta) = \alpha.\beta.\frac{1}{(1-\beta.t)^{\alpha+1}}, \\
M''(t) &= \left(\frac{1}{(1-\beta.t)^{\alpha+1}}\right)' = \alpha.\beta.\frac{1}{(1-\beta.t)^{\alpha+2}}(\alpha+1).\beta = \frac{\alpha.(\alpha+1).\beta^2}{(1-\beta.t)^{\alpha+2}}.
\end{aligned}$$

Полагаме $t = 0$ и получаваме първо $\mathbb{E}\, X = \alpha.\beta$ и за втория момент $\mathbb{E}\,(X^2) = \alpha.(\alpha+1).\beta^2$. Следователно дисперсията е
$\mathbb{V}ar\, X = \mathbb{E}\,(X^2) - (\mathbb{E}\, X)^2 = \alpha.(\alpha+1).\beta^2 - (\alpha.\beta)^2 = \alpha.\beta^2$.

2.4 Бета разпределение

Разпределението на случайна величина X с бета-разпределение се определя от два положителни параметъра (α_1, α_2). Разпределението е ограничено стандартно върху интервала $[0, 1]$. Означаваме с $X \sim Be(\alpha_1, \alpha_2)$.

Задачи и приложение. Бета разпределението е удобно за моделиране на разпределение върху краен интервал. Двата параметъра му придават значителна гъвкавост за представяне на разпределение на отношения - част от цялото. Например, ако случайните величини $Y_1 \sim \Gamma(\alpha_1, 1)$ и $Y_2 \sim \Gamma(\alpha_2, 1)$ са независими, то отношението на всяка към сумата им има бета разпределение

$$\frac{Y_1}{Y_1 + Y_2} \sim Be(\alpha_1, \alpha_2).$$

С линейна трансформация от вида $Y = (b-a).X + a$ се получава разпределение върху интервала $[a, b]$ (предполагаме $a < b$), което е удобен модел за продължителността на работи и етапи при управлението на проекти.

Функция на разпределение Плътността на разпределението на случайна величина $X \sim Be(\alpha_1, \alpha_2)$ има вида

$$f(x) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1).\Gamma(\alpha_2)} x^{\alpha_1-1}(1-x)^{\alpha_2-1}$$

за $x \in [0,1]$ и 0 извън него.

Моменти. Ще определим директно първия и втория момент.

$$\begin{aligned}
\mathbb{E}\,X &= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1).\Gamma(\alpha_2)} \int_0^1 x.x^{\alpha_1-1}(1-x)^{\alpha_2-1}\,dx = \\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1).\Gamma(\alpha_2)} . \frac{\Gamma(\alpha_1+\alpha_2+1)}{\Gamma(\alpha_1+\alpha_2+1)} . \frac{\Gamma(\alpha_1+1)}{\Gamma(\alpha_1+1)} . \int_0^1 x^{\alpha_1}(1-x)^{\alpha_2-1}\,dx = \\
&= \frac{\Gamma(\alpha_1+\alpha_2).\Gamma(\alpha_1+1)}{\Gamma(\alpha_1+\alpha_2+1).\Gamma(\alpha_1)} . \underbrace{\frac{\Gamma(\alpha_1+\alpha_2+1)}{\Gamma(\alpha_1+1).\Gamma(\alpha_2)} . \int_0^1 x^{\alpha_1}(1-x)^{\alpha_2-1}\,dx}_{=1} = \\
&= \frac{\alpha_1}{\alpha_1+\alpha_2}.
\end{aligned}$$

Аналогично

$$\begin{aligned}
\mathbb{E}\,(X^2) &= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1).\Gamma(\alpha_2)} \int_0^1 x^2.x^{\alpha_1-1}(1-x)^{\alpha_2-1}\,dx = \\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1).\Gamma(\alpha_2)} . \frac{\Gamma(\alpha_1+\alpha_2+2)}{\Gamma(\alpha_1+\alpha_2+2)} . \frac{\Gamma(\alpha_1+2)}{\Gamma(\alpha_1+2)} . \int_0^1 x^{\alpha_1+1}(1-x)^{\alpha_2-1}\,dx = \\
&= \frac{\Gamma(\alpha_1+\alpha_2).\Gamma(\alpha_1+2)}{\Gamma(\alpha_1+\alpha_2+2).\Gamma(\alpha_1)} . \underbrace{\frac{\Gamma(\alpha_1+\alpha_2+2)}{\Gamma(\alpha_1+2).\Gamma(\alpha_2)} . \int_0^1 x^{\alpha_1+1}(1-x)^{\alpha_2-1}\,dx}_{=1} = \\
&= \frac{(\alpha_1+1).\alpha_1}{(\alpha_1+\alpha_2+1).(\alpha_1+\alpha_2)}.
\end{aligned}$$

Накрая за дисперсията получаваме

$$\begin{aligned}
\mathbb{V}ar\,X &= \mathbb{E}\,(X^2) - (\mathbb{E}\,X)^2 = \\
&= \frac{(\alpha_1+1).\alpha_1}{(\alpha_1+\alpha_2+1).(\alpha_1+\alpha_2)} - \frac{\alpha_1^2}{(\alpha_1+\alpha_2)^2} = \\
&= \frac{(\alpha_1+1).\alpha_1.(\alpha_1+\alpha_2) - \alpha_1^2.(\alpha_1+\alpha_2+1)}{(\alpha_1+\alpha_2+1).(\alpha_1+\alpha_2)^2} = \\
&= \frac{\alpha_1^3+\alpha_1^2+\alpha_1^2.\alpha_2+\alpha_1.\alpha_2-\alpha_1^3-\alpha_1^2.\alpha_2-\alpha_1^2}{(\alpha_1+\alpha_2+1).(\alpha_1+\alpha_2)^2} = \\
&= \frac{\alpha_1.\alpha_2}{(\alpha_1+\alpha_2+1).(\alpha_1+\alpha_2)^2}.
\end{aligned}$$

2.5 Нормално разпределение

Разпределението на случайна величина X, когато е *нормално* се определя от два параметъра. Обикновено това са математическото очакване $\mathbb{E}\, X = \mu$ и дисперсията ѝ $\mathbb{V}ar\, X = \sigma^2$. Разпределението е върху цялата реална ос $(-\infty, +\infty)$ и симетрично спрямо очакването μ. Означаваме, че случайната величина X има нормално разпределение с $X \sim N(\mu, \sigma^2)$.

Задачи и приложение. Нормалното разпределение е най-често използвано като модел на отклонения или грешки при измервания на физическа величина. Методът на най-малките квадрати постулира, че сумата от квадратите на отклоненията (или грешките на измерване) трябва да е минимална, което е еквивалентно на твърдението, че модел на грешките са независими и еднакво нормално разпределени случайни величини.

Централната гранична теорема твърди, че сумата от независими и еднакво разпределени случайни величини се доближава до нормално разпределение с увеличаването на броя им. Твърдението е в сила и при доста по слаби условия. То е основание да предполагаме нормално разпределение при наблюдение, на което влияят множество разнопосочни и сравнително хомогенни по големина неотчетени фактори.

Функция на разпределение Плътността на стандартно нормално разпределение означаваме с ϕ има вида

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Тя е положителна за всяка стойност $x \in (-\infty, +\infty)$, симетрична спрямо началото и интегруема върху $(-\infty, +\infty)$. Стойността на нормиращата константа $\sqrt{2\pi}$ може да се определи от стойността на интеграла

$$\int_0^\infty e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}.$$

Функцията на разпределение на стандартното нормално разпределение не се изразява в явен вид

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du.$$

Моменти. Математическото очакване на стандартно разпределена случайна величина Z е $\mathbb{E}\, Z = 0$. Следва от симетрията на функцията $\phi(x)$ спрямо нулата.

Дисперсията на Z съвпада с втория момент $\mathbb{V}ar\, Z = \mathbb{E}\,(Z^2)$:

$$\begin{aligned}
\mathbb{E}\,(Z^2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 . e^{-\frac{x^2}{2}}\, dx = \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x . e^{-\frac{x^2}{2}}\, d(\frac{x^2}{2}) = \\
&= -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x .\, d(e^{-\frac{x^2}{2}}) =
\end{aligned}$$

$$= \underbrace{-\frac{1}{\sqrt{2\pi}}x.e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}}_{=0-0} + \underbrace{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx}_{=1} = 1.$$

Следователно за случайната величина Z със стандартно нормално разпределение можем да запишем $Z \sim N(0,1)$.

Линейна трансформация $X = a.Z + b$ на стандартното нормално разпределение има kима нормално разпределение с математическо очакване $\mathbb{E}\,X = \mathbb{E}\,(a.Z+b) = a.\underbrace{\mathbb{E}\,Z}_{=0} + \mathbb{E}\,b = b$.

Дисперсията ѝ е $\mathbb{V}ar\,X = \mathbb{V}ar\,(a.Z+b) = a^2.\mathbb{V}ar\,Z + \mathbb{V}ar\,b = a^2$.

Обратно, ако $X \sim N(\mu,\sigma^2)$, то $\frac{X-\mu}{\sigma}$ е с математическо очакване 0 и диспресия 1 или $\frac{X-\mu}{\sigma} \sim N(0,1)$.

Общият вид на плътността на нормално разпределена случайна величина $X \sim N(\mu,\sigma^2)$ се получава от равенството

$$F_X(x) = P(X \le x) = P(\underbrace{\frac{X-\mu}{\sigma}}_{\sim N(0,1)} \le \frac{x-\mu}{\sigma}) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Заместваме $u = \frac{v-\mu}{\sigma}$ в интеграла $\Phi(\frac{x-\mu}{\sigma})$: .

$$\frac{1}{\sqrt{2.\pi}}.\int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}}\,du = \frac{1}{\sqrt{2.\pi}}.\int_{-\infty}^{x} e^{-\frac{1}{2}.(\frac{v-\mu}{\sigma})^2}\,d\frac{v}{\sigma} = \frac{1}{\sqrt{2.\pi}.\sigma}.\int_{-\infty}^{x} e^{-\frac{(v-\mu)^2}{2.\sigma^2}}\,dv.$$

Получихми плътността на разпределение на случайната величина $X \sim N(\mu,\sigma^2)$:

$$f_X(x) = \frac{1}{\sqrt{2.\pi}.\sigma}.e^{-\frac{(x-\mu)^2}{2.\sigma^2}}.$$

Същия подход използваме за да определим разпределението на случайната величина $Y = Z^2$. Квадратът на стандартната случайна величина Z има разпределение върху $[0,+\infty)$. За положителни стойност на $x \ge 0$ имаме:

$$\begin{aligned} F_Y(x) &= P(Z^2 \le x) = P(|Z| \le \sqrt{x}) = 2.\int_0^{\sqrt{x}} e^{-\frac{u^2}{2}}\,du = \\ &= = 2.\frac{1}{\sqrt{2.\pi}}.\int_0^x e^{-\frac{v}{2}}\,\frac{dv}{2.\sqrt{v}} = \\ &= \frac{2^{-1/2}.}{\sqrt{\pi}}.\int_0^x v^{1/2-1}.e^{-\frac{v}{2}}\,dv. \end{aligned}$$

Получихме, че плътността на $Y = Z^2$ е плътност на $\Gamma(\alpha,\beta)$ разпределение за $\alpha = 1/2$ и $\beta = 2$ (и пътем показахме, че $\Gamma(1/2) = \sqrt{\pi}$).

Сумата на квадратите на m независими стандартни случайни величини $X_m = Z_1^2 + \cdots + Z_m^2$ също има разпределение $\Gamma(\alpha,\beta)$, като $\beta = 2$ се запазва и $\alpha = \underbrace{1/2 + \cdots + 1/2}_{m \text{ пъти}} = \frac{m}{2}$.

Заради изключително широкото си приложение в статистиката, този частен случай на гама-разпределение е удостоен с отделно име “хи-квадрат разпределение”, а параметърът m е “степени на свобода”.