Лекция VII - Непрекъснати съвместни разпределения.

Лекция VII - Непрекъснати съвместни разпределения.

- Съвместна плътност и функция на разпределение.
- Независимост в съвкупност
- Трансформации на случайни величини
- Разпределения използвани в статистиката

Съвместни разпределения

В предишната лекция разгледахме случайни величини, които са непрекъснати, но едномерни. Сега ще въведем съвместни разпределения на случайни величини. За да опростим записа, ще говорим предимно за двумерни разпределения. Многомерния случай се дефинира аналогично.

Навсякъде в текста (X_1, X_2) е случаен вектор, чиито компоненти са непрекъснати сл.в.

Съвместни разпределения

В предишната лекция разгледахме случайни величини, които са непрекъснати, но едномерни. Сега ще въведем съвместни разпределения на случайни величини. За да опростим записа, ще говорим предимно за двумерни разпределения. Многомерния случай се дефинира аналогично.

Навсякъде в текста (X_1, X_2) е случаен вектор, чиито компоненти са непрекъснати сл.в.

Дефиниция - Съвместна плътност

Съвместна плътност на (X_1, X_2) наричаме функцията $f_{X_1,X_2}(x_1, x_2)$, изпълняваща следните условия:

$$
\begin{aligned}
&1) \quad f_{X_1,X_2}(x_1, x_2) \geq 0 \\
&2) \quad \iint_{R^2} f_{X_1,X_2}(x_1, x_2)dx_1dx_2 = 1 \\
&3) \quad \mathrm{P}((X_1, X_2) \in D) = \iint_D f_{X_1,X_2}(x_1, x_2)dx_1dx_2
\end{aligned}
$$

Съвместни разпределения

В предишната лекция разгледахме случайни величини, които са непрекъснати, но едномерни. Сега ще въведем съвместни разпределения на случайни величини. За да опростим записа, ще говорим предимно за двумерни разпределения. Многомерния случай се дефинира аналогично.

Навсякъде в текста (X_1, X_2) е случаен вектор, чиито компоненти са непрекъснати сл.в.

Дефиниция - Съвместна плътност

Съвместна плътност на (X_1, X_2) наричаме функцията $f_{X_1,X_2}(x_1, x_2)$, изпълняваща следните условия:

$$1)\quad f_{X_1,X_2}(x_1, x_2) \geq 0$$

$$2)\quad \iint_{R^2} f_{X_1,X_2}(x_1, x_2)dx_1 dx_2 = 1$$

$$3)\quad \mathrm{P}((X_1, X_2) \in D) = \iint_{D} f_{X_1,X_2}(x_1, x_2)dx_1 dx_2$$

Съвместната плътност може да се разглежда и като непрекъснат аналог на съвместното разпределение на дискретни сл.в., което представяхме с таблица. Така 1) съответства на изискването вероятностите да са положителни, 2) сумата от всички вероятности да е едно, 3) дава връзката между случайния вектор и плътността (начина за пресмятане на вероятност).

Пример

Тества се бетонен детайл, като той е подложен на нарастваща сила, мерната единица е тон на квадратен сантиметър. Нека X е силата при която се появява първата пукнатина, а Y при която детайла се разрушава. Установено е, че $f_{X,Y}(x,y) = 24x(1-y)$ за $0 < x < y < 1$. Каква е вероятността детайлът да няма дефект при половин тон, но да се разруши при 3/4 тон?

Пример

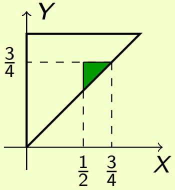
Тества се бетонен детайл, като той е подложен на нарастваща сила, мерната единица е тон на квадратен сантиметър. Нека X е силата при която се появява първата пукнатина, а Y при която детайла се разрушава. Установено е, че $f_{X,Y}(x, y) = 24x(1 - y)$ за $0 < x < y < 1$. Каква е вероятността детайлът да няма дефект при половин тон, но да се разруши при $3/4$ тон?

Вероятността която търсим е $\mathrm{P}(X > 1/2, Y < 3/4)$, съгласно условие 3) за да я пресметнем е достатъчно да интегрираме плътността върху областта, в която се изпълнява това събитие.

Пример

Теста се бетонен детайл, като той е подложен на нарастваща сила, мерната единица е тон на квадратен сантиметър. Нека X е силата при която се появява първата пукнатина, а Y при която детайла се разрушава. Установено е, че $f_{X,Y}(x,y) = 24x(1-y)$ за $0 < x < y < 1$. Каква е вероятността детайлът да няма дефект при половин тон, но да се разруши при 3/4 тон?

Вероятността която търсим е $\mathrm{P}(X > 1/2, Y < 3/4)$, съгласно условие 3) за да я пресметнем е достатъчно да интегрираме плътността върху областта, в която се изпълнява това събитие.



$$\mathrm{P}(X > 1/2, Y < 3/4) = \int_{1/2}^{3/4} \int_{x}^{3/4} 24x(1-y)dy\,dx \approx 0.1444$$

Съвместни разпределения

Подобно на дискретния случай и тук можем да намерим разпределението само на едната сл.в. Маргиналното разпределение се намира чрез интегриране по другата, аналог на сумирането в дискретния случай.

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)dx_2$$

В общия случай границите на интеграла са $(-\infty, \infty)$. В частност, в зависимост от вида на областта, е възможно границите на интегриране да са функции зависещи от x_1.

Съвместни разпределения

Подобно на дискретния случай и тук можем да намерим разпределението само на едната сл.в. Маргиналното разпределение се намира чрез интегриране по другата, аналог на сумирането в дискретния случай.

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_2$$

В общия случай границите на интеграла са $(-\infty,\infty)$. В частност, в зависимост от вида на областта, е възможно границите на интегриране да са функции зависещи от x_1.

Пример

В условията на предишния пример, нека намерим вероятността детайлът да няма дефект при половин тон, т.е. $\mathrm{P}(X > 1/2)$.

Съвместни разпределения

Подобно на дискретния случай и тук можем да намерим разпределението само на едната сл.в. Маргиналното разпределение се намира чрез интегриране по другата, аналог на сумирането в дискретния случай.

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)dx_2$$

В общия случай границите на интеграла са $(-\infty, \infty)$. В частност, в зависимост от вида на областта, е възможно границите на интегриране да са функции зависещи от x_1.

Пример

В условията на предишния пример, нека намерим вероятността детайлът да няма дефект при половин тон, т.е. $\mathrm{P}(X > 1/2)$. Това е вероятност, която зависи само от едната случайна величина, затова ще ни трябва нейното маргинално разпределение. Така за $0 < x < 1$ получаваме

$$f_X(x) = \int_x^1 24x(1-y)dy == 12x(x-1)^2$$

и следователно

$$\mathrm{P}(X > 1/2) = \int_{1/2}^1 12x(x-1)^2 dx = \ldots$$

Съвместни разпределения

Съвместна функция на разпределение

Функция на разпределение на случайния вектор (X_1, X_2), наричаме

$$F_{X_1,X_2}(x_1, x_2) = \mathrm{P}(X_1 < x_1, X_2 < x_2)$$

Съгласно 3) от дефиницията на плътност

$$F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(t_1, t_2) dt_1 dt_2$$

Следователно обратната връзка е от вида

$$f_{X_1,X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1,X_2}(x_1, x_2)}{\partial x_1 \, \partial x_2}$$

Съвместни разпределения

Съвместна функция на разпределение

Функция на разпределение на случайния вектор (X_1, X_2), наричаме

$$F_{X_1,X_2}(x_1, x_2) = \mathrm{P}(X_1 < x_1, X_2 < x_2)$$

Съгласно 3) от дефиницията на плътност

$$F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(t_1, t_2) dt_1 dt_2$$

Следователно обратната връзка е от вида

$$f_{X_1,X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1,X_2}(x_1, x_2)}{\partial x_1 \, \partial x_2}$$

Чрез функцията на разпределение обикновено се дефинира понятието "независимост", при това без значение дали сл.в. са дискретни или непрекъснати.

Дефиниция - Независимост на случайни величини.

Казваме, че сл.в. X_1 и X_2 са независими $X_1 \perp\!\!\!\perp X_2$, ако за всички $x_1, x_2 \in R$

$$F_{X_1,X_2}(x_1, x_2) = F_{X_1}(x_1) \, F_{X_2}(x_2)$$

Съвместни разпределения

В случая на непрекъснати сл.в. по-често използваме следната еквивалентна дефиниция, записана в термините на плътност.

Дефиниция - Независимост на случайни величини.

Казваме, че сл.в. X_1 и X_2 са независими $X_1 \perp\!\!\!\perp X_2$, ако за всички $x_1, x_2 \in R$

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2)$$

Подобно на независимост при случайни събития, понятието "независимост" може да бъде въведено и за повече от две сл.в.

Дефиниция - Независимост в съвкупност

Казваме, че сл.в. $X_1, X_2, \dots X_n$ са независими в съвкупност или накратко само независими, ако за всяко цяло число $2 \leq k \leq n$ и всеки набор от индекси $1 \leq i_1, i_2, \dots i_n$ е изпълнено

$$f_{X_{i_1},\dots X_{i_k}}(x_{i_1}, \dots, x_{i_k}) = f_{X_{i_1}}(x_{i_1})\, f_{X_{i_2}}(x_{i_2}) \dots f_{X_{i_k}}(x_{i_k})$$

Тази дефиниция означава, че който и случайни величини да изберем, съвместната им плътност ще бъде произведение от маргиналните.

Трансформации на случайни величини

В предишната лекция доказахме формула за смяна на променливите. Съществува аналогичен резултат в n-мерния случай. Ние ще го докажем в двумерния.

Твърдение

Нека X_1, X_2 са непрекъснати сл.в. със съвместна плътност $f_{X_1,X_2}(x_1, x_2)$. Нека

$$\begin{aligned} Y_1 &= g_1(X_1, X_2) \\ Y_2 &= g_2(X_1, X_2) \end{aligned}$$

като функциите g_1, g_2 задават взаимноеднозначно изображение на R^2 в R^2, т.е съществува обратно изображение $h_1(y_1, y_2), h_2(y_1, y_2)$ и двете изображения са непрекъснати. Нека, освен това, якобианът на смяната не е изроден.

$$J = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix} \neq 0$$

Тогава случайните величини Y_1, Y_2 имат съвместна плътност

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(\, h_1(y_1, y_2), h_2(y_1, y_2)\,)\, |J|$$

Док. Твърдението следва от теоремата за смяна на променливите в интеграл. За начало да означим

$$S = \left\{ (x_1, x_2) \in R^2 : \left| \begin{array}{l} g_1(x_1, x_2) < z_1 \\ g_2(x_1, x_2) < z_2 \end{array} \right. \right\}$$

Док. Твърдението следва от теоремата за смяна на променливите в интеграл. За начало да означим

$$S = \left\{ (x_1, x_2) \in R^2 : \left| \begin{array}{l} g_1(x_1, x_2) < z_1 \\ g_2(x_1, x_2) < z_2 \end{array} \right. \right\}$$

Ще намерим функцията на разпределение на новите случайни величини

$$F_{Y_1, Y_2}(z_1, z_2) = \mathrm{P}(Y_1 < z_1, Y_2 < z_2) = \mathrm{P}\left((x_1, x_2) \in S \right) =$$
$$= \iint_S f_{X_1, X_2}(x_1, x_2) dx_1 \, dx_2$$

Док. Твърдението следва от теоремата за смяна на променливите в интеграл. За начало да означим

$$S = \left\{ (x_1, x_2) \in R^2 : \left| \begin{array}{l} g_1(x_1, x_2) < z_1 \\ g_2(x_1, x_2) < z_2 \end{array} \right. \right\}$$

Ще намерим функцията на разпределение на новите случайни величини

$$F_{Y_1,Y_2}(z_1, z_2) = \mathrm{P}(Y_1 < z_1, Y_2 < z_2) = \mathrm{P}\,(\,(x_1, x_2) \in S\,) =$$
$$= \iint_S f_{X_1,X_2}(x_1, x_2) dx_1\, dx_2$$

Условията на теоремата ни гарантират, че в този интеграл може да се извърши смяна на променливите

$$\left| \begin{array}{l} x_1 = h_1(y_1, y_2) \\ x_2 = h_2(y_1, y_2) \end{array} \right.$$

Следователно

$$F_{Y_1,Y_2}(z_1, z_2) = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} f_{X_1,X_2,}(\,h_1(y_1, y_2), h_2(y_1, y_2)\,)\,|J|\, dy_1 dy_2$$

Като вземем предвид дефиницията на функция на разпределение, това равенство показва, че сл.в. Y_1, Y_2 имат съвместна плътност точно от търсения тип $\square$

Сума на нормално разпределени случайни величини

Това твърдение се използва за доказване на важни свойства на разпределенията и решаване на задачи, който възникват в статистиката. Ние ще докажем следното свойство на нормалното разпределение.

Сума на нормално разпределени случайни величини

Това твърдение се използва за доказване на важни свойства на разпределенията и решаване на задачи, който възникват в статистиката. Ние ще докажем следното свойство на нормалното разпределение.

Твърдение

Нека $X_1 \in N(\mu_1, \sigma_1^2)$ и $X_2 \in N(\mu_2, \sigma_2^2)$ са независими случайни величини, тогава

$$X_1 = X_2 \in N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

Сума на нормално разпределени случайни величини

Това твърдение се използва за доказване на важни свойства на разпределенията и решаване на задачи, който възникват в статистиката. Ние ще докажем следното свойство на нормалното разпределение.

Твърдение

Нека $X_1 \in N(\mu_1, \sigma_1^2)$ и $X_2 \in N(\mu_2, \sigma_2^2)$ са независими случайни величини, тогава

$$X_1 = X_2 \in N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

Док. В действителност ще докажем един частен случай на това твърдение, когато разпределението е стандартно нормално. В общия случай доказателството е същото като идея и изпълнение, но записът е претрупан с константи.

И така, нека $X_1, X_2 \in N(0,1)$, т.е. $f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_i^2}$ за $i = 1, 2$.
Ще докажем, че $X_1 + X_2 \in N(0, 2)$.

Сума на нормално разпределени случайни величини

Това твърдение се използва за доказване на важни свойства на разпределенията и решаване на задачи, който възникват в статистиката. Ние ще докажем следното свойство на нормалното разпределение.

Твърдение

Нека $X_1 \in N(\mu_1, \sigma_1^2)$ и $X_2 \in N(\mu_2, \sigma_2^2)$ са независими случайни величини, тогава

$$X_1 = X_2 \in N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

Док. В действителност ще докажем един частен случай на това твърдение, когато разпределението е стандартно нормално. В общия случай доказателството е същото като идея и изпълнение, но записът е претрупан с константи.

И така, нека $X_1, X_2 \in N(0,1)$, т.е. $f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_i^2}$ за $i = 1, 2$.
Ще докажем, че $X_1 + X_2 \in N(0, 2)$.

За да намерим съвместната плътност ще използваме фактът, че сл.в. X_1 и X_2 са независими, тогава

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2+x_2^2)}$$

Сума на нормално разпределени случайни величини

Ще направим смяна на променливите. Първата сл.в. $Y_1 = X_1 + X_2$ е тази, която ни интересува. Втората сл.в. $Y_2 = X_2$ въвеждаме изкуствено с цел да направим смяната обратима. Правата трансформация има вида

$$\left| \begin{array}{rl} y_1 = & x_1 + x_2 \\ y_2 = & x_2 \end{array} \right.$$

Сума на нормално разпределени случайни величини

Ще направим смяна на променливите. Първата сл.в. $Y_1 = X_1 + X_2$ е тази, която ни интересува. Втората сл.в. $Y_2 = X_2$ въвеждаме изкуствено с цел да направим смяната обратима. Правата трансформация има вида

$$\left| \begin{array}{rl} y_1 = & x_1 + x_2 \\ y_2 = & x_2 \end{array} \right.$$

Обратната трансформация е

$$\left| \begin{array}{rr} x_1 = & y_1 - y_2 \\ x_2 = & y_2 \end{array} \right.$$

Сума на нормално разпределени случайни величини

Ще направим смяна на променливите. Първата сл.в. $Y_1 = X_1 + X_2$ е тази, която ни интересува. Втората сл.в. $Y_2 = X_2$ въвеждаме изкуствено с цел да направим смяната обратима. Правата трансформация има вида

$$\left| \begin{array}{rl} y_1 = & x_1 + x_2 \\ y_2 = & x_2 \end{array} \right.$$

Обратната трансформация е

$$\left| \begin{array}{rr} x_1 = & y_1 - y_2 \\ x_2 = & y_2 \end{array} \right.$$

За якубианът получаваме

$$|J| = \left\| \begin{array}{rr} 1 & -1 \\ 0 & 1 \end{array} \right\| = 1$$

Сума на нормално разпределени случайни величини

Ще направим смяна на променливите. Първата сл.в. $Y_1 = X_1 + X_2$ е тази, която ни интересува. Втората сл.в. $Y_2 = X_2$ въвеждаме изкуствено с цел да направим смяната обратима. Правата трансформация има вида

$$\left| \begin{array}{rl} y_1 = & x_1 + x_2 \\ y_2 = & x_2 \end{array} \right.$$

Обратната трансформация е

$$\left| \begin{array}{rr} x_1 = & y_1 - y_2 \\ x_2 = & y_2 \end{array} \right.$$

За якубианът получаваме

$$|J| = \left\| \begin{array}{rr} 1 & -1 \\ 0 & 1 \end{array} \right\| = 1$$

Съгласно твърдението за трансформации на случайни величини, плътността на Y_1, Y_2 ще бъде

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2\pi} e^{-\frac{1}{2}((y_1 - y_2)^2 + y_2^2)} \,.\, 1$$

Сега остава да намерим маргиналната плътност на първата сл.в. Това става по стандартния начин с интегриране по другата.

Сума на нормално разпределени случайни величини

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[(y_1-y_2)^2+y_2^2\right]}\, dy_2$$

Ще решим този интеграл, като го сведем до интеграл от нормална плътност. За целта ще допълним израза в експонентата до точен квадрат по y_2

Сума на нормално разпределени случайни величини

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[(y_1-y_2)^2+y_2^2\right]}\, dy_2$$

Ще решим този интеграл, като го сведем до интеграл от нормална плътност. За целта ще допълним израза в експонентата до точен квадрат по y_2

$$(y_1 - y_2)^2 + y_2^2 = y_1^2 - 2y_1y_2 + 2y_2^2 = \left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}$$

Ще заместим изразът в интеграла и ще извадим пред него множителите, които не зависят от променливата на интегриране y_2.

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[(y_1-y_2)^2+y_2^2\right]}\, dy_2$$

Ще решим този интеграл, като го сведем до интеграл от нормална плътност. За целта ще допълним израза в експонентата до точен квадрат по y_2

$$(y_1 - y_2)^2 + y_2^2 = y_1^2 - 2y_1y_2 + 2y_2^2 = \left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}$$

Ще заместим изразът в интеграла и ще извадим пред него множителите, които не зависят от променливата на интегриране y_2.

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[\left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}\right]}\, dy_2 = \frac{e^{-\frac{y_1^2}{4}}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2}\, dy_2$$

Сума на нормално разпределени случайни величини

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[(y_1-y_2)^2+y_2^2\right]}\, dy_2$$

Ще решим този интеграл, като го сведем до интеграл от нормална плътност. За целта ще допълним израза в експонентата до точен квадрат по y_2

$$(y_1 - y_2)^2 + y_2^2 = y_1^2 - 2y_1y_2 + 2y_2^2 = \left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}$$

Ще заместим изразът в интеграла и ще извадим пред него множителите, които не зависят от променливата на интегриране y_2.

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[\left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}\right]}\, dy_2 = \frac{e^{-\frac{y_1^2}{4}}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2}\, dy_2$$

Сега ще направим смяна на променливите $t = \sqrt{2}y_2$. Ясно е, че границите на интегрирането се запазват, а $dy_2 = \frac{1}{\sqrt{2}}dt$.

Сума на нормално разпределени случайни величини

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[(y_1-y_2)^2+y_2^2\right]}\, dy_2$$

Ще решим този интеграл, като го сведем до интеграл от нормална плътност. За целта ще допълним израза в експонентата до точен квадрат по y_2

$$(y_1 - y_2)^2 + y_2^2 = y_1^2 - 2y_1y_2 + 2y_2^2 = \left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}$$

Ще заместим изразът в интеграла и ще извадим пред него множителите, които не зависят от променливата на интегриране y_2.

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left[\left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2 + \frac{y_1^2}{2}\right]}\, dy_2 = \frac{e^{-\frac{y_1^2}{4}}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\sqrt{2}\, y_2 - \frac{y_1}{\sqrt{2}}\right)^2}\, dy_2$$

Сега ще направим смяна на променливите $t = \sqrt{2} y_2$. Ясно е, че границите на интегрирането се запазват, а $dy_2 = \frac{1}{\sqrt{2}} dt$.

$$f_{Y_1}(y_1) = \frac{e^{-\frac{y_1^2}{4}}}{2\pi\sqrt{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(t - \frac{y_1}{\sqrt{2}}\right)^2}\, dt$$

Сума на нормално разпределени случайни величини

Да припомним, от предната лекция знаем, че интеграл от нормална плътност е единица, т.е. за всяко реално μ и $\sigma > 0$ е изпълнено

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{(t-\mu)^2}{2\sigma^2}}\, dx = 1.$$

В нашия случай $\frac{y_1}{\sqrt{2}}$ е константа от гледна точка на интегрирането, тогава можем да приемем $\mu = \frac{y_1}{\sqrt{2}}$ и $\sigma = 1$. Така, окончателно получаваме

$$f_{Y_1}(y_1) = \frac{e^{-\frac{y_1^2}{2\,.\,2}}}{\sqrt{2\pi}\sqrt{2}}$$

Което явно е нормална плътност $N(0,2)$, т.е. $Y1 = X_1 + X_2 \in N(0,2)$. $\square$

Сума на нормално разпределени случайни величини

Да припомним, от предната лекция знаем, че интеграл от нормална плътност е единица, т.е. за всяко реално μ и $\sigma > 0$ е изпълнено

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{(t-\mu)^2}{2\sigma^2}}\, dx = 1.$$

В нашия случай $\frac{y_1}{\sqrt{2}}$ е константа от гледна точка на интегрирането, тогава можем да приемем $\mu = \frac{y_1}{\sqrt{2}}$ и $\sigma = 1$. Така, окончателно получаваме

$$f_{Y_1}(y_1) = \frac{e^{-\frac{y_1^2}{2\,.\,2}}}{\sqrt{2\pi}\sqrt{2}}$$

Което явно е нормална плътност $N(0,2)$, т.е. $Y1 = X_1 + X_2 \in N(0,2)$. $\square$

Доказаното свойство по индукция се обобщава за повече от две сл.в.

Твърдение

Нека $X_i \in N(\mu_i, \sigma_i^2)$, $i = 1, \dots n$ и случайните величини са независими в съвкупност, тогава

$$X_1 + \dots + X_n \in N\left(\mu_1 + \dots + \mu_n,\ \sigma_1^2 + \dots + \sigma_n^2\right)$$

Гама разпределение

По нататък ще опишем свойствата на няколко разпределения, които имат важно приложение в статистиката.

Гама разпределение - $X \in \Gamma(\alpha, \beta)$

Казваме, че случайна величина X има гама разпределение с параметри $\alpha > 0$ и $\beta > 0$, ако плътността и има вида:

$$f_X(x) = \begin{cases} \dfrac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

където $\Gamma(\alpha)$ е Ойлеровата гама функция, т.е. $\Gamma(\alpha) = \displaystyle\int_0^\infty x^{\alpha-1} e^{-x} dx$.

Гама разпределение

По нататък ще опишем свойствата на няколко разпределения, които имат важно приложение в статистиката.

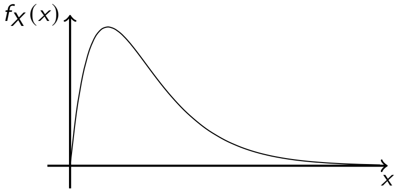
Гама разпределение - $X \in \Gamma(\alpha, \beta)$

Казваме, че случайна величина X има гама разпределение с параметри $\alpha > 0$ и $\beta > 0$, ако плътността и има вида:

$$f_X(x) = \begin{cases} \dfrac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

където $\Gamma(\alpha)$ е Ойлеровата гама функция, т.е. $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$.

Най общо плътността изглежда по следния начин. Като в зависимост от параметрите, кривата може да се доближи до вида на експоненциалното разпределение (при малки стойности на α) или на нормалното (при големи стойности на α).



Гама разпределение

Гама разпределението е от тези семейства, които са затворени относно сума. Други подобни са нормално, биномно, поасоново разпределение.

Твърдение

Нека случайните величини $X_1 \in \Gamma(\alpha_1, \beta)$ и $X_2 \in \Gamma(\alpha_2, \beta)$ са независими. Тогава $X_1 + X_2 \in \Gamma(\alpha_1 + \alpha_2, \beta)$.

Гама разпределение

Гама разпределението е от тези семейства, които са затворени относно сума. Други подобни са нормално, биномно, поасоново разпределение.

Твърдение

Нека случайните величини $X_1 \in \Gamma(\alpha_1, \beta)$ и $X_2 \in \Gamma(\alpha_2, \beta)$ са независими. Тогава $X_1 + X_2 \in \Gamma(\alpha_1 + \alpha_2, \beta)$.

Няма да доказваме това твърдение. Доказателството отново се основава на формулата за смяна на променливите, то е подобно на това от аналогичното твърдение за сума на нормални сл.в.,но е по тежко аналитично и се използват свойства на познатите от анализа Гама и Бета функция на Ойлер.

Гама разпределение

Гама разпределението е от тези семейства, които са затворени относно сума. Други подобни са нормално, биномно, поасоново разпределение.

Твърдение

Нека случайните величини $X_1 \in \Gamma(\alpha_1, \beta)$ и $X_2 \in \Gamma(\alpha_2, \beta)$ са независими. Тогава $X_1 + X_2 \in \Gamma(\alpha_1 + \alpha_2, \beta)$.

Няма да доказваме това твърдение. Доказателството отново се основава на формулата за смяна на променливите, то е подобно на това от аналогичното твърдение за сума на нормални сл.в.,но е по тежко аналитично и се използват свойства на познатите от анализа Гама и Бета функция на Ойлер.

Този резултат се обобщава по индукция за повече от две сл.в.

Твърдение

Нека случайните величини $X_i \in \Gamma(\alpha_i, \beta)$ $i = 1, 2, \ldots, n$ са независими в съвкупност. Тогава $X_1 + X_2 + \ldots + X_n \in \Gamma(\alpha_1 + \alpha_2 + \ldots + \alpha_n, \beta)$.

Гама разпределение

Гама разпределението може да се разглежда като обобщение на експоненциалното. Гама функцията е непрекъсната функция, която в целите числа взима за стохности факториели - $\Gamma(n) = (n-1)!$, като $\Gamma(1) = 1$. Лесно се вижда, че $\Gamma(1, \lambda)$ има плътност $f(x) = \lambda e^{-\lambda x}$, която е от експоненциален вид, т.е $\Gamma(1, \lambda) \equiv \mathcal{E}x(\lambda)$.

Твърдение

Нека случайните величини $X_i \in \mathcal{E}x(\lambda)$ $i = 1, 2, \ldots, n$ са независими в съвкупност. Тогава $X_1 + X_2 + \ldots + X_n \in \Gamma(n, \lambda)$.

Док. $X_i \in \mathcal{E}x(\lambda)$ е еквивалентно на $X_i \in \Gamma(1, \lambda)$. Твърдението следва директно от предишното твърдение. $\square$

Използвайки тази връзка лесно можем да намерим очакването и дисперсията на гама разпределението, поне в случая когато $\alpha = n$ е цяло число. Нека $Y = X_1 + X_2 + \ldots + X_n$, като $X_i \in \mathcal{E}x(\lambda)$. Знаем че $\mathrm{E}X_i = \frac{1}{\lambda}$ и $\mathrm{D}X_i = \frac{1}{\lambda^2}$. От свойствата на очакване и дисперсията за $Y \in \Gamma(n, \lambda)$ получаваме

$$\mathrm{E}Y = \mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}X_i = \frac{n}{\lambda} \qquad \mathrm{D}Y = \mathrm{D}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{D}X_i = \frac{n}{\lambda^2}$$

Не е трудно да се покаже, че и в общия случай, ако $Y \in \Gamma(\alpha, \beta)$ то $\mathrm{E}Y = \frac{\alpha}{\beta}$ и $\mathrm{D}Y = \frac{\alpha}{\beta^2}$.

Хи-квадрат разпределение

При решаването на редица статистически задачи се налага да се изследват суми от квадрати на нормални случайни величини. Оказва се, че тези суми имат хи-квадрат разпределение. Броят на събираемите n е прието да се наричат „степени на свобода“.

Ние ще въведем хи-квадрат разпределението, като частен случай на гама разпределение.

Хи-квадрат разпределение

При решаването на редица статистически задачи се налага да се изследват суми от квадрати на нормални случайни величини. Оказва се, че тези суми имат хи-квадрат разпределение. Броят на събираемите n е прието да се наричат „степени на свобода“.

Ние ще въведем хи-квадрат разпределението, като частен случай на гама разпределение.

Хи-квадрат разпредление - $X \in \chi^2(n)$

Казваме, че сл.в. има хи-квадрат разпределение с n „степени на свобода“, $X \in \chi^2(n)$, ако $X \in \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$.

$$f_X(x) = \begin{cases} \dfrac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Хи-квадрат разпределение

При решаването на редица статистически задачи се налага да се изследват суми от квадрати на нормални случайни величини. Оказва се, че тези суми имат хи-квадрат разпределение. Броят на събираемите n е прието да се наричат „степени на свобода“.

Ние ще въведем хи-квадрат разпределението, като частен случай на гама разпределение.

Хи-квадрат разпредление - $X \in \chi^2(n)$

Казваме, че сл.в. има хи-квадрат разпределение с n „степени на свобода“, $X \in \chi^2(n)$, ако $X \in \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$.

$$f_X(x) = \begin{cases} \dfrac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Ясно е, че графиката на плътността е идентична с тази на гама разпределението, а също така

$$\mathrm{E}X = n \qquad \mathrm{D}X = 2n$$

Хи-квадрат разпределение

При решаването на редица статистически задачи се налага да се изследват суми от квадрати на нормални случайни величини. Оказва се, че тези суми имат хи-квадрат разпределение. Броят на събираемите n е прието да се наричат „степени на свобода“.

Ние ще въведем хи-квадрат разпределението, като частен случай на гама разпределение.

Хи-квадрат разпредление - $X \in \chi^2(n)$

Казваме, че сл.в. има хи-квадрат разпределение с n „степени на свобода“, $X \in \chi^2(n)$, ако $X \in \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$.

$$f_X(x) = \begin{cases} \dfrac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Ясно е, че графиката на плътността е идентична с тази на гама разпределението, а също така

$$\mathrm{E}X = n \qquad \mathrm{D}X = 2n$$

Основното свойство, поради което хи-квадрат разпределението се използва е формулирано в следващото твърдение.

Хи-квадрат разпределение

Твърдение

Нека сл.в. X_i, $i = 1, 2, \ldots, n$ са независими в съвкупност, стандартно нормално разпределени, т.е. $X_i \in N(0, 1)$. Тогава

$$X_1^2 + X_2^2 + \ldots + X_n^2 \in \chi^2(n)$$

Хи-квадрат разпределение

Твърдение

Нека сл.в. X_i, $i = 1, 2, \dots, n$ са независими в съвкупност, стандартно нормално разпределени, т.е. $X_i \in N(0, 1)$. Тогава

$$X_1^2 + X_2^2 + \ldots + X_n^2 \in \chi^2(n)$$

Док. За начало ще направимм доказателството за една случайна величина. Нека $X \in N(0, 1)$, ще покажем че $Y = X^2 \in \chi^2(1)$.

Твърдение

Нека сл.в. X_i, $i = 1, 2, \dots, n$ са независими в съвкупност, стандартно нормално разпределени, т.е. $X_i \in N(0, 1)$. Тогава

$$X_1^2 + X_2^2 + \dots + X_n^2 \in \chi^2(n)$$

Док. За начало ще направимм доказателството за една случайна величина. Нека $X \in N(0, 1)$, ще покажем че $Y = X^2 \in \chi^2(1)$.

Смяната $Y = X^2$ е по-сложна от разглежданите досега, тъи като функцията не е еднозначно обратима. На всяка стойност на Y съответстват две стойности на X. В този случай, не можем да приложим директно формулата за смяна на променливите вместо това ще разгледаме функцията на разпределение на Y. Нека $y \geq 0$ е произволно фиксирано число

$$F_Y(y) = \mathrm{P}(Y < y) = \mathrm{P}(X^2 < y) = \mathrm{P}(-\sqrt{y} < X < \sqrt{y})$$

Хи-квадрат разпределение

Твърдение

Нека сл.в. X_i, $i = 1, 2, \dots, n$ са независими в съвкупност, стандартно нормално разпределени, т.е. $X_i \in N(0, 1)$. Тогава

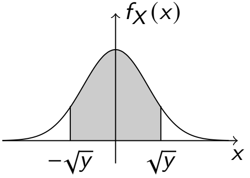
$$X_1^2 + X_2^2 + \dots + X_n^2 \in \chi^2(n)$$

Док. За начало ще направимм доказателството за една случайна величина. Нека $X \in N(0, 1)$, ще покажем че $Y = X^2 \in \chi^2(1)$.

Смяната $Y = X^2$ е по-сложна от разглежданите досега, тъи като функцията не е еднозначно обратима. На всяка стойност на Y съответстват две стойности на X. В този случай, не можем да приложим директно формулата за смяна на променливите вместо това ще разгледаме функцията на разпределение на Y. Нека $y \geq 0$ е произволно фиксирано число

$$F_Y(y) = \mathrm{P}(Y < y) = \mathrm{P}(X^2 < y) = \mathrm{P}(-\sqrt{y} < X < \sqrt{y})$$

Сл.в. X е нормално разпределена и нейната плътност е симетрична. Следователно $\mathrm{P}(-\sqrt{y} < X < \sqrt{y}) = 2\,\mathrm{P}(0 < X < \sqrt{y})$.



Хи-квадрат разпределение

Ще изразим функцията на разпределение на Y, чрез тази на X

$$F_Y(y) = 2\,\mathrm{P}(0 < X < \sqrt{y}) = 2\,\left[F_X(\sqrt{y}) - F_X(0)\right]$$

Хи-квадрат разпределение

Ще изразим функцията на разпределение на Y, чрез тази на X

$$F_Y(y) = 2\,\mathrm{P}(0 < X < \sqrt{y}) = 2\,\left[F_X(\sqrt{y}) - F_X(0)\right]$$

За да намерим плътността на Y е достатъчно да диференцираме това равенство. Ще припомним, че $F_X(0) = 1/2$, така от формулата за диференциране на сложна функция следва

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = 2\,f_X(\sqrt{y})\,\frac{1}{2\sqrt{y}}$$

Хи-квадрат разпределение

Ще изразим функцията на разпределение на Y, чрез тази на X

$$F_Y(y) = 2\,\mathrm{P}(0 < X < \sqrt{y}) = 2\,\left[F_X(\sqrt{y}) - F_X(0)\right]$$

За да намерим плътността на Y е достатъчно да диференцираме това равенство. Ще припомним, че $F_X(0) = 1/2$, така от формулата за диференциране на сложна функция следва

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = 2\, f_X(\sqrt{y})\, \frac{1}{2\sqrt{y}}$$

Но $X \in N(0,1)$ което означава, че $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Тогава

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}}\, \frac{1}{\sqrt{y}} = \frac{y^{\frac{1}{2}-1}\, e^{-\frac{y}{2}}}{2^{\frac{1}{2}}\, \Gamma\left(\frac{1}{2}\right)}$$

Хи-квадрат разпределение

Ще изразим функцията на разпределение на Y, чрез тази на X

$$F_Y(y) = 2\,\mathrm{P}(0 < X < \sqrt{y}) = 2\,\left[F_X(\sqrt{y}) - F_X(0)\right]$$

За да намерим плътността на Y е достатъчно да диференцираме това равенство. Ще припомним, че $F_X(0) = 1/2$, така от формулата за диференциране на сложна функция следва

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = 2\,f_X(\sqrt{y})\,\frac{1}{2\sqrt{y}}$$

Но $X \in N(0,1)$ което означава, че $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Тогава

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{\sqrt{y}} = \frac{y^{\frac{1}{2}-1}\,e^{-\frac{y}{2}}}{2^{\frac{1}{2}}\,\Gamma\left(\frac{1}{2}\right)}$$

В последното равенство използвахме още едно от свойствата на Гама функция $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Получената плътност $f_Y(y)$ е от типа $\chi^2(1)$.

Хи-квадрат разпределение

Ще изразим функцията на разпределение на Y, чрез тази на X

$$F_Y(y) = 2\,\mathrm{P}(0 < X < \sqrt{y}) = 2\,\left[F_X(\sqrt{y}) - F_X(0)\right]$$

За да намерим плътността на Y е достатъчно да диференцираме това равенство. Ще припомним, че $F_X(0) = 1/2$, така от формулата за диференциране на сложна функция следва

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = 2\, f_X(\sqrt{y})\, \frac{1}{2\sqrt{y}}$$

Но $X \in N(0,1)$ което означава, че $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Тогава

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{\sqrt{y}} = \frac{y^{\frac{1}{2}-1}\, e^{-\frac{y}{2}}}{2^{\frac{1}{2}}\,\Gamma\left(\frac{1}{2}\right)}$$

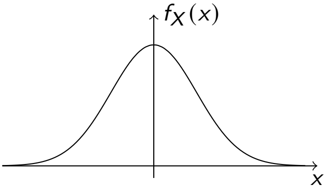
В последното равенство използвахме още едно от свойствата на Гама функция $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Получената плътност $f_Y(y)$ е от типа $\chi^2(1)$.

Сега ще разгледаме общия случай. Нека $X_i \in N(0,1)$, $i = 1, 2, \ldots, n$, от току що доказаното $X_i^2 \in \chi^2(1) \equiv \Gamma(\frac{1}{2}, \frac{1}{2})$. Съгласно твърдението за сума на гама разпределени сл.в. (стр.14)

$$X_1^2 + X_2^2 + \ldots + X_n^2 \in \Gamma\left(\frac{n}{2}, \frac{1}{2}\right) \equiv \chi^2(n) \qquad \square$$

Разпределение на Стюдънт

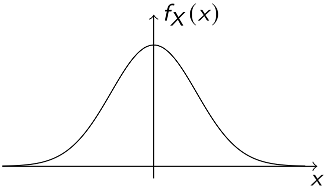
Разпределението на Стюдънт е важно за решаването на много статистически задачи. Нарича се още t-разпределение и съкратено се записва $X \in t(n)$ където n - "степени на свобода" е естествено число.



Плътността е камбановидна крива. Визуално тя трудно може да бъде различена от нормалната плътност, при големи стойности на n на практика съвпада с нея.

Разпределение на Стюдънт

Разпределението на Стюдънт е важно за решаването на много статистически задачи. Нарича се още t-разпределение и съкратено се записва $X \in t(n)$ където n - "степени на свобода" е естествено число.



Плътността е камбановидна крива. Визуално тя трудно може да бъде различена от нормалната плътност, при големи стойности на n на практика съвпада с нея.

Няма да извеждаме свойствата на това разпределение, само ще споменем случая, в който възниква.

Твърдение

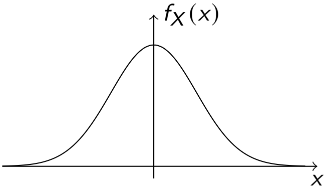
Нека $Z \in N(0,1)$ и $X \in \chi^2(n)$ са независими сл.в. Тогава

$$\frac{Z}{\sqrt{X/n}} \in t(n)$$

Доказателството се извършва стандартно със смяна на променливите, подобно на приведените по-горе доказателства.

Разпределение на Стюдънт

Разпределението на Стюдънт е важно за решаването на много статистически задачи. Нарича се още t-разпределение и съкратено се записва $X \in t(n)$ където n - "степени на свобода" е естествено число.



Плътността е камбановидна крива. Визуално тя трудно може да бъде различена от нормалната плътност, при големи стойности на n на практика съвпада с нея.

Няма да извеждаме свойствата на това разпределение, само ще споменем случая, в който възниква.

Твърдение

Нека $Z \in N(0,1)$ и $X \in \chi^2(n)$ са независими сл.в. Тогава

$$\frac{Z}{\sqrt{X/n}} \in t(n)$$

Доказателството се извършва стандартно със смяна на променливите, подобно на приведените по-горе доказателства.