СЕМ, Л-08
(2020-11-19)

Съвместно разпределение на дискретни случайни величини

Нека X, Y са две дискретни случайни величини. Интересуваме се от (X, Y).
Дефиниция. Нека (X, Y) са две дискретни случайни величини. Тогава таблицата по-долу се наричаме съвместно разпределение на X и Y:

	x_1	x_2	$\cdots$	x_k	$\cdots$	
y_1	p_{11}	p_{21}	$\cdots$	p_{n1}	$\cdots$	$\Sigma_i p_{i1}$
y_2	p_{12}	p_{22}	$\cdots$	p_{n2}	$\cdots$	$\Sigma_i p_{i2}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
y_k	p_{1k}	p_{2k}	$\cdots$	p_{nk}	$\cdots$	$\Sigma_i p_{ik}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	$\Sigma_j p_{1j}$	$\Sigma_j p_{2j}$	$\cdots$	$\Sigma_j p_{nj}$	$\cdots$	

Където $0 \leq p_{ij} = \mathbb{P}(X = x_i,\ Y = y_j)$, за $\forall i, j \in \{\text{Table Indexes}\}$. $\Sigma_{i,j} p_{ij} = 1$.

Маргинално разпределение на X:

x_1	x_2	$\cdots$	x_k	$\cdots$
$\Sigma_j p_{1j}$	$\Sigma_j p_{2j}$	$\cdots$	$\Sigma_j p_{nj}$	$\cdots$

Разпределенията само на X или само на Y се наричат маргинални разпределения.

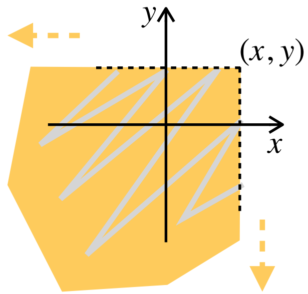
$\oplus$ Хвърляме два стандартни честни зара $\{1, 2, 3, 4, 5, 6\}$. Нека X е броят шестици, а Y е броят единици. Търси се (X, Y). За едно хвърляне може да имаме минимум 0 аксимум 2 шестици или единици.

	$X = 0$	1	2	
$Y = 0$	$\left(\frac{4}{6}\right)^2$	$\binom{2}{1} \cdot \frac{1}{6} \cdot \frac{4}{6}$	$\left(\frac{1}{6}\right)^2$	$\frac{25}{36}$
1	$\binom{2}{1} \cdot \frac{1}{6} \cdot \frac{4}{6}$	$\binom{2}{1} \cdot \frac{1}{6} \cdot \frac{1}{6}$	0	$\frac{10}{36}$
2	$\left(\frac{1}{6}\right)^2$	0	0	$\frac{1}{36}$
	$\frac{25}{36}$	$\frac{10}{36}$	$\frac{1}{36}$	

Дефиниция (Функция на разпределение на случайни величини). Нека (X, Y) се състои от *произволни* случайни величини. Тогава:

$$F_{XY}(x, y) = \mathbb{P}(\{X < x\} \cap \{Y < y\}) = \mathbb{P}(X < x,\ Y < y),$$

дефинирана за $\forall (x, y) \in \mathbb{R}^2$ се нарича функция на разпределение на X и Y. Като $F_X(x) = F_{XY}(x, \infty)$ и $F_Y = F_{XY}(\infty, y)$



Дефиниция (Независимост $\perp\!\!\!\perp$). Произволни случайни величини X и Y са независими ($X \perp\!\!\!\perp Y$) тогава и само тогава, когато е изпълнено:

$$F_{XY}(x, y) = F_X(x) \cdot F_Y(y) = F_{XY}(x, \infty) \cdot F_{XY}(\infty, y),\ \forall (x, y) \in \mathbb{R}^2.$$

Ковариация на две случайни величини X и Y

Линейна зависимост (например Y зависи линейно от X) между двете случайни величини имаме когато $Y = aX + b$.

Ковариацията е мярка за посоката на линейна връзка.

Дефиниция (Ковариация). Нека X и Y са случайни величини с $\text{Var}[X] < \infty$ и $\text{Var}[Y] < \infty$ (с крайни дисперсии). Тогава:

$$\text{Cov}[X, Y] = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$$

наричаме ковариация между X и Y. Ковариацията измерва как две случайни величини се изменят заедно.

Интуиция:

- Положителна ковариация: Когато X е над средната си стойност, Y също има тенденция да е над средната си
- Отрицателна ковариация: Когато X е над средната си, Y има тенденция да е под средната си
- Нулева ковариация: Няма линейна зависимост (но може да има нелинейна!)

Алтернативна формула (по-лесна за изчисление).

$$\mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Това е формулата, която най-често използваме за изчисления.

Доказателство:

$$\begin{aligned}\mathrm{Cov}[X, Y] = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right] &= \\ = \mathbb{E}(XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[XY]) & \qquad \mathbb{E} \text{ е линеен функционал} \\ = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] &\end{aligned}$$

Следствие: Ако $X \perp\!\!\!\perp Y$ (са независими), то $\mathrm{Cov}(X, Y) = 0$.

Доказателство: Ако $X \perp\!\!\!\perp Y$, то $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \Rightarrow \underbrace{\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]}_{\mathrm{Cov}(X,Y)} = 0$.

Смяна на променливите: ако $X \longmapsto 10X$, $Y \longmapsto 10Y$, то

$$\mathrm{Cov}[10X,\ 10Y] = 100 \cdot \mathrm{Cov}[X,\ Y]$$

Свойства

- Симетрия: $\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X]$
- Ковариация със себе си е дисперсия: $\mathrm{Cov}[X, X] = \mathrm{Var}[X]$
- Линейност: $\mathrm{Cov}[aX + b, cY + d] = ac \cdot \mathrm{Cov}[X, Y]$
- Адитивност: $\mathrm{Cov}[X + Z, Y] = \mathrm{Cov}[X, Y] + \mathrm{Cov}[Z, Y]$

Дефиниция (Коефициент на корелация на Пиърсън) (Карл Пиърсън (1857–1936), Великобритания). Нека X и Y са случайни величини с $\mathrm{Var}[X] < \infty$ и $\mathrm{Var}[Y] < \infty$ (с крайни дисперсии). Тогава:

$$\rho[X, Y] = \frac{\mathrm{Cov}[X, Y]}{\sigma_X \sigma_Y} = \frac{\mathbb{E}\left[(X - \mu_X)(Y - \mu_Y)\right]}{\sigma_X \sigma_Y}$$

се нарича коефициент на корелация между X и Y.

Корелацията измерва силата и посоката на линейната връзка между две случайни величини.

Твърдение. Нека X и Y са случайни величини с крайни дисперсии и нека

$$\overline{X} = \frac{X - \mathbb{E}[X]}{\sigma_X} \text{ и } \overline{Y} = \frac{Y - \mathbb{E}[Y]}{\sigma_Y}.$$

Тогава са изпълнени следните равенства:

- $\rho[X, Y] = \mathbb{E}[\overline{X}\,\overline{Y}]$
- $\mathbb{E}[\overline{X}] = \mathbb{E}[\overline{Y}] = 0$

- $\mathrm{Var}[\overline{X}] = \mathbb{E}[\overline{X}^2] = \mathrm{Var}[\overline{Y}] = \mathbb{E}[\overline{Y}^2] = 1$

Доказателство:

- $$\rho[X,Y] = \frac{\mathrm{Cov}[X,Y]}{\sigma_x\sigma_Y} = \frac{(X-\mathbb{E}[X])(Y=\mathbb{E}[Y])}{\sigma_x\sigma_Y}$$
$$= \mathbb{E}\left[\left(\frac{X-\mathbb{E}[X]}{\sigma_X}\right)\left(\frac{Y-\mathbb{E}[Y]}{\sigma_Y}\right)\right] = \mathbb{E}[\overline{X}\,\overline{Y}]$$

- $\mathbb{E}[X] = \mathbb{E}\left[\dfrac{X-\mathbb{E}[X]}{\sigma_X}\right] = \dfrac{1}{\sigma_X}\cdot\mathbb{E}[X-\mathbb{E}[X]] = 0$

- $\mathrm{Var}[\overline{X}] = \mathrm{Var}\left[\dfrac{X-\mathbb{E}[X]}{\sigma_X}\right] = \dfrac{1}{\sigma_X^2}\cdot\mathrm{Var}[X-\underbrace{\mathbb{E}[X]}_{const.}] = \dfrac{1}{\mathrm{Var}[X]}\cdot\mathrm{Var}[X] = 1$

Свойства

- Диапазон: $-1 \le \rho \le 1$
- $\rho = 1 \rightarrow$ перфектна положителна линейна зависимост
- $\rho = -1 \rightarrow$ перфектна отрицателна линейна зависимост
- $\rho = 0 \rightarrow$ липса на линейна зависимост
- Безразмерна величина (няма единици)

Колкото е по-голяма е абсолютната стойност на коефициента на корелация, толкова по-добра е линейната зависимост между X и Y.

Доказателство:

$(\overline{X}-\overline{Y})^2$ е неотрицателно случайна величина . Следователно:

$$\begin{aligned}
&\mathbb{E}[(\overline{X}-\overline{Y})^2] \ge 0\\
&\Rightarrow \mathbb{E}[\overline{X}^2] + \mathbb{E}[\overline{Y}^2] - 2\mathbb{E}[\overline{X}\,\overline{Y}] = 2\left(1-\rho[X,Y]\right)\\
&\Rightarrow \rho[X,Y] \le 1
\end{aligned}$$

Аналогично,

$$\begin{aligned}
&\mathbb{E}(\overline{X}+\overline{Y})^2 \ge 0\\
&\Rightarrow \mathbb{E}[\overline{X}^2] + \mathbb{E}[\overline{Y}^2] + 2\mathbb{E}[\overline{X}\,\overline{Y}] = 2\left(1+\rho[X,Y]\right)\\
&\Rightarrow \rho[X,Y] \ge -1
\end{aligned}$$

Ще докажем, че ако $|\rho(X,Y)| = 1 \Leftrightarrow \exists a,b : a,b \in \mathbb{R} \wedge Y = aX + b$

$$
\begin{aligned}
&Y = aX + b \Leftrightarrow \\
&Y - \mathbb{E}[Y] = a(X - \mathbb{E}[X]) + a\mathbb{E}[X] + b - \mathbb{E}[Y] \Leftrightarrow \\
&\frac{Y - \mathbb{E}[Y]}{\sigma_Y} = \underbrace{\frac{a}{\sigma_Y} \cdot \sigma_X}_{v} \cdot \frac{X - \mathbb{E}[X]}{\sigma_X} + \underbrace{\frac{a\mathbb{E}[X] + b - \mathbb{E}[Y]}{\sigma_Y}}_{w} \Leftrightarrow \\
&\overline{Y} = v\overline{X} + w
\end{aligned}
$$

$$
0 = \mathbb{E}[\overline{Y}] = \mathbb{E}[v\overline{X} + w] = v\,\underbrace{\mathbb{E}[\overline{X}]}_{=0} + w \Rightarrow w = 0 \Rightarrow \overline{Y} = v\overline{X}
$$

$$
\mathrm{Var}[\overline{Y}] = v^2 \cdot \mathrm{Var}[\overline{X}] \Rightarrow v^2 = 1 \text{ или } v = \pm 1.
$$

$$
Y = aX + b \Leftrightarrow \overline{Y} = v\overline{X} \text{ и } v^2 = 1
$$

$(\Rightarrow)$ Нека $Y = aX + b$ или $\overline{Y} = v\overline{X}$. Тогава

$\rho[X, Y] \overset{\text{твърдение}}{=} \mathbb{E}[\overline{X}\,\overline{Y}] = v\mathbb{E}[\overline{X}^2] = v \Rightarrow \rho[X, Y] = \pm 1.$

$(\Leftarrow)$Нека $|\rho(X, Y)| = 1$. Тогава $|\mathbb{E}[\overline{X}\,\overline{Y}]| = 1$.

Да допуснем, че $\mathbb{E}[\overline{X}\,\overline{Y}] = 1$ (аналогично и за другия случай, в който $\mathbb{E}[\overline{X}\,\overline{Y}] = -1$).

Тогава,

$$
\begin{aligned}
&\mathbb{E}[(\overline{X} - \overline{Y})^2] = \mathbb{E}[\overline{X}^2] + \mathbb{E}[\overline{Y}^2] - 2\mathbb{E}[\overline{X}\,\overline{Y}] = 1 + 1 - 2 \cdot 1 = 0 \Rightarrow \\
&\mathbb{E}[(\overline{X} - \overline{Y})^2] = 0 \Rightarrow \overline{X} = \overline{Y}
\end{aligned}
$$

тъй като и $\mathrm{Var}[\overline{X} - \overline{Y}] = 0$.

Условно математическо очакване (УМО)

Знаем, че функцията $g(a) = \mathbb{E}[(X-a)^2]$, за $a \in \mathbb{R}$ се минимизира тогава и само тогава когато $a = \mathbb{E}[X]$. Тоест $\min\limits_{a\in\mathbb{R}} g(a) = \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathrm{Var}[X]$.

Доказателство:

Нека $\mu = \mathbb{E}[X]$ е математическото очакване на X.

Имаме:

$$\mathbb{E}[(X-a)^2] = \mathbb{E}[X^2 - 2aX + a^2] = \mathbb{E}[X^2] - 2a\mathbb{E}[X] + a^2.$$

Това е квадратична функция на a:

$$f(a) = a^2 - 2a\mu + \mathbb{E}[X^2]$$

Производната по a:

$$f'(a) = 2a - 2\mu$$

Приравняваме на нула:

$$2a - 2\mu = 0 \Rightarrow a = \mu$$

Втората производна е $f''(a) = 2 > 0$, така че това е минимум.

Горното доказателство е чисто аналитично, но може да подходим и чрез метод на „допълването“ за дисперсията:

Известно е, че:

$$\begin{aligned}\mathbb{E}[(X-a)^2] &= \mathbb{E}[(X-\mu+\mu-a^2)] \\ &= \mathbb{E}[(X-\mu)^2] + 2(\mu-a)\cdot\mathbb{E}[X-\mu] + (\mu-a)^2.\end{aligned}$$

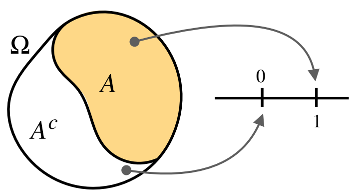
Тъй като $\mathbb{E}[X-\mu] = 0$, получаваме:

$$\mathbb{E}[(X-a)] = \mathrm{Var}[X] + (\mu-a)^2.$$

Първият член $\mathrm{Var}[X]$ не зависи от a, а вторият член $(\mu-a)^2 \geq 0$. Минимумът се достига при $a = \mu$, тогава $(\mu-a)^2 = 0$.

$\square$

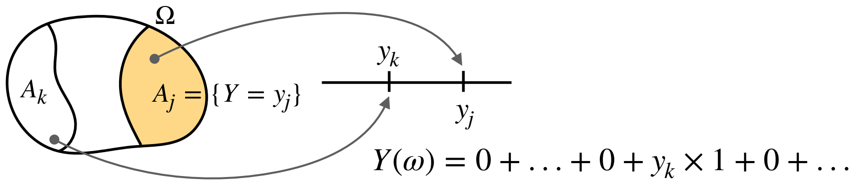
Нека $Y \sim \mathrm{Ber}(p)$. Тоест, $Y = \begin{cases} 1, & \text{с вероятност } p \\ 0, & \text{с вероятност } q = 1-p \end{cases}$



Да означим още събитията: $A = \{Y=1\}$ и $A^c = \{Y=0\}$.
Въвеждаме индикаторните функции: $Y = \mathbf{1}_A = \mathbf{1}_{\{Y=1\}}$.
Тогава, $p = \mathbb{E}[Y] = \mathbb{E}[\mathbf{1}_A] = \mathbb{E}[\mathbf{1}_{\{Y=1\}}] = \mathbb{P}(A) = \mathbb{P}(Y=1)$.
По-общо, ако имаме вероятностно пространство $(\Omega, \mathscr{A}, \mathbb{P})$ и $C \subseteq \mathscr{A}$, то $\mathbb{E}[\mathbf{1}_C] = \mathbb{P}(C)$.

Ако A и B са множества, то $\mathbf{1}_A\mathbf{1}_B = 1\mathbf{1}_{A\cap B}$. Следователно $\mathbb{E}[\mathbf{1}_A\mathbf{1}_B] = \mathbb{P}(A\cap B)$. Удобно е да записваме вероятностите като очакване на индикаторни функции, тъй като знаем, че очакването е линеен функционал и това може доста да ни помогне в някои случаи.

Ако Y е дискретна случайна величина, то $Y = \sum_i y_i \cdot \mathbf{1}_{A_i}$, където A_i е пълна група от събития и $\mathbb{P}(A_i) = \mathbb{P}(Y = y_i)$, $A_i = \{Y = y_i\}$.



⊕ Имаме нова случайна величина X и наблюдаваме Y, където $A = \{Y = 1\}$.
(Пример: X е клиент влязъл в магазин, а Y е дали клиента е мъж или жена)

$g: \{0, 1\} \to \mathbb{R}$ е функция.

$$\min_g \mathbb{E}[X - g(Y)]^2 = ?$$

От всички функции g искаме да вземем тази, която минимизира квадратичната грешка.

$$\mathbb{E}[X - f(Y)]^2,\ f(x) = ?$$

$g(Y) = a \times 1_A + b \times 1_{A^c} = aY + b(1 - Y)$, тъй като $1 - Y = \mathbf{1}_{A^c}$.

взаимно изключващи се събития $\mathbf{1}_A \cap \mathbf{1}_{A^c} = \varnothing \Rightarrow \mathbb{E}[\mathbf{1}_A\mathbf{1}_{A^c}] = 0$

$$\min_{a,b} \mathbb{E}[X - a\mathbf{1}_A - b\mathbf{1}_{A^c}]^2 =$$
$$= \min_{a,b}(\mathbb{E}[X^2] - a^2\mathbb{E}[\mathbf{1}_A] + b^2\mathbb{E}[\mathbf{1}_{A^c}] - 2a\mathbb{E}[X\mathbf{1}_A] - 2b\mathbb{E}[X\mathbf{1}_{A^c}] - 0)$$
$$= f(a,b)$$

Интересуваме се от $\min_{a,b} f(a,b) \Rightarrow \begin{cases} 0 = \dfrac{df}{da} = 2a\mathbb{E}[\mathbf{1}_A] - 2\mathbb{E}[X\mathbf{1}_A] \\ 0 = \dfrac{df}{db} = 2b\mathbb{E}[\mathbf{1}_{A^c}] - 2\mathbb{E}[X\mathbf{1}_{A^c}] \end{cases} \Rightarrow$

$$\boxed{a = \frac{\mathbb{E}[X\mathbf{1}_A]}{\mathbb{E}[\mathbf{1}_A]} \text{ и } b = \frac{\mathbb{E}[X\mathbf{1}_{A^c}]}{\mathbb{E}[\mathbf{1}_{A^c}]}}$$

$$g(Y) = \underbrace{\frac{\mathbb{E}[X\mathbf{1}_A]}{\mathbb{P}(\mathbf{1}_A)}}_{a} \cdot \mathbf{1}_A + \underbrace{\frac{\mathbb{E}[X\mathbf{1}_{A^c}]}{\mathbb{P}(\mathbf{1}_{A^c})}}_{b} \cdot \mathbf{1}_{A^c} \overset{\text{екв.}}{\leftrightarrow} \frac{\mathbb{E}[XY]}{\mathbb{P}(Y=1)} \cdot \mathbf{1}_{\{Y=1\}} + \frac{\mathbb{E}[1-Y]}{\mathbb{P}(Y=0)} \cdot \mathbf{1}_{\{Y=0\}}$$

Дефиниция (Условно математическо очакване (УМО)). Нека X и Y са две случайни величини. Тогава:

$$\mathbb{E}[X\,|\,Y] = f(y): \ \min_g \mathbb{E}[X - g(Y)]^2 = \mathbb{E}[X - f(Y)^2]$$

$\oplus\ X = \mathbf{1}_B$

$$\mathbb{E}[X\,|\,Y] = \mathbb{E}[\mathbf{1}_B\,|\,Y] = \frac{\mathbb{E}\mathbf{1}_B\mathbf{1}_A}{\mathbb{P}(A)}\cdot\mathbf{1}_A + \frac{\mathbb{E}\mathbf{1}_B\mathbf{1}_{A^c}}{\mathbb{P}(A^c)}\cdot\mathbf{1}_{A^c} = = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)}\cdot\mathbf{1}_A + \frac{\mathbb{P}(A^c\cap B)}{\mathbb{P}(A^c)}\cdot\mathbf{1}_{A^c}$$
$$= \underbrace{\mathbb{P}(B\,|\,A)}_{\substack{\text{най-доброто}\\ \text{приближение на } X\\ \text{когато } Y=1}} \cdot\mathbf{1}_A + \underbrace{\mathbb{P}(B\,|\,A^c)}_{\substack{\text{най-доброто}\\ \text{приближение на } X\\ \text{когато } Y=0}}\cdot\mathbf{1}_{A^c}$$

Твърдение. Нека X и Y са случайни величини, като Y е дискретна.

Y	Y_1	$\dots$	y_i	$\dots$
$\mathbb{P}$	p_1	$\dots$	p_i	$\dots$

$$Y = \sum_i y_i\mathbf{1}_{A_j},\ \ A_i = \{Y = y_i\},\ \ \mathbb{P}(A_i) = p_i.$$

Тогава $\mathbb{E}[X\,|\,Y] = \sum_i \frac{\mathbb{E}[X\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)}\cdot\mathbf{1}_{A_i}$. Количеството $\mathbb{E}[X\,|\,Y=y_i] = \frac{\mathbb{E}[X\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)}$ се нарича условно очакване на X при положение (условие), че $Y = y_i$.

$\oplus\ X = \mathbf{1}_B,\ B = \{X = 1\}$

$$\mathbb{E}[\mathbf{1}_B\,|\,Y=y_i] = \frac{\mathbb{E}[\mathbf{1}_B\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)} = \frac{\mathbb{P}(B\cap A_i)}{\mathbb{P}(A_i)} = \mathbb{P}(B\,|\,A_i) = \mathbb{P}[X=1\,|\,Y=y_i]$$

$\oplus\ X = \sum_i x_i\mathbf{1}_{B_i}$, X е дискретна случайна величина.

$$Y = \sum_i y_i\mathbf{1}_A \text{ и } p_{ij} = \mathbb{P}(X = x_i\cap Y = y_j)$$

$$\mathbb{E}[X\,|\,Y] = \sum_i \frac{\mathbb{E}[X\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)}\cdot\mathbf{1}_{A_i}$$

$$\underbrace{\mathbb{E}[X\mathbf{1}_{A_i}]} = \mathbb{E}\left[\sum_i x_i\mathbf{1}_{B_i}\right] = \sum_i x_i\mathbb{E}[\mathbf{1}_{B_i}\mathbf{1}_{A_i}] =$$
$$= \sum_i x_i\cdot\mathbb{P}(B_i\cap A_i) = \sum_i x_i\cdot p_{ij}$$

$$\mathbb{E}[X \mid Y = y_j] = \frac{\sum_i x_i \cdot p_{ij}}{\mathbb{P}(A_j)} = \sum_i x_i \frac{\mathbb{P}(A_j \cap B_i)}{\mathbb{P}(A_j)}$$
$$= \sum_i x_i \cdot \mathbb{P}(B_i \mid A_j) = \sum_i x_i \cdot \mathbb{P}(X = x_i \mid Y = y_i)$$

Твърдение. Нека X и Y са случайни величини, като Y е дискретна. Тогава $\mathbb{E}[X \mid Y]$ е дискретна случайна величина и $\mathbb{E}[X \mid Y] = \sum_i \mathbb{E}[X \mid Y = y_i] \cdot \mathbf{1}_{A_i}$.

$\mathbb{E}[X \mid Y]$	$\dots$	$\mathbb{E}[X \mid Y = y_i]$
$\mathbb{P}$	$\dots$	$\mathbb{P}(A_i)$

Свойства на условното математическо очакване

Теорема. Нека X, Y са случайни величини и Z е дискретна случайна величина $Z = \sum_i z_i \mathbf{1}_{A_i}$. Тогава:

a) Линейност: $\mathbb{E}[aX + bY \mid Z] = a \cdot \mathbb{E}[X \mid Z] + b \cdot \mathbb{E}[B \mid Z]$

b) Независимост: Ако $X \perp\!\!\!\perp Z$ (X е независима от Z), то $\mathbb{E}[X \mid Z] = \mathbb{E}[X]$

c) Точност: Ако $X = g(Z)$ (X е Z-измерима), то $\mathbb{E}[X \mid Z] = g(Z)$

d) Итеративно очакване (Tower property): $\mathbb{E}\big[\mathbb{E}[X \mid Z]\big] = \mathbb{E}[X]$

e) Изваждане на известното: $\mathbb{E}\big[f(U, Z) \mid Z = z_j\big] = \mathbb{E}[f(U, z_i)]$, където

- U е конкретна случайна величина $U = X$;
- U е вектор от случайни величини $U = (X_i)_{i \geq 1}$
- U е редица от случайни величини $U = \{X_1, \dots, X_n\}$, ако $U \perp\!\!\!\perp Z$.

Доказателство:

a) Линейност:

$$\mathbb{E}[aX + bY \mid Z] = \sum_i \frac{\mathbb{E}\big[(aX + bY) \cdot \mathbf{1}_{A_i}\big]}{\mathbb{P}(A_i)} \cdot \mathbf{1}_{A_i} = \sum_i \frac{a\mathbb{E}[X\mathbf{1}_{A_i}] + b\mathbb{E}[Y\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)} \cdot \mathbf{1}_{A_i} =$$
$$= a \underbrace{\sum_i \frac{\mathbb{E}[X\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)} \cdot \mathbf{1}_{A_i}} + b \underbrace{\sum_i \frac{\mathbb{E}[Y\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)} \cdot \mathbf{1}_{A_i}} = a\mathbb{E}[X \mid Z] + b\mathbb{E}[Y \mid Z].$$

b) Независимост:

Нека X е дискретна (ще го докажем само в този случай, макар и да е вярно и ако не е дискретна)

$$\Rightarrow \mathbb{E}[X\,|\,Z] = \sum_i \underbrace{\mathbb{E}[X\,|\,Z=z_i]}\cdot \mathbf{1}_{A_i} = \mathbb{E}[X\,|\,Z=z_i] = \sum_i x_i \cdot \mathbb{P}(X=x_i\,|\,Z=z_i) =$$

$$\overset{X\perp\!\!\!\perp Z}{=} \sum_i \frac{\mathbb{P}(X=x_i)\cancel{\mathbb{P}(Z=z_i)}}{\cancel{\mathbb{P}(Z=z_i)}} = \underbrace{\sum_i \mathbb{E}[\mathbf{1}_{A_i}]}_{\text{пълна група от събития}} = \mathbb{E}[X]\,.$$

c) Точност:

$$\mathbb{E}[X\,|\,Z] = \sum_i \frac{\mathbb{E}\left[g(Z)\cdot \mathbf{1}_{A_i}\right]}{\mathbb{P}(A_i)}\cdot \mathbf{1}_{A_i} = S,$$

имаме, че $g(Z)\cdot \mathbf{1}_{A_j} = g(Z)\cdot \mathbf{1}_{\{Z=z_j\}} = g(z_j)\cdot \mathbf{1}_{\{Z=z_j\}}$

$$\Rightarrow S = \sum_i \frac{\mathbb{E}\left[g(z_j)\cdot 1_{A_i}\right]}{\mathbb{P}(A_i)}\cdot \mathbf{1}_{A_i} = \sum_i g(z_j)\cdot \frac{\cancel{\mathbb{P}(A_i)}}{\cancel{\mathbb{P}(A_i)}}\cdot \mathbf{1}_{A_i} = g(Z) = X.$$

d) Итеративно очакване (Tower property):

Нека X е дискретна.

$$\mathbb{E}[X\,|\,Z] = \sum_i \mathbb{E}[X\,|\,Z=z_i]\cdot \mathbf{1}_{A_i}$$

$$\mathbb{E}\left[\mathbb{E}[X\,|\,Z]\right] = \mathbb{E}\left[\sum_i \mathbb{E}[X\,|\,Z=z_i]\cdot \mathbf{1}_{A_i}\right] = \sum_i \mathbb{E}[X\,|\,Z=z_i]\mathbb{E}[\mathbf{1}_{A_i}] =$$

$$= \sum_i \mathbb{E}[X\,|\,Z=z_i]\cdot \mathbb{P}(Z=z_j) = \sum_i \frac{\mathbb{E}[X1_{A_i}]}{\cancel{\mathbb{P}(A_i)}}\cdot \cancel{\mathbb{P}(A_i)} =$$

$$= \mathbb{E}\left[\sum_i X\mathbf{1}_{A_i}\right] = \mathbb{E}[X]\underbrace{\sum_i \mathbf{1}_{A_i}}_{=1} = \mathbb{E}[X]\,.$$

e) Изваждане на известното:

⊕ Имаме следния модел:

В даден магазин влизат клиенти, като знаем, че мъжете закупуват нещо с вероятност $1/3$, а жените с вероятност $2/3$.

$$X = \{\text{закупува или не}\} = \begin{cases} 1, \text{ с вероятност } \mathbb{P}(Y) \\ 0, \text{ с вероятност } \mathbb{P}(Y)' \end{cases}$$

$$Z = \begin{cases} \frac{1}{3}, & \text{мъж (вероятност 1/2)} \\ \frac{2}{3}, & \text{жена (вероятност 1/2)} \end{cases}$$

$$X = Y_1 \cdot \mathbf{1}_{\{Y_1 = \frac{1}{3}\}} + Y_2 \cdot \mathbf{1}_{\{Y = \frac{2}{3}\}},$$

където $Y_1 \in \mathrm{Ber}\left(p = \dfrac{1}{3}\right)$, а $Y_2 \in \mathrm{Ber}\left(p = \dfrac{2}{3}\right)$.

$$X = f(Y_1, Y_2, Z) = \begin{cases} Y_1, & \text{при } Z = \frac{1}{3} \\ Y_2, & \text{при } Z = \frac{2}{3} \end{cases}$$

$$\begin{aligned}
\mathbb{E}[X] = \mathbb{P}(X = 1) &= \mathbb{E}\big[\mathbb{E}[X \,|\, Z]\big] \\
&= \mathbb{E}\left[X \,|\, Z = \frac{1}{3}\right] \cdot \mathbb{P}\left(Z = \frac{1}{3}\right) + \mathbb{E}\left[X \,|\, Z = \frac{2}{3}\right] \cdot \mathbb{P}\left(Z = \frac{2}{3}\right) = \\
&= \mathbb{E}\left[Y_1 \,|\, Z = \frac{1}{3}\right] \cdot \underbrace{\frac{1}{2} + \frac{1}{2}}_{\text{мъж или жена}} \cdot \mathbb{E}\left[Y_2 \,|\, Z = \frac{2}{3}\right] = \\
&= \frac{1}{2}\left(\mathbb{E}[Y_1] + \mathbb{E}[Y_2]\right) = \frac{1}{2}\left(\frac{1}{3} + \frac{2}{3}\right) = \frac{1}{2}.
\end{aligned}$$

$\oplus$ $U = (X_i)_{i \geq 1}$, X_i са независими една от друга случайни величини и $X_i \sim \mathrm{Ber}(p)$ (хвърляне на НЕчестна монета с вероятност p за ези и q за тура).

Нека $W \sim \mathrm{Geom}(r)$ и W не зависи от U. Търсим $\mathbb{E}\left[\sum_{i=1}^{W+1} X_i\right]$.

$f(U, W) = \sum_{i=1}^{w+1} X_i$, ако $W = w$. Тогава:

$$
\begin{aligned}
\mathbb{E}\sum_{i=1}^{W+1} &= \mathbb{E}\big[f(U,W)\big] = \mathbb{E}\big[\mathbb{E}[f(U,W)\,|\,W]\big] \\
&= \mathbb{E}\sum_{w=0}^{\infty}\mathbb{E}\big[f(U,W)\,|\,W=w\big]\cdot\mathbf{1}_{W=w} \\
&= \mathbb{E}\left[\sum_{w=0}^{\infty}\mathbb{E}\big[f(U,w)\big]\cdot\mathbf{1}_{W=w}\right] = \mathbb{E}\sum_{w=0}^{\infty}\left[\mathbb{E}\sum_{i=1}^{w+1}X_i\right]\cdot\mathbf{1}_{W=w} \\
&= \sum_{w=0}^{\infty}\mathbb{E}\sum_{i=1}^{w+1}X_i\cdot E[\mathbf{1}_{W=w}] = \sum_{w=0}^{\infty}\mathbb{E}\sum_{i=1}^{w+1}X_i\cdot\mathbb{P}(W=w) \\
&= \sum_{w=0}^{\infty}(w+1)\cdot p\cdot(1-r)^w\cdot r = pr\sum_{w=0}^{\infty}(w+1)(1-r)^w = T.
\end{aligned}
$$

$$
\frac{1}{(1-x)^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1+x}\right) \overset{|x|<1}{=} \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=0}^{\infty}x^n\right) = \sum_{n=1}^{\infty}n\,x^{n-1} = \sum_{n=0}^{\infty}(n+1)x^n
$$

$$
\Rightarrow T \overset{x=1-r}{=} pr\frac{1}{\big(1-(1-r)\big)^2} = pr\cdot\frac{1}{r^2} = \frac{p}{r}.
$$

Условни разпределения

Нека X и Y са дискретни случайни величини. Тогава под разпределение на X при условие $Y = y_j$ се разбира следната таблица:

$X\,|\,Y=y_j$	$\dots$	x_i	$\dots$
$\mathbb{P}$	$\dots$	$\mathbb{P}(X=x_i\,|\,Y=y_j)$	$\dots$

$$
\sum_j \mathbb{P}(X=x_i\,|\,Y=y_j) = 1,\ \forall j
$$

⊕ Хвърляме два зара $\{1, 2, 3, 4, 5, 6\}$. Нека X е броят шестици, а Y е броят единици. Търси се (X, Y). За едно хвърляне може да имаме 0, 1, 2 шестици или единици.

	$X=0$	1	2	
$Y=0$	$\left(\frac{4}{6}\right)^2$	$\binom{2}{1}\cdot\frac{1}{6}\cdot\frac{4}{6}$	$\left(\frac{1}{6}\right)^2$	$\mathbb{P}(Y=0)=\frac{25}{36}$
1	$\binom{2}{1}\cdot\frac{1}{6}\cdot\frac{4}{6}$	$\binom{2}{1}\cdot\frac{1}{6}\cdot\frac{1}{6}$	0	$\mathbb{P}(Y=1)=\frac{10}{36}$
2	$\left(\frac{1}{6}\right)^2$	0	0	$\mathbb{P}(Y=2)=\frac{1}{36}$
	$\mathbb{P}(X=0)=\frac{25}{36}$	$\mathbb{P}(X=1)=\frac{10}{36}$	$\mathbb{P}(X=2)=\frac{1}{36}$	

Условно разпределение:

	$X = 0$	1	2	
$Y = 0$	$\left(\frac{4}{6}\right)^2 \frac{4}{5} \cdot \frac{4}{5}$	$\frac{1}{5} \cdot \frac{4}{5} \cdot 2$	$\frac{1}{5} \cdot \frac{1}{5}$	$\sum_{i=0}^{2} \mathbb{P}(X \mid Y = i) = 1$
1	$\frac{5+5-2}{5+5}$	$\frac{1+1}{5+5}$	0	$\dots$
2	1	0	0	$\dots$