Привет !! Лекция 6

* характеристики случайных величин
* неравенства

опр $\mathrm{Cov}(X,Y) = E(X^c \cdot Y^c) = E\left[(X-E(X))\cdot(Y-E(Y))\right]$

$Y^c = Y - E(Y)$

$X^c = X - E(X)$

смысл:

$\mathrm{Cov}(X,Y)$ – задаёт геометрию

$\langle X,Y\rangle = \mathrm{Cov}(X,Y)$

билинейность

- $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$

симметричность

- $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

Аналог косинуса угла между СВ:

$$\cos(\angle a,b) = \frac{\langle a,b\rangle}{\|a\|\cdot\|b\|}$$

$\|a\| = \sqrt{\langle a,a\rangle}$

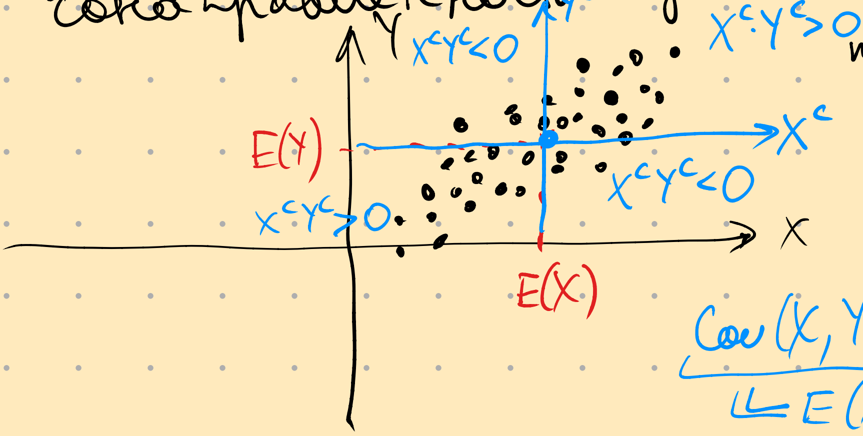
$\|b\| = \sqrt{\langle b,b\rangle}$

опр. Корреляция X и Y

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Cov}(X,X)\cdot\mathrm{Cov}(Y,Y)}} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}}$$

Что хорошего в Cov?

* легко считать
* определяет геометрию
* знак Cov определяет (в среднем) сонаправленность изменений.



n – число экспер-в

(X_i, Y_i)

$\mathrm{Cov}(X,Y) > 0$

$\Leftarrow E(X^cY^c)$

$Var(X) = E((X^c)^2) = Cov(X,X)$

(кван)

$\boxed{\text{формулы}}$

линейность

$Var(X+Y) = Cov(X+Y, X+Y) = Cov(X, X+Y) + Cov(Y, X+Y) =$

$= \underbrace{Cov(X,X)} + \underbrace{Cov(X,Y) + Cov(Y,X)} + \underbrace{Cov(Y,Y)} =$

$= Var(X) + 2Cov(X,Y) + Var(Y)$

$$\boxed{Var(X+Y) = Var(X) + 2Cov(X,Y) + Var(Y)}$$

9 класс

$(a+b)^2 = a^2 + 2ab + b^2$!!

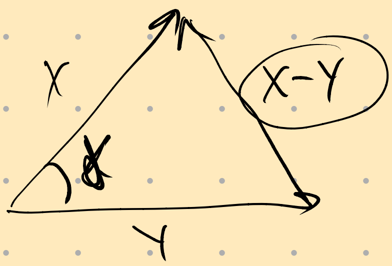
теорема косинусов.

$Y \to (-Y)$

$Var(X-Y) = Var(X) + 2Cov(X,-Y) + Var(-Y)$

$$\boxed{Var(X-Y) = Var(X) - 2Cov(X,Y) + Var(Y)}$$

$Var(X-Y) = Var(X) - 2Corr(X,Y) \cdot \sqrt{Var X \cdot Var Y} + Var(Y)$



$Var(X)$ — аналог квадр. длины X

$\sqrt{Var X}$ — аналог длины X

$c^2 = a^2 - 2\cos\cancel{\gamma}\, ab + b^2$

Критерий независимости

X и Y независимы $\iff E(f(X) \cdot g(Y)) = E(f(X)) \cdot E(g(Y))$ для $\forall f, g$.

$$\text{cov}(R,L) = E\big[(R-E(R))\cdot(L-E(L))\big] =$$

раскрою все скобки

$$= E\big[\underbrace{R\cdot L}_{\text{СВ}} - \underbrace{E(R)}_{\text{const}}\cdot\underbrace{L}_{\text{СВ}} - R\cdot E(L) + \underbrace{E(R)\cdot E(L)}_{\text{const}}\big] =$$

$E(42)=42$

$E(\text{const})=\text{const}$

$$= E(RL) - E(R)\cdot E(L) - E(R)\cdot E(L) + E(R)\cdot E(L)$$

$$\text{cov}(R,L) = E(RL) - E(R)\cdot E(L)$$

Критерий независимости СВ в терм-х cov

X и Y незав $\iff$ $\forall f,g$ $\text{cov}(f(X),g(Y))=0$

$\text{cov}(X,Y)=0$

$\text{cov}(X^2,Y^3)=0$ ...

$\text{cov}(\cos X, 3-Y)=0$

...

Следствие

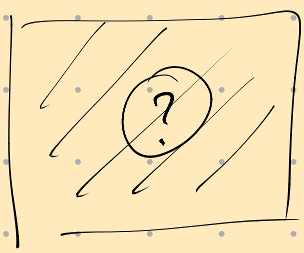
если X и Y незав, то

дост. услов

$$\text{Var}(X+Y)=\text{Var}(X-Y)=\text{Var}(X)+\text{Var}(Y)$$

(т. Пифагора)

Энтропия.



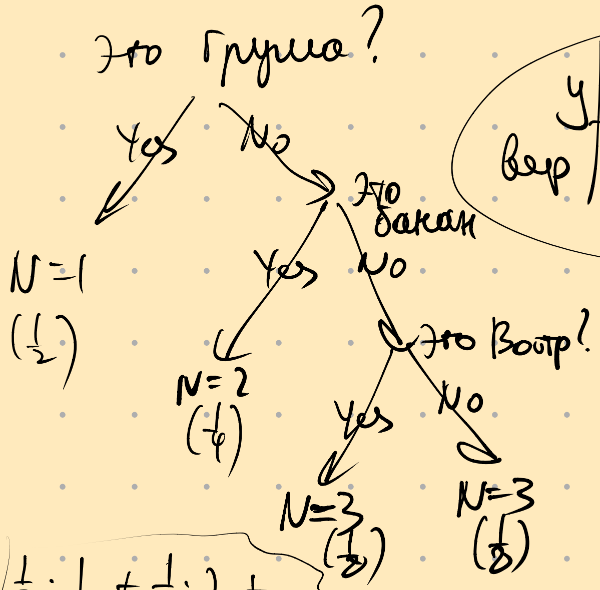
x	Арбуз	Банан	Виногр.	Груша
вер	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$

Игра. можно зад-ть вопросы (да/нет)

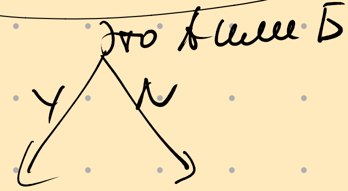
Цель: узнать X за мин-ое $\boxed{\text{ожидаемое}}$ число вопросов

$\min E(N)$

N - число вопросов



y	А	Б	В	Г
вер	1/4	1/4	1/4	1/4



$$E(N) = \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 3 + \frac{1}{8}\cdot 3 = \frac{14}{8} \quad [\text{бит}]$$

инт. опр.

$H(X)$ энтропия СВ X — это минимальное ожидаемое кол-во бинарных вопросов, затрачиваемое на получение инфо о значении X.

опр.

$$H(X) = \sum P(X=x)\cdot \log_{(\frac{1}{2})} P(X=x)$$

$$= -\sum P(X=x)\cdot \log_2 P(X=x) \quad [\text{бит}]$$

$\boxed{\times\,\text{const}}$

$$= -\sum P(X=x)\cdot \ln P(X=x) \quad [\text{нат}]$$

natural bit

Упр. сколько бит в нате?

$$\frac{\log_2 x}{\ln x} = ?$$

Упр. X — дискретная СВ [прин. счётное кол-во знач]

$$H(2X+2025) \overset{?}{=} H(X)$$

$2x+2025$

x	1	2	3	4
вер	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

y	2027	2029	2031	2033
	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

$$H(X) = H(Y)$$

<u>Опр</u> совместная энтропия $H(X,Y)$

<u>интуиция</u>: мин. ожидаемое кол-во вопросов, чтобы узнать и X, и Y.

$$V = \begin{pmatrix} X \\ Y \end{pmatrix}$$

[коот] $$H(X,Y) = -\sum_v P(V=v) \cdot \ln P(V=v) =$$

$$= -\sum_{x,y} P(X=x, Y=y) \cdot \ln P(X=x, Y=y)$$

<u>Теорема.</u>

(1) * $H(X,Y) \geq H(X)$

(2) * Если X и Y незав, то $\underline{H(X,Y) = H(X) + H(Y)}$

<u>док-во.</u> (2):

X и Y незав $\quad P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$

$$\ln P(X=x, Y=y) = \ln P(X=x) + \ln P(Y=y)$$

$$H(X,Y) = -\sum_{x,y} P(X=x) \cdot P(Y=y) \cdot \left(\ln P(X=x) + \ln P(Y=y)\right)$$

$$1 = \sum_y P(Y=y)$$

$$= -\sum_x \sum_y P(X=x) \cdot P(Y=y) \cdot \ln P(X=x)$$

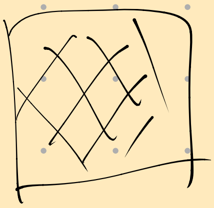
$$-\sum_x \sum_y P(X=x) P(Y=y) \cdot \ln P(Y=y) =$$

$$= -\sum_x P(X=x) \cdot \ln P(X=x)$$

$$-\sum_y P(Y=y) \cdot \ln P(Y=y)$$

$$= H(X) + H(Y)$$

* кросс-энтропия.



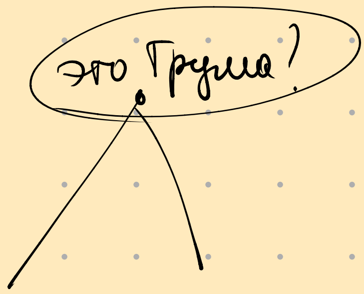
x	А	Б	В	Г
модель $q(x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$
истина $p(x)$	$1/4$	$1/4$	$1/4$	$1/4$

$CE(p \| q)$ — минимальное ожидаемое кол-во вопросов, чтобы узнать X, если истинные вер-сти — p, а минимизируем мы ошибочно думая, что вер-сти равны q.

„щ q в p"

Опр. $$CE(p\|q) = -\sum_x P(X=x) \ln Q(X=x)$$

$$H(p) = -\sum_x P(X=x) \cdot \ln P(X=x)$$



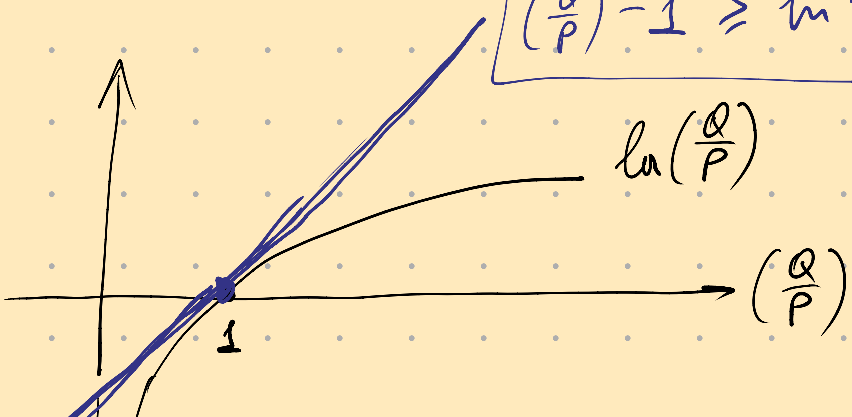
← дерево вопросов адаптируем под модельную Q

Теорема $$CE(p\|q) \geq H(p)$$

док-во.

Шаг!

$$\boxed{\left(\frac{Q}{P}\right) - 1 \geq \ln \frac{Q}{P}}$$



$$Q - P \geq P \cdot (\ln Q - \ln P)$$

$$\boxed{P - P\ln P \leq Q - P\ln Q} \quad \forall P, Q > 0$$

Шаг 2

$$\sum (p_i - p_i \ln p_i) \leq \sum (q_i - p_i \ln q_i)$$

$\sum p_i = 1 = \sum q_i$

$$1 - \sum p_i \ln p_i \leq 1 - \sum p_i \ln q_i$$

$$1 + H(p) \leq 1 + CE(p \,\|\, q)$$

$\boxtimes$

Условная энтропия.

* $H(Y|X)$ — мин. ожид-ое кол-во шагов чтобы узнать Y, если мы уже знаем X.

опр.

$$H(Y|X) = -\sum_{x,y} \underbrace{P(X=x, Y=y)}_{\text{безусл}} \cdot \ln \underbrace{P(Y=y \,|\, X=x)}_{\text{услов}}$$

Теоремы:

* $H(Y|X) \leq H(Y)$

* если X и Y незав, то $H(Y|X) = H(Y)$

* $H(X) + H(Y|X) = H(X, Y)$

опр. Дивергенция Кульбака-Лейблера

$KL = D_{KL}$

<u>инт:</u> дополнит-ое кол-во вопросов из-за несовпадения модели q

и реальных
вер-стей p

$$D_{KL}(p \| q) = CE(p \| q) - H(p)$$

Теор:

$$\boxed{D_{KL}(p \| q) \geq 0}$$