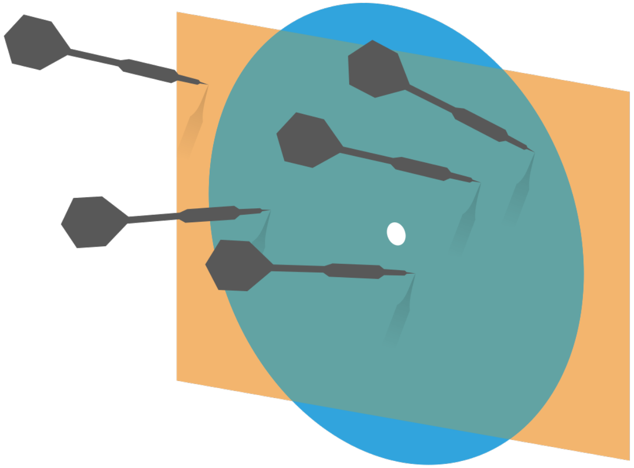
Probability and Statistics for Computer Science



Credit: wikipedia

"I have now used each of the terms mean, variance, covariance and standard deviation in two slightly different ways." ---Prof. Forsythe

Hongye Liu, Teaching Assistant Prof, CS361, UIUC, 9.17.2020

No pair in a hand of 5 from 52 cards

The probability of drawing hands of 5-cards that have no pairs.

(no replacement)

* Consider the order doesn't matter

$|E| = ?$

$$\frac{52 \times 48 \times \cdots \times 36}{5!} = \binom{13}{5} \cdot 4^5$$

$|\Omega| = ?$

No pair in a hand of 5 from 52 cards

The probability of drawing hands of 5-cards that have no pairs.

(no replacement)

* Consider the order doesn't matter

$|E| = ?$ $\binom{13}{5} \cdot 4^5$ → decide on suits for 5 cards ♢ ♡ ..

→ choose the number from [1, 2, 3 . . . 10, J, Q, K]

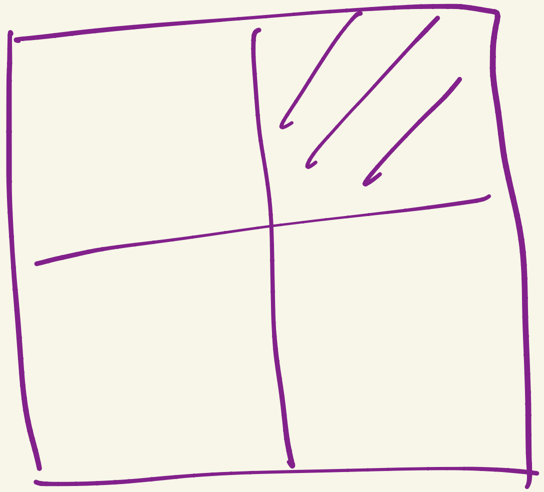
$|\Omega| = ?$ $\binom{52}{5}$

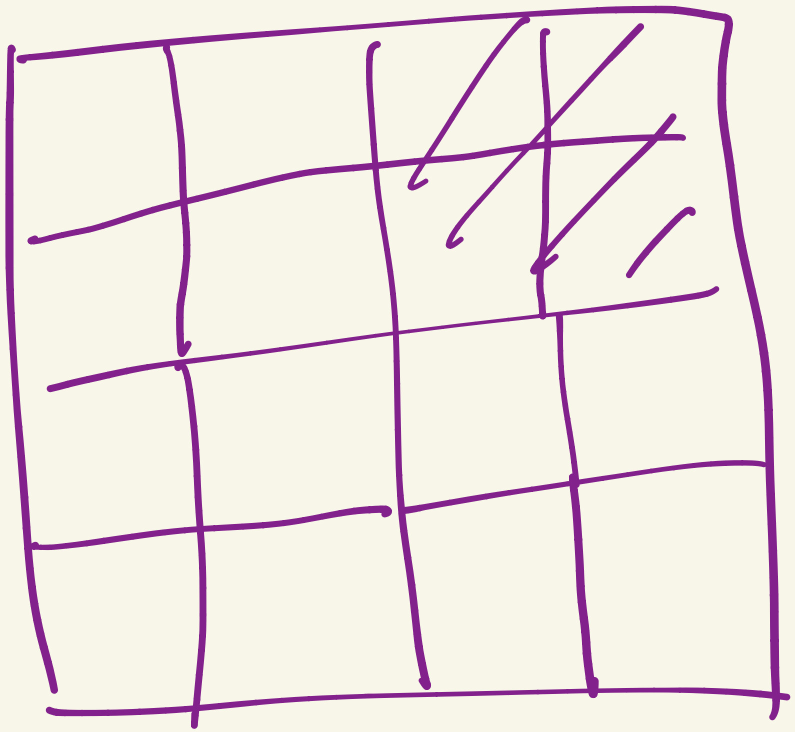
$|\Omega|$ # outcomes in the sample space

$|E|$ # outcomes in E

$$p = \frac{|E|}{|\Omega|}$$

make sure we have equal prob. for each outcome, apply the same principle for both E & Ω





i.e. take one card
that happens to be a ♡

$$\frac{13}{52}$$

→ 13 choices to pick a heart

→ 52 choices to pick a card

$$\frac{1}{4}$$

, out of 4 suits

Last time

Random Variable

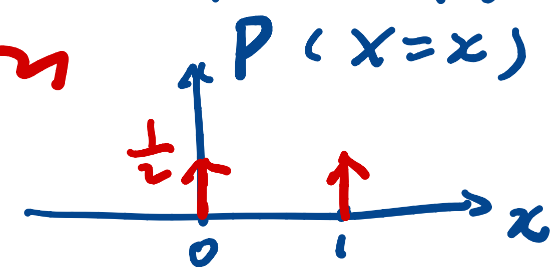
$X(\omega)$

$\omega \longrightarrow x$

$X = \begin{cases} 1 & \omega = \text{head} \\ 0 & \omega = \text{tail} \end{cases}$

* Definition
* Probability distribution

PDF, CDF



* Conditional probability distri.

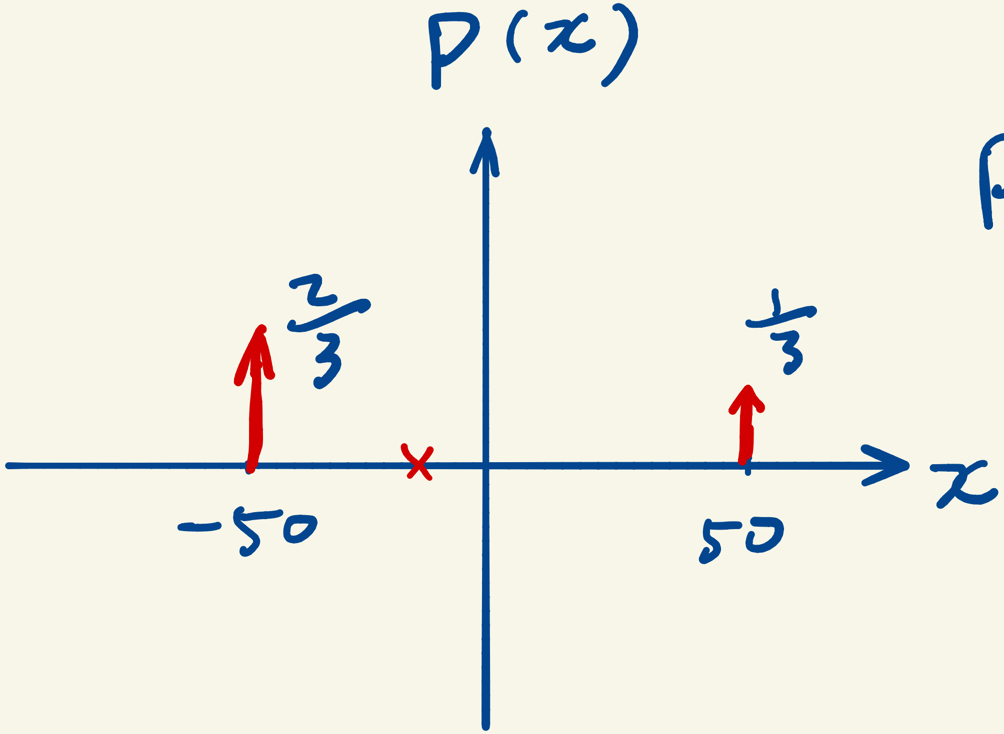
$P(X|Y)$

$P(X=x_0)$

$\to P(\{\omega \text{ s.t. } X(\omega)=x_0\})$

$$X(\omega) = \begin{cases} 50 & \omega = \text{head} \\ -50 & \omega = \text{tail} \end{cases}$$

$50 \to$ 50 dollars on the bet



$$P(X=x) = \begin{cases} \frac{1}{3} & x = 50 \\ \frac{2}{3} & x = -50 \\ 0 & \text{otherwise} \end{cases}$$

Summarize!!

Objectives

Random Variable (R.V.)

Definition

properties

* Expected value
* Variance & Covariance
* Markov's Inequality

X

$f(X)$

$f(X, Y)$

Expected value (Discrete case)

- The **expected value** (or **expectation**) of a random variable X is

$$E[X] = \sum_x x P(x)$$

P (X = x)

The expected value is a weighted sum of the values X can take

all

Expected value

- The **expected value** of a random variable X is

<= 1

1 heads

0 tails

$$E[X] = \sum_x x \boxed{P(x)}$$

$1 \times \frac{1}{2} + 0 \times \frac{1}{2}$

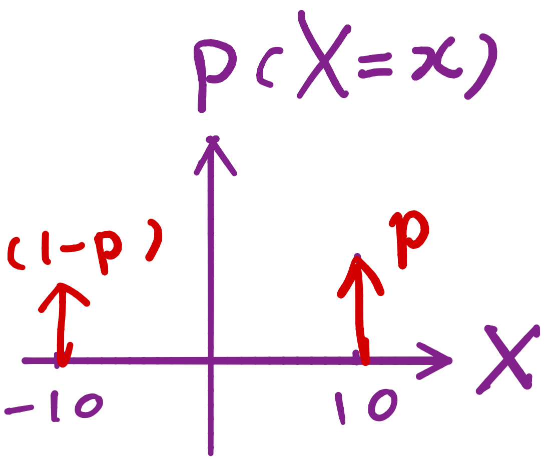
$\sum_x p(x) = 1$

The expected value is a weighted sum of all the values X can take

Expected value: profit

- A company has a project that has ***p*** probability of earning 10 million and ***1-p*** probability of losing 10 million.

- Let X be the return of the project.



$$E[X] = 10\cdot p + (-10)\times(1-p)$$

$$= 20p - 10 \geq 0$$

$$p \geq \frac{1}{2}$$

Solve at home

Cookies

$\underbrace{\quad\quad\quad}_{\$1 \text{ each}}$

$\boxed{\text{chocolate}} \downarrow$ \$2 each

A) random draw 1 from 4 items

Expected value = ?

B) random draw 1 twice with replacement independently

[if the two draws are the same, you get the prize.]

Expected value = ?

Linearity of Expectation

- For random variables X and Y and constants k,c
  - Scaling property
    $$E[kX] = kE[X]$$
  - Additivity
    $$E[X+Y] = E[X] + E[Y]$$
  - And $E[kX+c] = kE[X]+c$

Linearity of Expectation

- Proof of the additive property

$$E[X+Y] = E[X] + E[Y]$$

$S = X + Y$

$$E[X+Y] = E[S] = \sum_{s} s\,P(s)$$

$$= \sum_{\{s = x+y\}} s \sum_{\{s = x+y\}} P(x,y)$$

$$= \sum_{x} \sum_{y} (x+y) P(x,y)$$

$P(S=s)$

$= P(X=x, Y=y \text{ s.t. } s = x+y)$

$$E[X+Y] = E[S] = \sum_s s\,P(s)$$

$P(S=s)$
$= P(X=x, Y=y$
$\quad \& \; s = x+y)$

$$= \sum_{\{s=x+y\}} s \sum_{\{s=x+y\}} P(x,y)$$

$$= \sum_x \sum_y (x+y) P(x,y)$$

$$= \sum_x \sum_y x\,p(x,y) + \sum_x \sum_y y\,p(x,y)$$

$$= \sum_x x \underline{\sum_y p(x,y)} + \sum_y \sum_x y\,p(x,y)$$

$$= \sum_x x\,\underline{p(x)} + \sum_y y \underline{\sum_x p(x,y)}$$

$$= \underline{\sum_x x\,p(x)} + \underline{\sum_y y\,P(y)}$$

$$= E[X] + E[Y]$$

$E[X_1 + X_2 + \cdots]$

$= E[X_1] + E[X_2] + \cdots$

Q. What's the value?

- What is $E[E[X]+1]$?

A. $E[X]+1$ B. 1 C. 0

$= Y$

$E[Y+1]$

$= E[Y]+1$

$= E[E[X]]+1$

$= E[X]+1$

$P(Y) = 1$

$E[Y] = Y \times 1$

$= E[X]+1$

Expected value of a function of X

- If $\boldsymbol{f}$ is a function of a random variable X, then $Y = \boldsymbol{f}(X)$ is a random variable too
- The expected value of $Y = \boldsymbol{f}(X)$ is

$$E[Y] = E[f(X)] = \sum_x f(x)P(x)$$

The exchange of variable theorem

$f(X)=Y$

$E[Y]=\sum_y yP(y) \longrightarrow$

If each $x \to$ each y Bi-ject

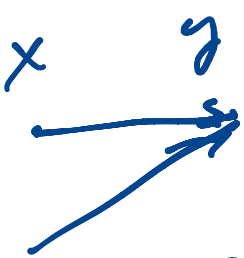
$E[Y]=\sum yP(x)$

$\because P(Y=y)=P(X=x)$

$= \sum_x f(x)P(x)$

if some multiple $x \to$ one y II

single $x \to$ single y I



$E[Y = \sum y\,P(y) = \sum_{y \in I} y\,P(x) + \sum_{y \in II} y\,P(y)$

$y \cdot P(y)$

$= y \cdot \sum P(x) = \sum_x y\,P(x)$

$P(y)$

$= \sum P(x)$

Expected time of cat

- A cat moves with random constant speed **V**, either 5mile/hr or 20mile/hr with equal probability, what's the expected time for it to travel 50 miles?

$$T = \frac{D}{V} = f(V)$$

$$E[T] = \sum_{V} f(v)P(v) \neq \frac{D}{E[V]}$$

$$= \frac{D}{V_1} \cdot P(V_1) + \frac{D}{V_2} \cdot P(V_2)$$

$$= \frac{50}{5} \times \frac{1}{2} + \frac{50}{20} \times \frac{1}{2} = 6.25$$

Q: Is this statement true?

If there exists a constant such that $\boldsymbol{P}(X \geq a) = 1$, then $E[X] \geq a$. It is:

A. True

B. False

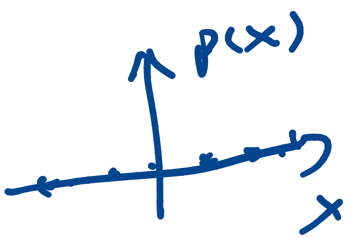
$$E[X] = \sum x\,p(x) = \sum_{x<a} x\,\cancel{p(x)}^{0} + \sum_{x\geq a} x\,p(x)$$

$$\geq \sum_{x\geq a} a\,p(x) = a \underbrace{\cancel{\sum_{x\geq a} p(x)}}^{1}$$

Variance and standard deviation

- The variance of a random variable X is

$f(X) = (X - E[X])^2$



$$\boxed{var[X] = E[(\underline{X - E[X]})^2]}$$

- The standard deviation of a random variable X is

$\{x_i\}$

$$std[X] = \sqrt{var[X]}$$

$var(\{x_i\}) = \frac{\sum^{N} (x_i - \mu)^2}{N}$

Properties of variance

- For random variable X and constant k

$$var[X] \geq 0$$

$$var[kX] = k^2 var[X]$$

A neater expression for variance

- Variance of Random Variable X is defined as:

$Y = (X - E[X])^2$

$= E[f(x)]$

$$var[X] = E[(X - E[X])^2]$$

$= E[Y] = \sum_y y\,p(y)$

$= \sum_x (x - E[x])^2 \cdot p(x)$

- It's the same as:

$$\boxed{var[X] = E[X^2] - E[X]^2}$$

$$X(\omega) = \begin{cases} 1 & \omega = \text{Head} \\ 0 & \omega = \text{tail} \end{cases}$$

$$X - E[X] = \begin{cases} 1 - \frac{1}{2} & \omega = \text{Head} \\ 0 - \frac{1}{2} & \omega = \text{tail} \end{cases}$$

$$(X - E[X])^2 = \begin{cases} \left(\frac{1}{2}\right)^2 = \frac{1}{4} & \omega = H \\ \left(-\frac{1}{2}\right)^2 = \frac{1}{4} & \omega = T \end{cases}$$

$$\cancel{\boxed{X = X + 1}}$$

A neater expression for variance

$$var[X] = E[(X - E[X])^2]$$

A neater expression for variance

$$var[X] = E[(X - E[X])^2]$$

number

$$var[X] = E[(X-\mu)^2] \quad where\ \mu = E[X]$$

$$= E[X^2 - 2\mu X + \mu^2]$$

$$= E[X^2] + E[-2\mu X] + E[\mu^2]$$

$$= E[X^2] - 2\mu E[X] + E[\mu^2]$$

$$= E[X^2] - 2(E[X])^2 + E[E[X]^2]$$

$$E[X]^2$$

A neater expression for variance

$$var[X] = E[(X - E[X])^2]$$

$$var[X] = E[(X - \mu)^2] \quad where\ \mu = E[X]$$

$$= E[X^2 - 2X\mu + \mu^2]$$

$$= E[X^2] - E[X]^2$$

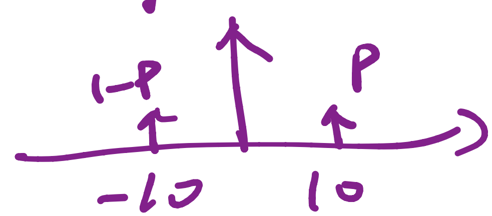
Variance: the profit example

- For the profit example, what is the variance of the return? We know E[X]= 20p-10

$$X = \begin{cases} 10 & s \\ -10 & f \end{cases}$$

$$var[X] = E[X^2] - (E[X])^2$$

$$var[X] = 100 - (20p-10)^2$$



$$E[X^2] = \sum_x x^2 \cdot p(x)$$

$$= 10^2 p(X=10) + (-10)^2 \cdot p(X=-10)$$

$$= 100p + 100(1-p)$$

Motivation for covariance

- Study the relationship between random variables
- Note that it's the un-normalized correlation
- Applications include the fire control of radar, communicating in the presence of noise.

Covariance

- The **covariance** of random variables X and Y is

$$cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

- Note that

$$cov(X, X) = E[(X - E[X])^2] = var[X]$$

A neater form for covariance

- A neater expression for **covariance** (similar derivation as for variance)

$$cov(X,Y) = E[XY] - E[X]E[Y]$$

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

Correlation coefficient is normalized covariance

- The correlation coefficient is

$$corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

- When X, Y takes on values with equal probability to generate data sets $\{(x,y)\}$, the correlation coefficient will be as seen in Chapter 2.

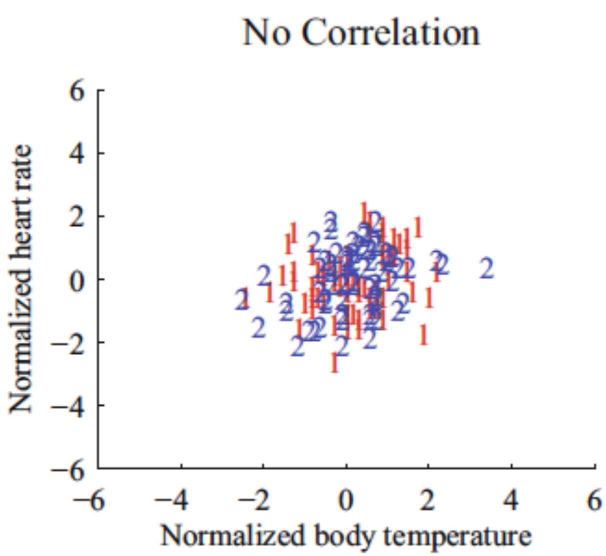
Correlation coefficient is normalized covariance

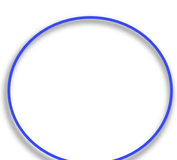
- The correlation coefficient can also be written as:

$$corr(X, Y) = \frac{E[XY] - E[X]E[Y]}{\sigma_X \sigma_Y}$$

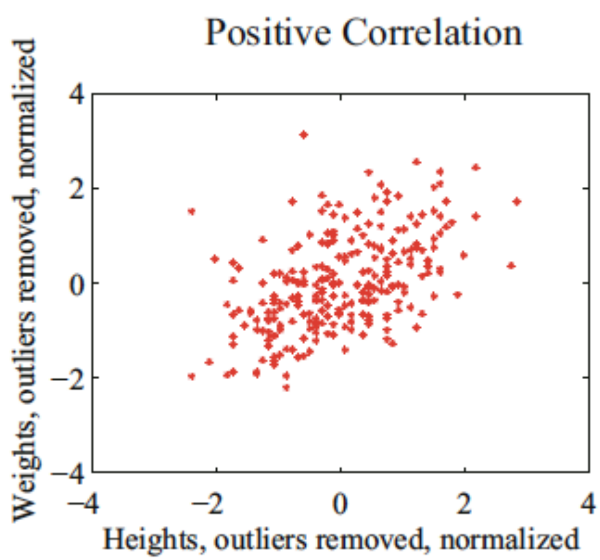
Covariance seen from scatter plots

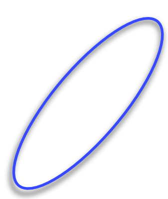
Zero Covariance



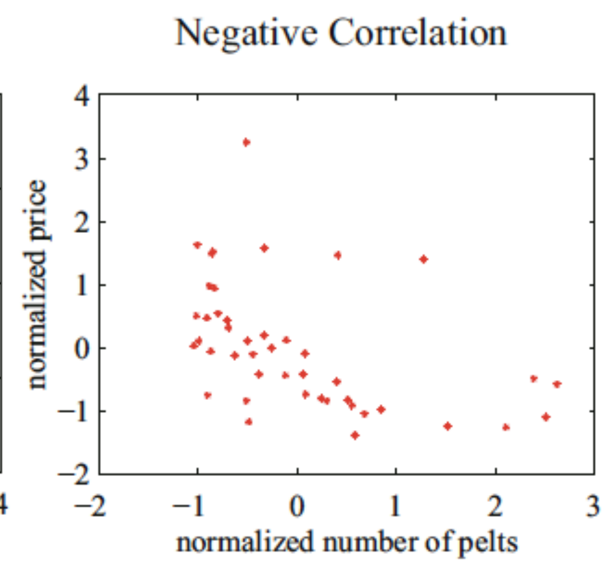


Positive Covariance





Negative Covariance



Credit: Prof.Forsyth

When correlation coefficient or covariance is zero

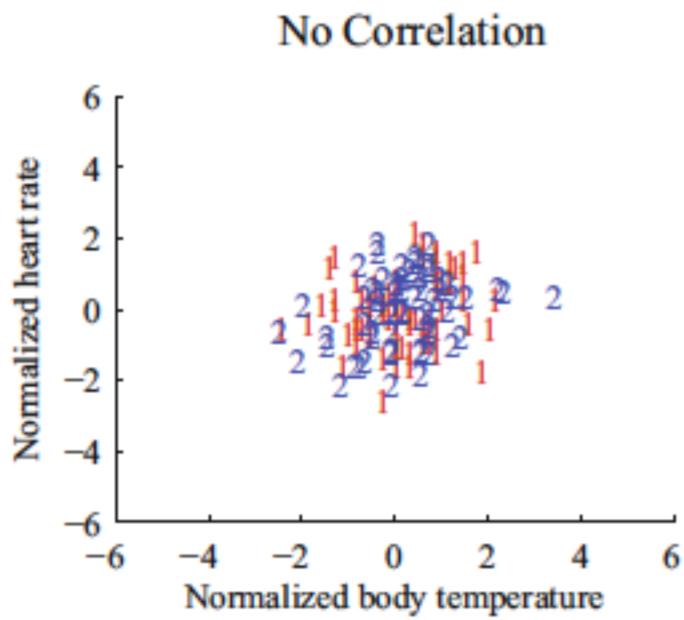
- The covariance is 0!

- That is:

cov(X,Y)

$$=E[XY] - E[X]E[Y] = 0$$

$$E[XY] = E[X]E[Y]$$



- This is a necessary property of independence of random variables * (not equal to independence)

not sufficient

Variance of the sum of two random variables

$$var[X + Y] = var[X] + var[Y] + 2cov(X, Y)$$

Extra pt. in HW

These are equivalent:

(I) $\mathrm{cov}(X, Y) = 0$; $\mathrm{corr}(X, Y) = 0$

(II) $E[XY] = E[X]E[Y]$

(III) $\mathrm{var}[X+Y] = \mathrm{var}[X] + \mathrm{var}[Y]$

uncorrelated!!

Properties of independence in terms of expectations

- $$E[XY] = E[X]E[Y]$$

If X, Y are independent

then $\begin{cases} \text{cov}(X,Y)=0;\ \text{corr}(X,Y)=0 \\ E[XY]=E[X]E[Y] \\ \text{var}[X+Y]=\text{var}[X]+\text{var}[Y] \end{cases}$

Q: What is this expectation?

- We toss two identical coins A & B independently for three times and 4 times respectively, for each head we earn $1, we define X is the earning from A and Y is the earning from B. What is E(XY)?

A. $2 B. $3 C. $4

Uncorrelated vs Independent

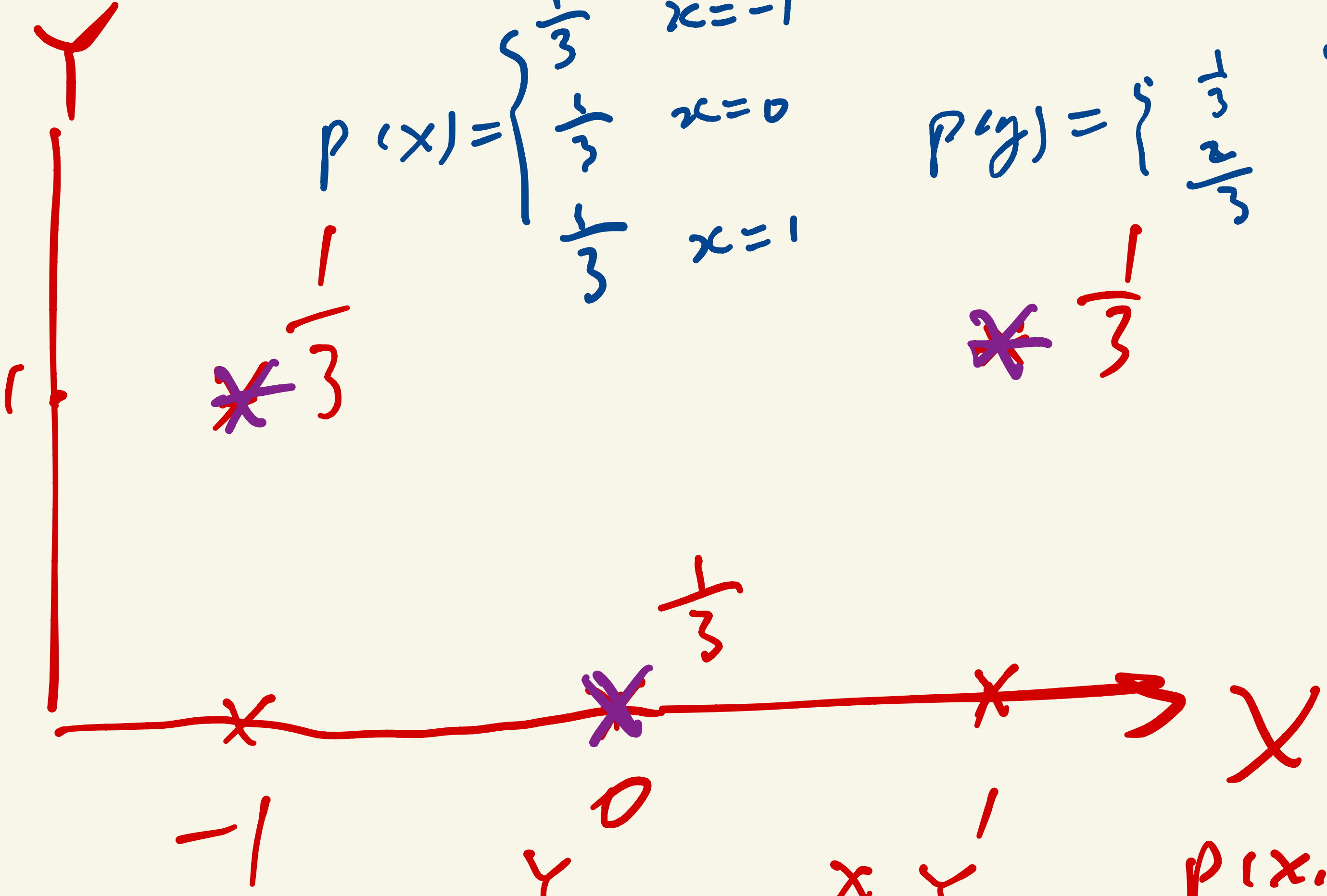
- If two random variables are uncorrelated, does this mean they are independent? Investigate the case X takes -1, 0, 1 with equal probability and $Y=X^2$.

$X \backslash Y$	0	1
-1	0	$\frac{1}{3}$
0	$\frac{1}{3}$	0
1	0	$\frac{1}{3}$

$Y = X^2$

$$p(x) = \begin{cases} \frac{1}{3} & x=-1 \\ \frac{1}{3} & x=0 \\ \frac{1}{3} & x=1 \end{cases}$$

$$p(y) = \begin{cases} \frac{1}{3} & y=0 \\ \frac{2}{3} & y=1 \end{cases}$$



X	Y	XY	$p(x,y)$
1	1	1	$\frac{1}{3}$
-1	1	-1	$\frac{1}{3}$
0	0	0	$\frac{1}{3}$

$E[XY] = 0$

$p(x,y) \neq p(x)p(y)$

i.e. for $x=0$ $y=0$

$LHS = \frac{1}{3}$

$RHS = \frac{1}{9}$

Assignments

- Finish Chapter 4 of the textbook
- Next time: Proof of Chebyshev inequality & Weak law of large numbers, Continuous random variable

Additional References

- Charles M. Grinstead and J. Laurie Snell "Introduction to Probability”
- Morris H. Degroot and Mark J. Schervish "Probability and Statistics”

See you next time

See You!