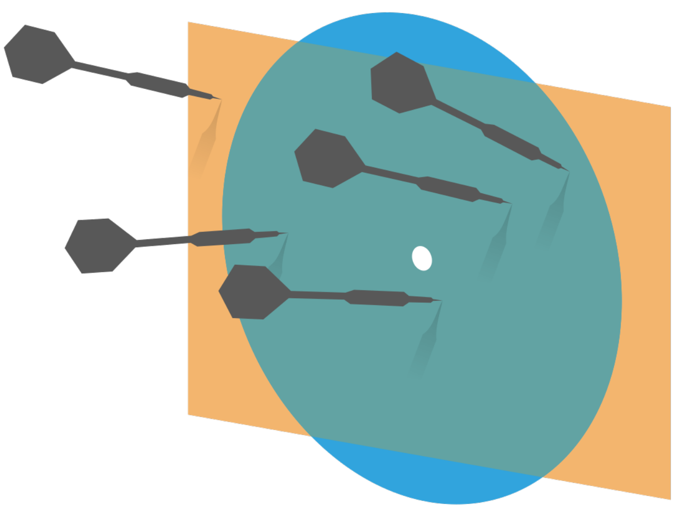
Probability and Statistics for Computer Science



Credit: wikipedia

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

Last time

- Random Variable
- Probability distribution
- Cumulative distribution
- Conditional probability and joint probability

Objectives

- Random Variable
  - *Independence of random variables*
  - *Expected value*
  - *Variance & covariance*

prod for indep. $P(A \cap B) = P(A)P(B)$

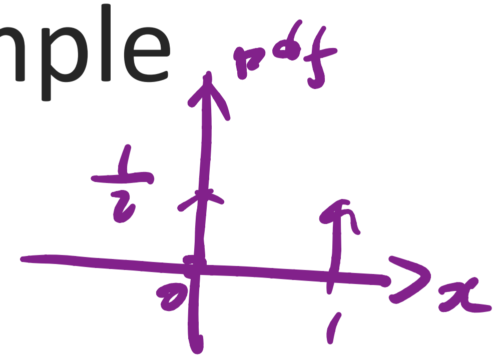
Independence of random variables

- Random variable X *and* Y are independent if

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

  $$\boxed{P(x,y) = P(x)P(y) \quad for\ all\ x\ and\ y}$$

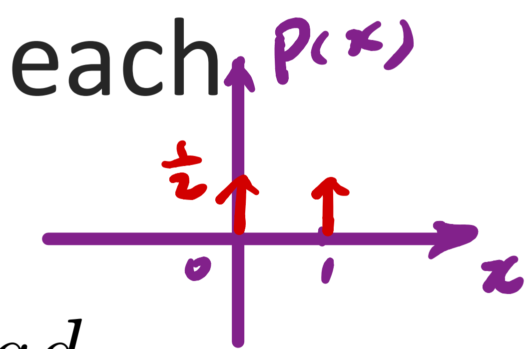
- In the previous coin toss example
  - Are X and Y independent?
  - Are ***S*** and ***D*** independent?



$S = X + Y \quad D = X - Y$

Joint Probability Example

- Tossing a fair coin twice, we define random variable X *and* Y for each toss.



$$X(\omega) = \begin{cases} 1 & outcome\ of\ \omega\ is\ head \\ 0 & outcome\ of\ \omega\ is\ tail \end{cases}$$

$S = X + Y$

$$Y(\omega) = \begin{cases} 1 & outcome\ of\ \omega\ is\ head \\ 0 & outcome\ of\ \omega\ is\ tail \end{cases}$$

$D = X - Y$

Joint probability distribution example

$P(x, y)$

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

$P(y=0) = P(y=0 \cap X=0) + P(y=0 \cap X=1)$

Y \ X	0	1	$P(y)$	
0	$P(x=0, y=0)$ $P(0,0)$ $\frac{1}{4}$	$P(x=1, y=0)$ $P(1,0)$ $\frac{1}{4}$	$\frac{1}{2}$	$= P(y=0)$
1	$P(0,1)$ $\frac{1}{4}$	$P(1,1)$ $\frac{1}{4}$	$\frac{1}{2}$	$= P(y=1)$

	0	1
$P(x)$	$\frac{1}{2}$	$\frac{1}{2}$
	$P(x=0)$	$P(x=1)$

$P(x, y) \overset{\checkmark}{=} P(x)P(y)$

Joint probability distribution example

$P(s,d)$

S \ D	-1	0	1	$P(s)$
0	0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\boxed{0}$	$\frac{1}{4}$	$\boxed{\frac{1}{2}}$
2	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$P(d)$	$\frac{1}{4}$	$\boxed{\frac{1}{2}}$	$\frac{1}{4}$	

$P(s=1 \cap d=0) = 0$

$P(s=1) = \frac{1}{2}$

$P(d=0) = \frac{1}{2}$

$P(s,d) \neq P(s)P(d)$

Joint probability distribution example

$P(s,d)$

S \ D	-1	0	1	$P(s)$
0	0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\boxed{0}$	$\frac{1}{4}$	$\boxed{\frac{1}{2}}$
2	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$P(d)$	$\frac{1}{4}$	$\boxed{\frac{1}{2}}$	$\frac{1}{4}$	

$$P(S=1, D=0) \neq P(S=1)P(D=0)$$

After class

Bayes rule for random variable

- Bayes rule for events generalizes to random variables

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$\boxed{\begin{aligned} P(x|y) &= \frac{P(y|x)P(x)}{P(y)} \\ &= \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)} \end{aligned}}$$

Total Probability

After class

Conditional probability distribution example

$$P(s|d) = \frac{P(s,d)}{P(d)}$$

S \ D	-1	0	1
0	0	$\frac{1}{2}$	0
1	1	0	1
2	0	$\frac{1}{2}$	0

After class

Conditional probability distribution example

$$P(s|d) = \frac{P(s,d)}{P(d)}$$

S \ D	-1	0	1
0	0	$\frac{1}{2}$	0
1	1	0	1
2	0	$\frac{1}{2}$	0

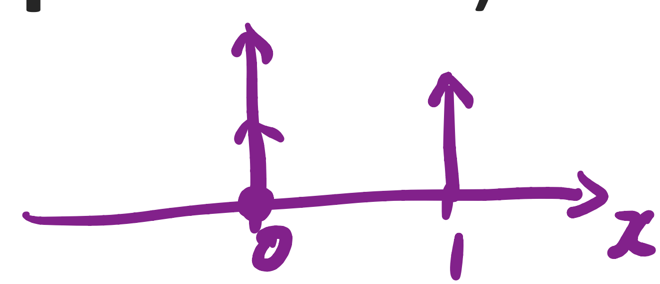
$$P(D=-1|S=1) = \frac{P(S=1|D=-1)P(D=-1)}{P(S=1)} = \frac{1 \times \frac{1}{4}}{\frac{1}{2}}$$

Three important facts of Random variables

- Random variables have **probability functions**
- Random variables can be **conditioned** on events or other random variables
- Random variables have **averages**

Expected value

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



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

$p(X=1) = p$

$p(X=0) = 1-p$

$1 \times p + 0 \times (1-p) = p$

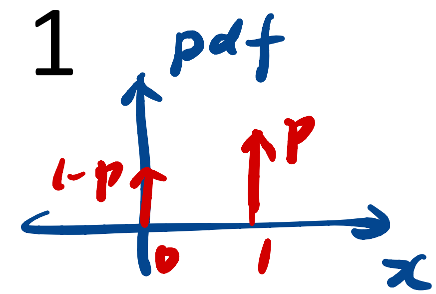
The expected value is a weighted sum of **all** the values X can take

Expected value

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

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

<= 1



E[x] = 0.5 if p = 1/2

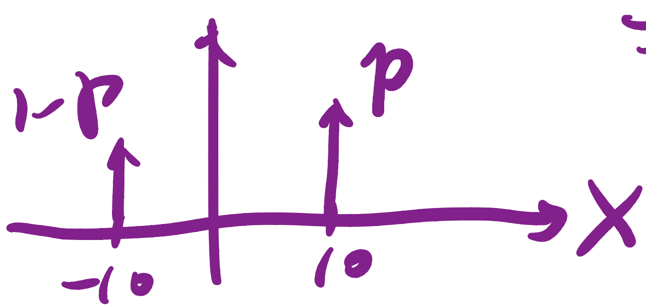
Theoretical mean

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. What is $E[X]$?

$$E[X] = \sum x\,p(x)$$
$$= 10p + (-10)(1-p)$$
$$= 20p - 10$$

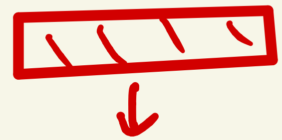


After class

Cookie

$1 each

Chocolate



$2 each

A) random draw 1 out of 4 — Expected value ?

B) random draw 1 <u>twice</u> with replacement
when the two 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 $E[kX + c] = kE[X] + c$

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

Linearity of Expectation

- Proof of the additive property

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

$S = X + Y$

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

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

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

$P(s) = P(S=s)$

Note. $P(S=s) = 0$ if $x+y \neq s$

X, Y, Z are Bernoulli with $p = 0.5$

$E[X + 2Y - Z]$

$= E[X] + 2E[Y] - E[Z]$

$= \cancel{0.5} + 2 \times 0.5 \cancel{- 0.5} = 1$

$E[X] = E[Y] = E[Z]$

$= p \times 1 + (1-p) \cdot 0$

$= p = 0.5$

Proof conti.

$$E[X+Y] = \sum_x \sum_y (x+y)P(x,y)$$

$$E[aX+bY+cZ] = ?$$

$$= \sum_x \sum_y xP(x,y) + \sum_x \sum_y yP(x,y)$$

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

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

$$= E[X] + \sum_y y P(y)$$

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

Q. What's the value?

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

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

Expected value of a function of X

- If $\boldsymbol{f}$ is a function of a random variable X, then $\underline{Y = \boldsymbol{f}(X)}$ is a random variable too $E[f(X)]$

- The expected value of $Y = \boldsymbol{f}(X)$ is $Y = f(X)$

$$E[f(X)] = ?$$

$$\underline{E[Y]} = \sum_y \underline{y P(y)}$$

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 exchang of variable theorem

$$E[f(x)] = E[Y] = \sum y\,P(y)$$

if $\{x\} \to \{y\}$ is bijection $P(y) = P(x)$

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

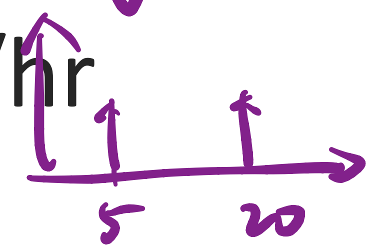
if several x in $\{x_s\}$ → one y_s value

$$E[Y] = \sum_{y \in I} y\,P(x) + y_s\,P(y_s)$$
$$= \sum_{y \in I} y\,P(x) + y_s \sum_{x \in \{x_s\}} P(x)$$
$$= \sum y\,P(x) = \sum_x f(x)\,P(x)$$

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?

V



$E[T] \quad T = \frac{50}{V}$

$E[T] = \sum_{v} \frac{50}{v} \cdot p(v)$

$= \frac{50}{5} \times \frac{1}{2} + \frac{50}{20} \times \frac{1}{2} \neq \frac{50}{E[V]}$

after class

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

Variance and standard deviation

- The variance of a random variable X is

$E[f(X)]$

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

- The standard deviation of a random variable X is

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

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

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:

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

- It's the same as:

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

A neater expression for variance

$\mu = E[X]$

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

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

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

$$E[2X\mu] \stackrel{?}{=} 2\mu\, E[X] \qquad E[\mu^2] = \mu^2$$

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

$$= E[X^2] - 2\mu\, \overset{\mu}{\cancel{E[X]}} + \mu^2$$

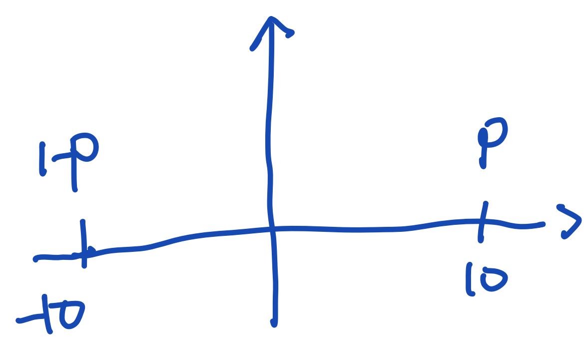
$$E[X^2] - \mu^2$$

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

Variance: the profit example

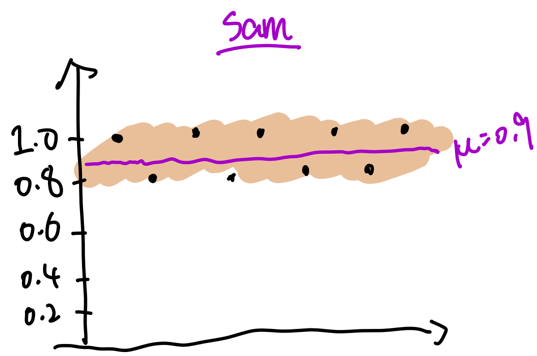
- For the profit example, what is the variance of the return? We know $E[X]$= 20p-10

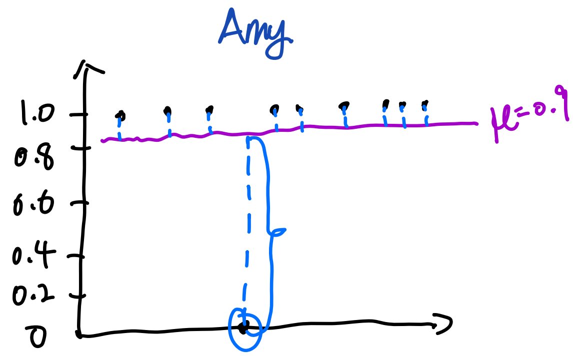
$$var[X] = \boxed{E[X^2]} - \boxed{(E[X])^2}$$



$$= \sum_{\boxed{x}} x^2 \cdot p(x) - (20p-10)^2$$

$$= \boxed{10^2 \cdot p + (-10)^2 (1-p)} - (20p-10)^2$$





$$Var_{sam} \ll Var_{Amy}$$

Motivation for covariance

- Study the relationship between random variables
- Note that it's the un-normalized correlation

  unbounded.

  $[-1, 1]$
- 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

Let $Y = X$

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

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

A neater form for covariance

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

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

$$Var(X) = cov(X,X) = E[X \cdot X] - E[X] \cdot E[X]$$

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

Correlation coefficient is normalized covariance

- The correlation coefficient is

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

- 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.

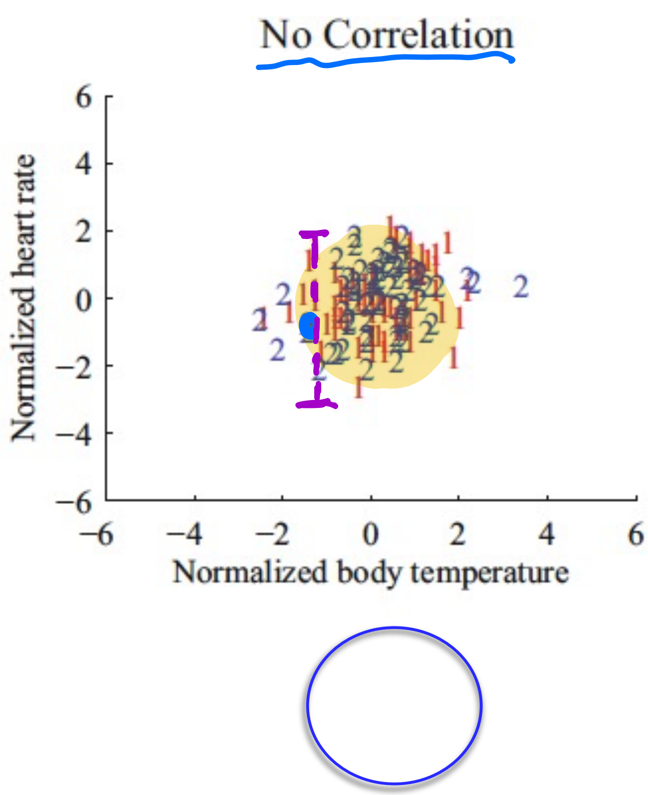
Correlation coefficient is normalized covariance

- The correlation coefficient can also be written as:

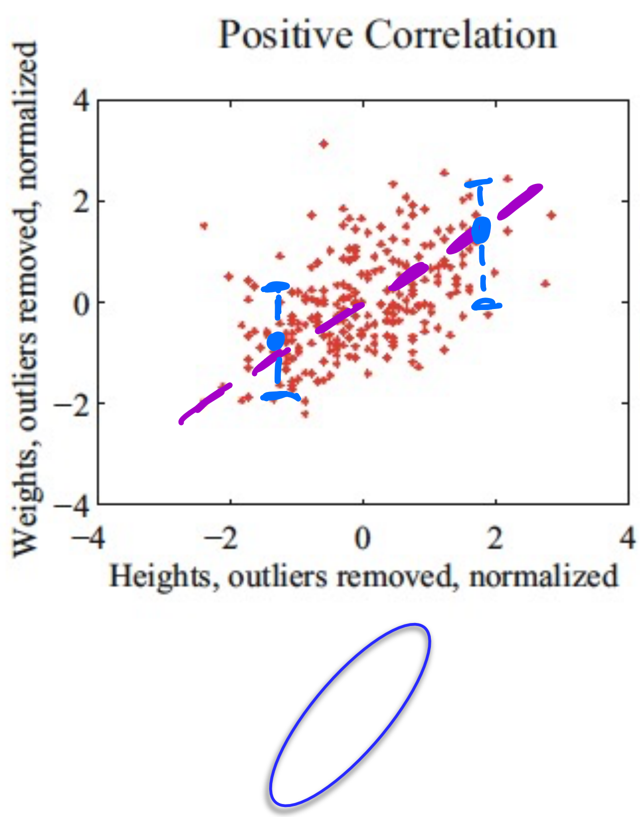
$$corr(X,Y) = \frac{\overbrace{E[XY] - E[X]E[Y]}^{Cov(X,Y)}}{\sigma_X \sigma_Y}$$

Correlation seen from scatter plots

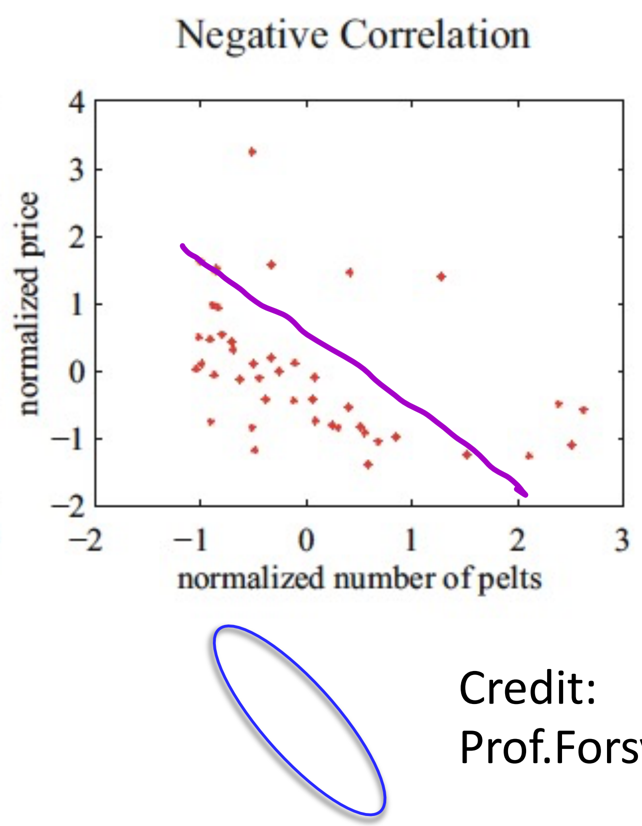
Zero Correlation



Positive correlation



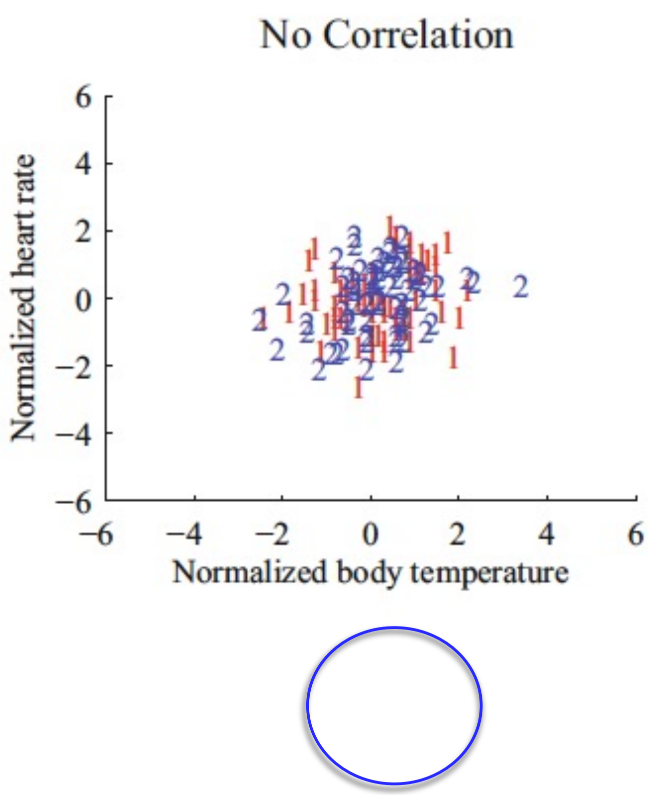
Negative correlation



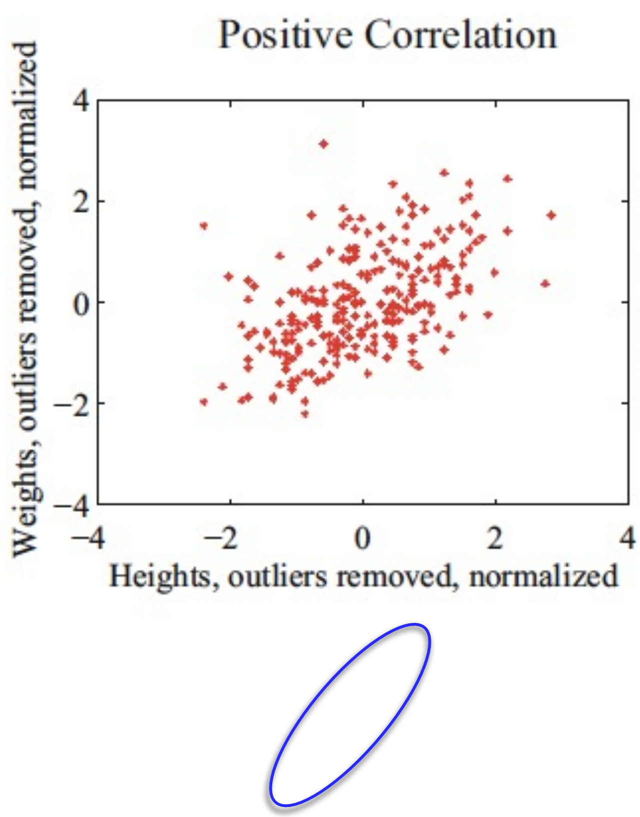
Credit: Prof.Forsyth

Covariance seen from scatter plots

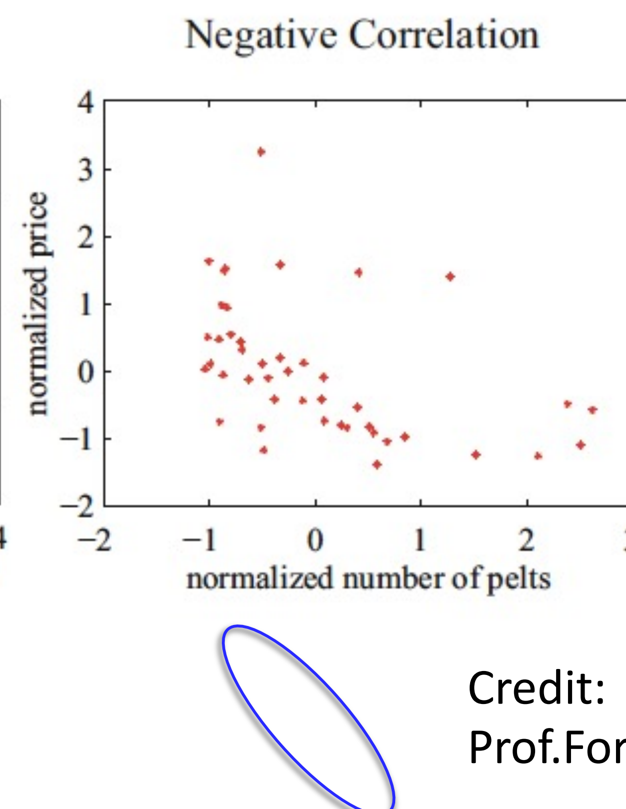
Zero Covariance



Positive Covariance



Negative Covariance



Credit: Prof.Forsyth

When correlation coefficient or covariance is zero

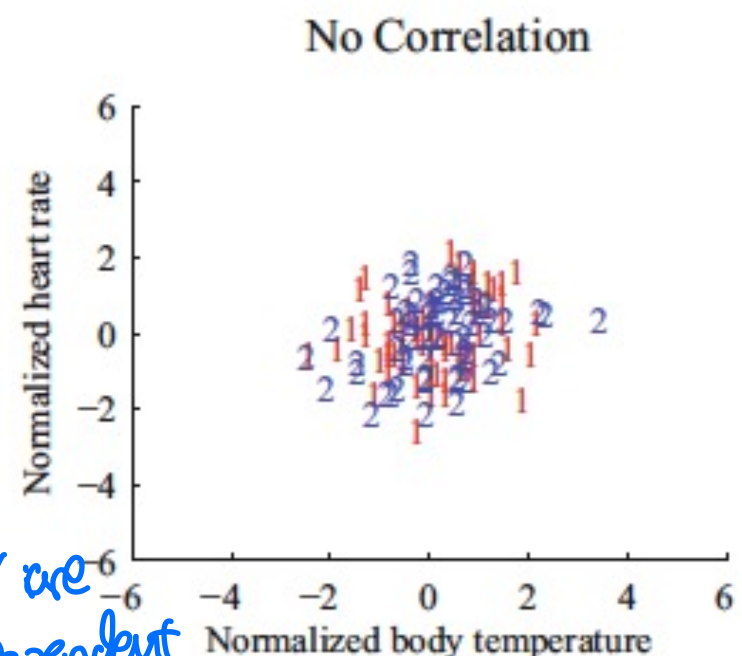
- The covariance is 0!
- That is:

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

$E[XY]-\boxed{E[X]\cdot E[Y]}=0$

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

$$\boxed{E[XY]=E[X]E[Y]}$$ if X, Y are independent



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

Variance of the sum of two random variables

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

If events X & Y are independent, then

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

$$X = \begin{cases} 1 & 0.5 \\ 2 & 0.5 \end{cases}$$

$$\sum_x x \cdot p(x) \cdot \boxed{2} = \boxed{2} \sum_x x \cdot p(x)$$

$$\mathbb{E}[XY] = \sum_y \sum_x xy \cdot \boxed{P(x,y)}$$

If x, y are independent, $P(x,y) = P(x) \cdot P(y) \ \forall x, y$

$$= \sum_y \boxed{\sum_x} \cdot x \cdot \boxed{P(x)} \cdot y \cdot \boxed{P(y)}$$

$$= \boxed{\sum_y y \cdot P(y)} \cdot \boxed{\sum_x x P(x)} = \boxed{\mathbb{E}[Y]} \cdot \boxed{\mathbb{E}[X]}$$

Assignments

- Finish week4 module   Quiz 3 at 4:30pm
- Next time: Markov and 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!