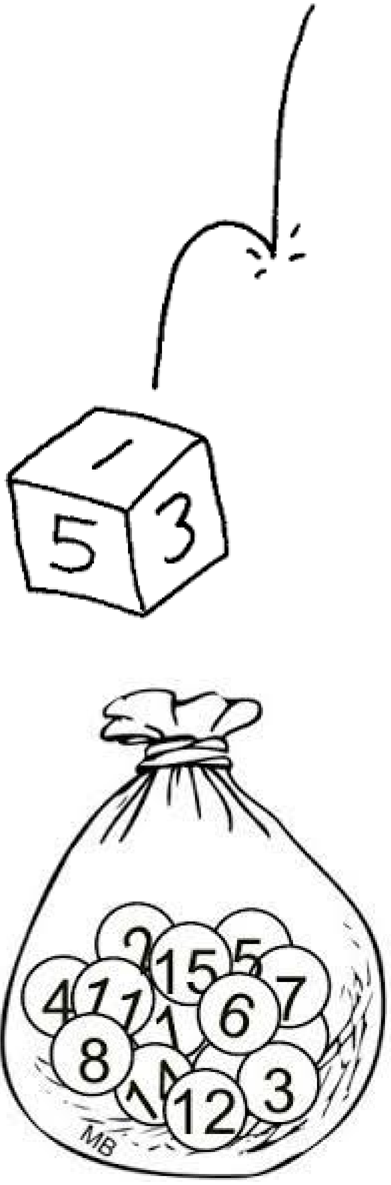
Lecture 12

- Covariance (协方差)
- Correlation Coefficient (相关系数)
- Moment (矩统计量)



Covariance (协方差)

The covariance between X and Y is defined by

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

Two ways to calculate covariance:

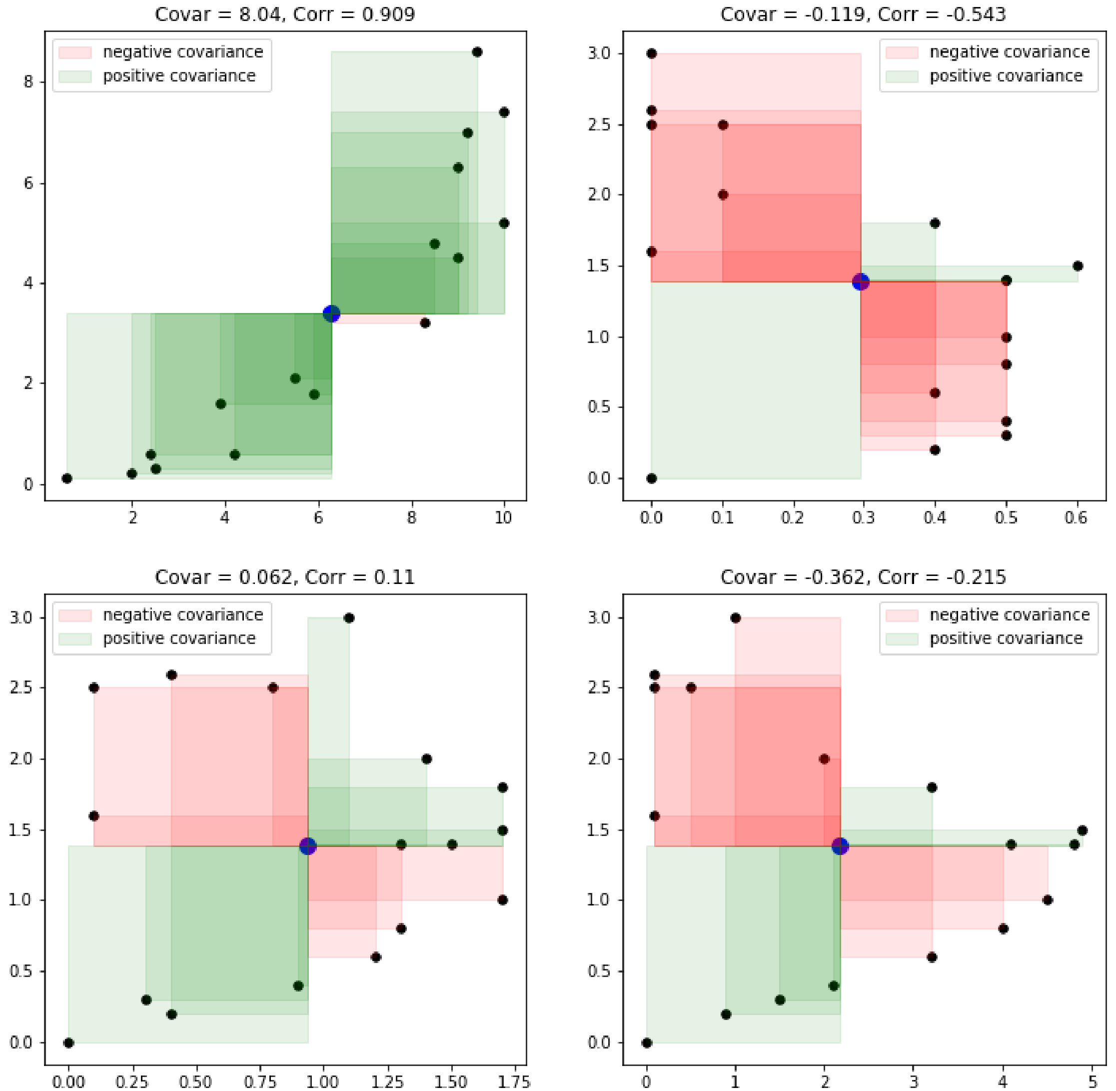
1. $$\begin{aligned}\mathrm{Cov}(X,Y) &= E\{XY - E[X]Y - XE[Y] + E[X]E[Y]\} \\ &= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y]\end{aligned}$$

2. $$D(X+Y) = D(X) + D(Y) + 2 \cdot \mathrm{Cov}(X,Y)$$

$$\Rightarrow \mathrm{Cov}(X,Y) = \frac{1}{2}[D(X+Y) - D(X) - D(Y)]$$

Covariance explained

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



- **Blue point**: $C[E(X), E(Y)]$
- **Sample point**: $P(X,Y)$
- **Edge of square**: $X - E(X)$ & $Y - E(Y)$
- **Area of square**: $|[X-E(X)][Y-E(Y)]|$
- **Color of square**: sign of $[X-E(X)][Y-E(Y)]$

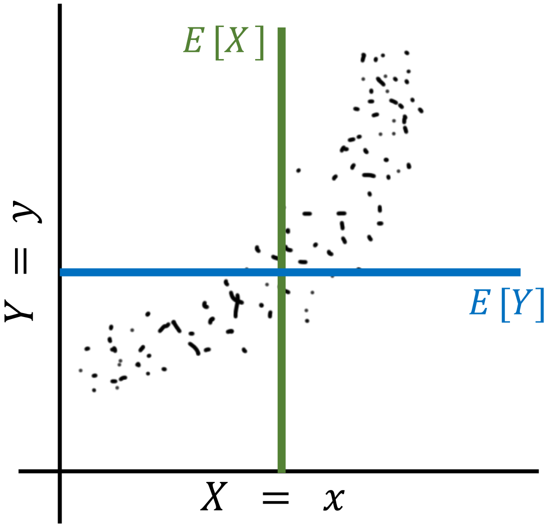
$\Rightarrow$ **Covariance**: $\mathrm{Cov}(X,Y) = E\{[X-E(X)][Y-E(Y)]\} = \dfrac{\text{green area} - \text{red area}}{\text{\# data point}}$

Quick test

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

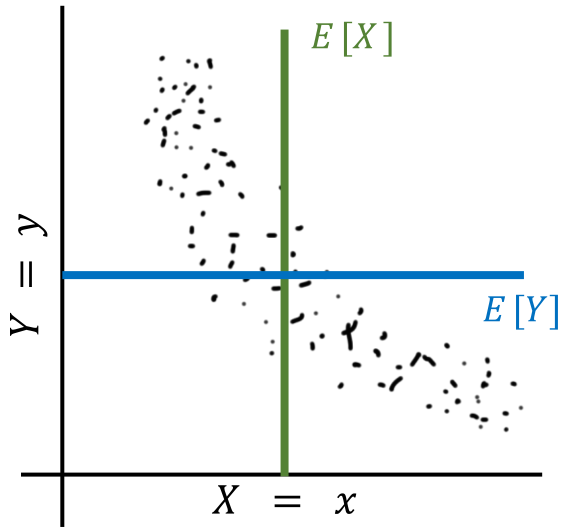
Is the covariance positive, negative, or zero?

1.



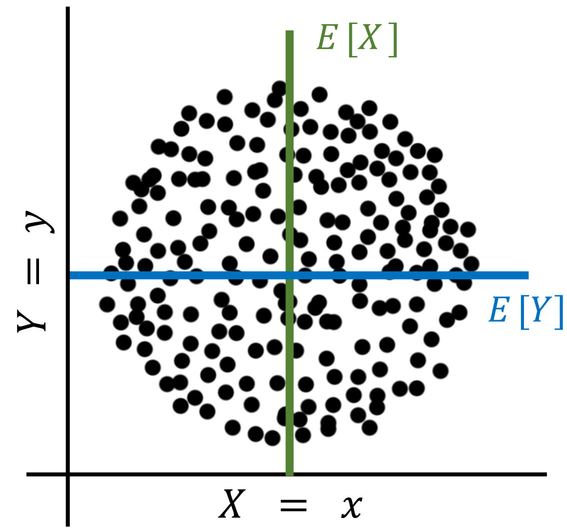
positive

2.



negative

3.



zero

Properties of covariance

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

1. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$, exchangeable, symmetry

2. $\mathrm{Cov}(X,X) = D(X)$, reduced to variance

3. $\mathrm{Cov}(aX,Y) = a\cdot\mathrm{Cov}(X,Y)$ multiplicative rule
   $\mathrm{Cov}(aX+b,Y) = a\cdot\mathrm{Cov}(X,Y)$, non-linearity

4. $\mathrm{Cov}(X_1+X_2,Y) = \mathrm{Cov}(X_1,Y) + \mathrm{Cov}(X_2,Y)$, summation

5. If X and Y are independent,

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

Ex. Given R.V.s (X, Y) follows joint PDF

$$f(x,y) = \begin{cases} 8xy, & 0 \le x \le y \le 1, \\ 0, & \text{otherwise} \end{cases}$$

Find $\mathrm{Cov}(X, Y)$. Hint: $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$.

Ex. Given R.V.s (X, Y) follows joint PDF

$$f(x,y) = \begin{cases} 8xy, & 0 \le x \le y \le 1, \\ 0, & \text{otherwise} \end{cases}$$

Find $\mathrm{Cov}(X,Y)$.　　Hint: $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$.

Sol.

$$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy = \int_0^1 dx \int_x^1 x \cdot 8xydy = \frac{8}{15}$$

$$E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy = \int_0^1 dx \int_x^1 y \cdot 8xydy = \frac{4}{5}$$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_0^1 dx \int_x^1 xy \cdot 8xydy = \frac{4}{9}$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{4}{9} - \frac{4}{5}\cdot\frac{8}{15} = \frac{4}{225}$$

Zero covariance does not imply independence

Let X take on values $\{-1, 0, 1\}$ with equal probability 1/3.

Define $Y = \begin{cases} 1 & \text{if } X = 0 \\ 0 & \text{otherwise} \end{cases}$

Y \ X	-1	0	1	
0	1/3	0	1/3	2/3
1	0	1/3	0	1/3
	1/3	1/3	1/3	

Marginal PMF of Y, $p_Y(y)$

Marginal PMF of X, $p_X(x)$

Show that $\text{Cov}(X, Y) = 0$
but
$P\{Y = y_0 | X = x_0\} \neq P(Y = y_0)$

$$E[X] = \frac{1}{3} \times (-1 + 0 + 1) = 0$$
$$E[Y] = \frac{1}{3} \times 1 + \frac{2}{3} \times 0 = \frac{1}{3}$$
$$E[XY] = \frac{1}{3} \times (-1 \cdot 0 + 0 \cdot 1 + 1 \cdot 0) = 0$$
$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0$$

$\Longrightarrow$ X, Y are with zero covariance.

$P\{Y = 0 | X = 1\} = 1, P\{Y = 0 | X = 0\} = 0$
$\neq$
$P(Y = 0) = 2/3$

$\Longrightarrow$ X, Y are not independent.

Covariance matrix

Def. The covariance matrix of a n-dimensional R.V. $(X_1, X_2, \dots, X_n)$ can be written as

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

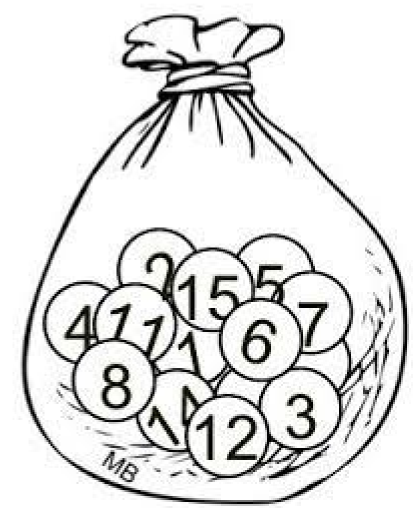
where $c_{ij} = \mathrm{Cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}$,

$$i, j = 1,2, \dots, n.$$

The element $c_{ij} = c_{ji}$ indicates the covariance between the i-th and j-th dimensions of $(X_1, X_2, \dots, X_n)$.

Lecture 12

- Covariance (协方差)
- Correlation Coefficient (相关系数)
- Moment (矩统计量)



Correlation Coefficient (相关系数)

The correlation coefficient between X and Y is defined by

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{E\{[X-E(X)][Y-E(Y)]\}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

where $D(X) > 0, D(Y) > 0$.

Note:

1. ρ_{XY} only measures the linear correlation;
2. $-1 \leq \rho_{XY} \leq 1$;
3. $|\rho_{xy}| = 1 \Leftrightarrow P\{Y = a + bX\} = 1$;
4. X and Y are independent $\Rightarrow X, Y$ are uncorrelated ($\rho_{XY} = 0$);
5. X, Y are uncorrelated ($\rho_{XY} = 0$) $\Leftrightarrow \mathrm{Cov}(X,Y) = 0$
$$\Leftrightarrow E(XY) = E(X)E(Y)$$
$$\Leftrightarrow D(X+Y) = D(X) + D(Y)$$

Correlation Coefficient explained

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$ (Given $D(X) = \sigma_X, D(Y) = \sigma_Y$ are known constant.)

$$= \frac{E\{[X-E(X)][Y-E(Y)]\}}{\sigma_X \sigma_Y} = E\left\{\left[\frac{X-E(X)}{\sigma_X}\right]\left[\frac{Y-E(Y)}{\sigma_Y}\right]\right\} = \mathrm{Cov}(X^*, Y^*)$$

(The covariance of normalized variable X^* and Y^*)

Moreover:

$$\rho_{XY} = \frac{E\{[X-E(X)][Y-E(Y)]\}}{\sqrt{E\{[X-E(X)]^2\}}\sqrt{E\{[Y-E(Y)]^2\}}} = \pm\sqrt{\frac{E\{[X-E(X)][Y-E(Y)]\}^2}{E\{[X-E(X)]^2\}E\{[Y-E(Y)]^2\}}}$$

If $X - E(X)$ and $Y - E(Y)$ are with the same sign, $\rho_{XY} = 1$;

If $X - E(X)$ and $E(Y) - Y$ are with the opposite sign, $\rho_{XY} = -1$;

Otherwise, $-1 < \rho_{XY} < 1$.

Covariance vs. Correlation

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

	Covariance	Correlation
Definition	Measure of linear correlation	Scaled version of covariance
Values	$(-\infty, +\infty)$	$[-1, 1]$
Change in scale	Affects covariance	Does not affect correlation
Unit-free measure	No	Yes

Differ only in scale!

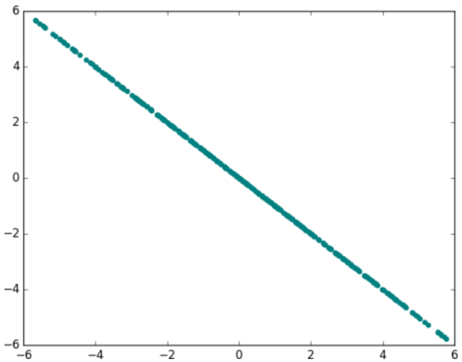
https://www.wallstreetmojo.com/correlation-vs-covariance/

Quick test

What is the correlation coefficient ρ_{XY}?

A. $\rho_{XY} = 1$
B. $\rho_{XY} = -1$
C. $\rho_{XY} = 0$
D. Other

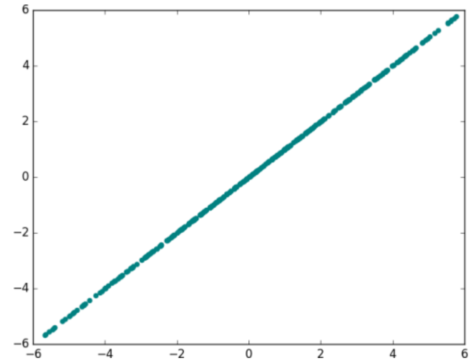
1.



B

$Y = aX + b,\ a < 0$

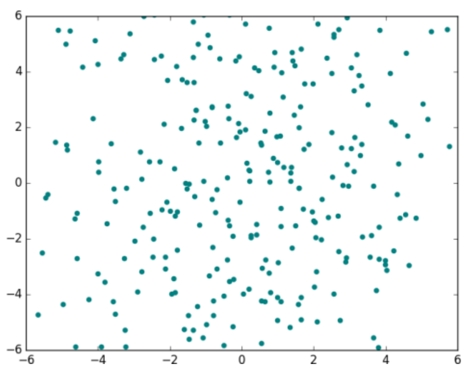
2.



A

$Y = aX + b, a > 0$

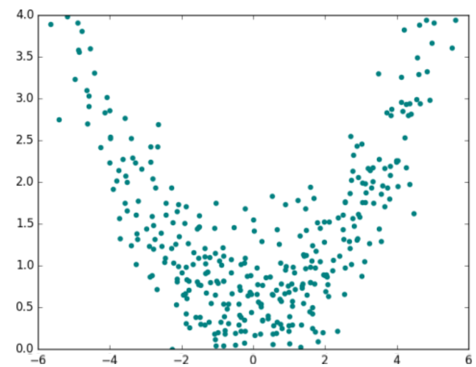
3.



C

uniformly distributed

4.



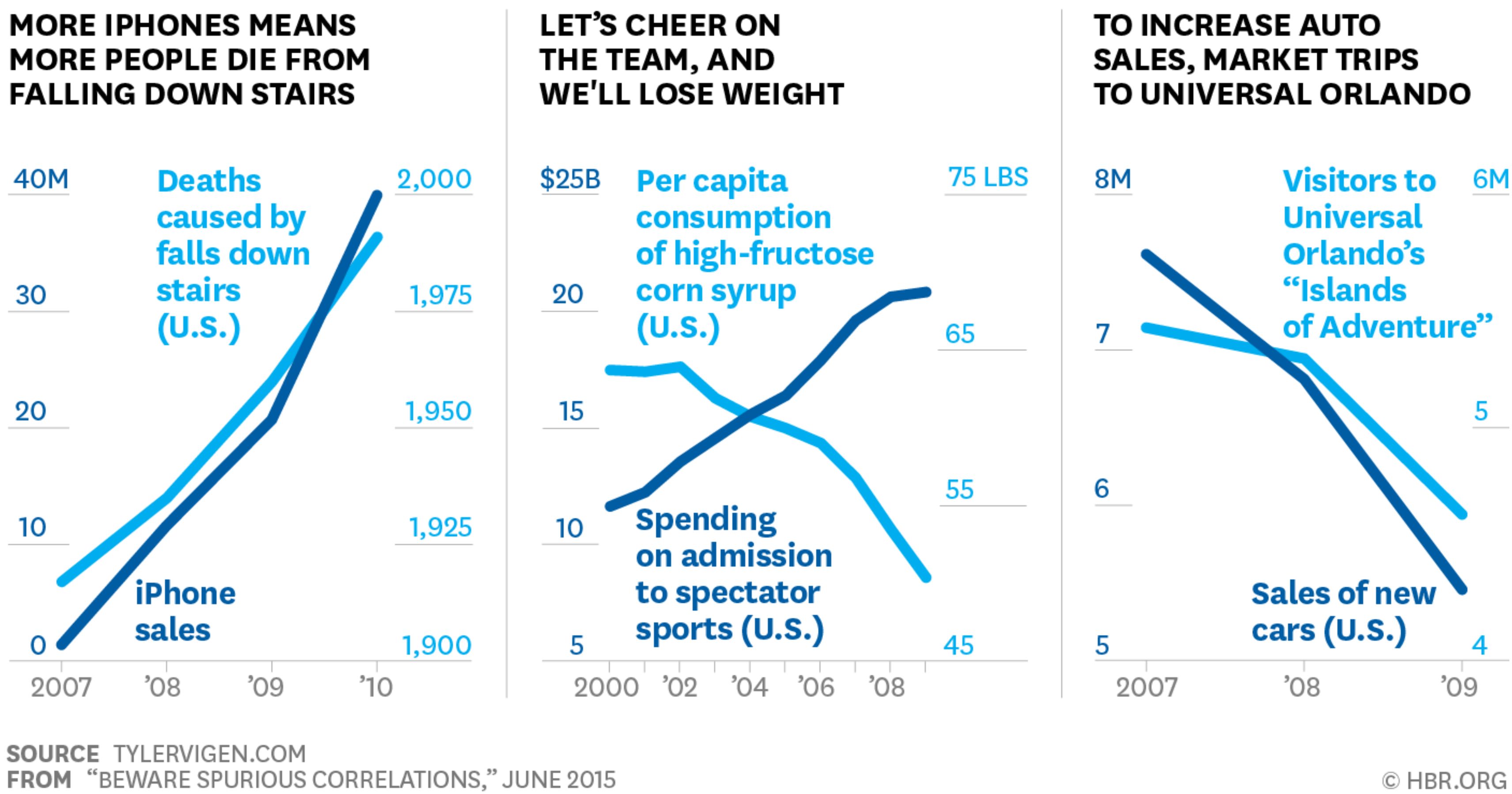
C

$Y = X^2$

nonlinearly related even $\rho_{XY} = 0$.

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y]$$
$$= E[X^3] - E[X]E[X^2] = 0 - 0 \cdot E[X^2] = 0.$$

Beware Spurious Correlations



https://hbr.org/2015/06/beware-spurious-correlations

Ex. Given $X \sim \mathcal{N}(1, 3^2), Y \sim \mathcal{N}(0, 4^2), \rho_{XY} = -\frac{1}{2}$, and $Z = \frac{X}{3} - \frac{Y}{2}$, find $D(Z), \rho_{XZ}$.

$$D(Z) = D\left(\frac{X}{3} - \frac{Y}{2}\right) = D\left(\frac{X}{3}\right) - 2\mathrm{Cov}\left(\frac{X}{3}, \frac{Y}{2}\right) + D\left(\frac{Y}{2}\right)$$

$$\rho_{XZ} = \frac{\mathrm{Cov}(X, Z)}{\sqrt{D(X)}\sqrt{D(Z)}}$$

Ex. If $\rho_{XY} = 0.5$, $E(X) = E(Y) = 0$, $E(X^2) = E(Y^2) = 2$, find $E[(X+Y)^2]$.

$$E[(X+Y)^2] = E(X^2) + E(Y^2) + \mathbf{2E(XY)}$$

Ex. Given $X \sim \mathcal{N}(1, 3^2), Y \sim \mathcal{N}(0, 4^2), \rho_{XY} = -\frac{1}{2}$, and $Z = \frac{X}{3} - \frac{Y}{2}$, find $D(Z), \rho_{XZ}$.

Sol.

$$
\begin{aligned}
D(Z) = D\left(\frac{X}{3} - \frac{Y}{2}\right) &= D\left(\frac{X}{3}\right) - 2\mathrm{Cov}\left(\frac{X}{3}, \frac{Y}{2}\right) + D\left(\frac{Y}{2}\right) \\
&= \frac{1}{9} D(X) - 2 \cdot \frac{1}{2} \cdot \frac{1}{3} \mathrm{Cov}(X, Y) + \frac{1}{4} D(Y) \\
&= \frac{1}{9} D(X) - \frac{1}{3} \rho_{XY} \sqrt{D(X)} \sqrt{D(Y)} + \frac{1}{4} D(Y) \\
&= \frac{1}{9} \cdot 9 - \frac{1}{3}\left(-\frac{1}{2}\right) \sqrt{9} \sqrt{16} + \frac{1}{4} \cdot 16 = 7
\end{aligned}
$$

$$
\begin{aligned}
\rho_{XZ} &= \frac{\mathrm{Cov}(X, Z)}{\sqrt{D(X)}\sqrt{D(Z)}} = \frac{1}{3 \cdot \sqrt{7}} \mathrm{Cov}\left(X, \frac{X}{3} - \frac{Y}{2}\right) \\
&= \frac{1}{3 \cdot \sqrt{7}} \left[\mathrm{Cov}\left(X, \frac{X}{3}\right) - \mathrm{Cov}\left(X, \frac{Y}{2}\right)\right] = \frac{1}{3 \cdot \sqrt{7}} \left[\frac{1}{3} \mathrm{Cov}(X, X) - \frac{1}{2} \mathrm{Cov}(X, Y)\right] \\
&= \frac{1}{3 \cdot \sqrt{7}} \left[\frac{1}{3} \cdot 9 - \frac{1}{2}\left(-\frac{1}{2}\right) \cdot 3 \cdot 4\right] = \frac{2}{\sqrt{7}}
\end{aligned}
$$

Ex. If $\rho_{XY} = 0.5$, $E(X) = E(Y) = 0$, $E(X^2) = E(Y^2) = 2$, find $E[(X+Y)^2]$.

$E[(X+Y)^2] = E(X^2) + E(Y^2) + 2E(XY)$

Sol. $D(X) = E(X^2) - E(X)^2 = 2$.

Similarly, $D(Y) = 2$.

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{D(X)\cdot D(Y)}} = \frac{E(XY)}{2} = 0.5$$

$\Rightarrow E(XY) = 1$

$$E[(X+Y)^2] = E(X^2) + E(Y^2) + 2E(XY) = 2+2+2 = 6$$

Uncorrelated but not independent

X and Y are independent $\Rightarrow X, Y$ are uncorrelated $(\rho_{XY} = 0)$
X and Y are independent $\not\Leftarrow X, Y$ are uncorrelated $(\rho_{XY} = 0)$

Is this possible?

Ex. Let (X, Y) follow joint uniform distribution in a circle with radius R and center $(0,0)$. Show that X and Y are uncorrelated, but X and Y are not independent.

Sol.

$$f(x,y) = \begin{cases} \dfrac{1}{\pi R^2}, & x^2 + y^2 \leq R^2 \\ 0, & \text{otherwise} \end{cases}$$

1. Show that X, Y are uncorrelated, i.e., $\rho_{XY} = 0$

$$E(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x,y) dx dy = \int_{-R}^{+R} dy \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} x \frac{1}{\pi R^2} dx = 0$$

$$E(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x,y) dx dy = \int_{-R}^{+R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} y \frac{1}{\pi R^2} dy = 0$$

$$E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy \cdot f(x,y) dx dy = \int_{-R}^{+R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} xy \cdot \frac{1}{\pi R^2} dy = 0$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0 \quad \Rightarrow \quad \rho_{XY} = 0$$

$\Rightarrow X, Y$ are uncorrelated.

2. Show that X, Y are not independent, i.e., $f(x,y) \neq f_X(x) f_Y(y)$

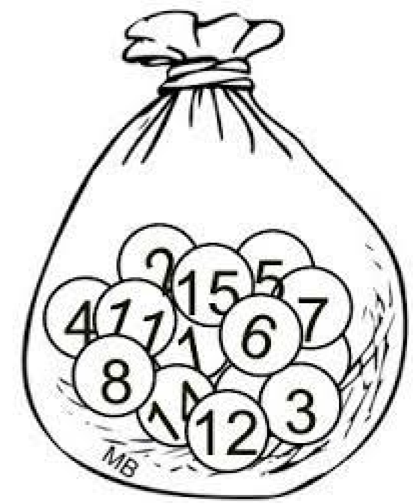
$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\, dy = \begin{cases} \dfrac{2\sqrt{R^2 - x^2}}{\pi R^2}, & -R \leq x \leq R \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\, dx = \begin{cases} \dfrac{2\sqrt{R^2 - y^2}}{\pi R^2}, & -R \leq y \leq R \\ 0, & \text{otherwise} \end{cases}$$

$\Rightarrow f(x,y) \neq f_X(x) f_Y(y) \;\; \Rightarrow \;\; X, Y$ are not independent

Lecture 12

- Covariance (协方差)
- Correlation Coefficient (相关系数)
- **Moment (矩统计量)**



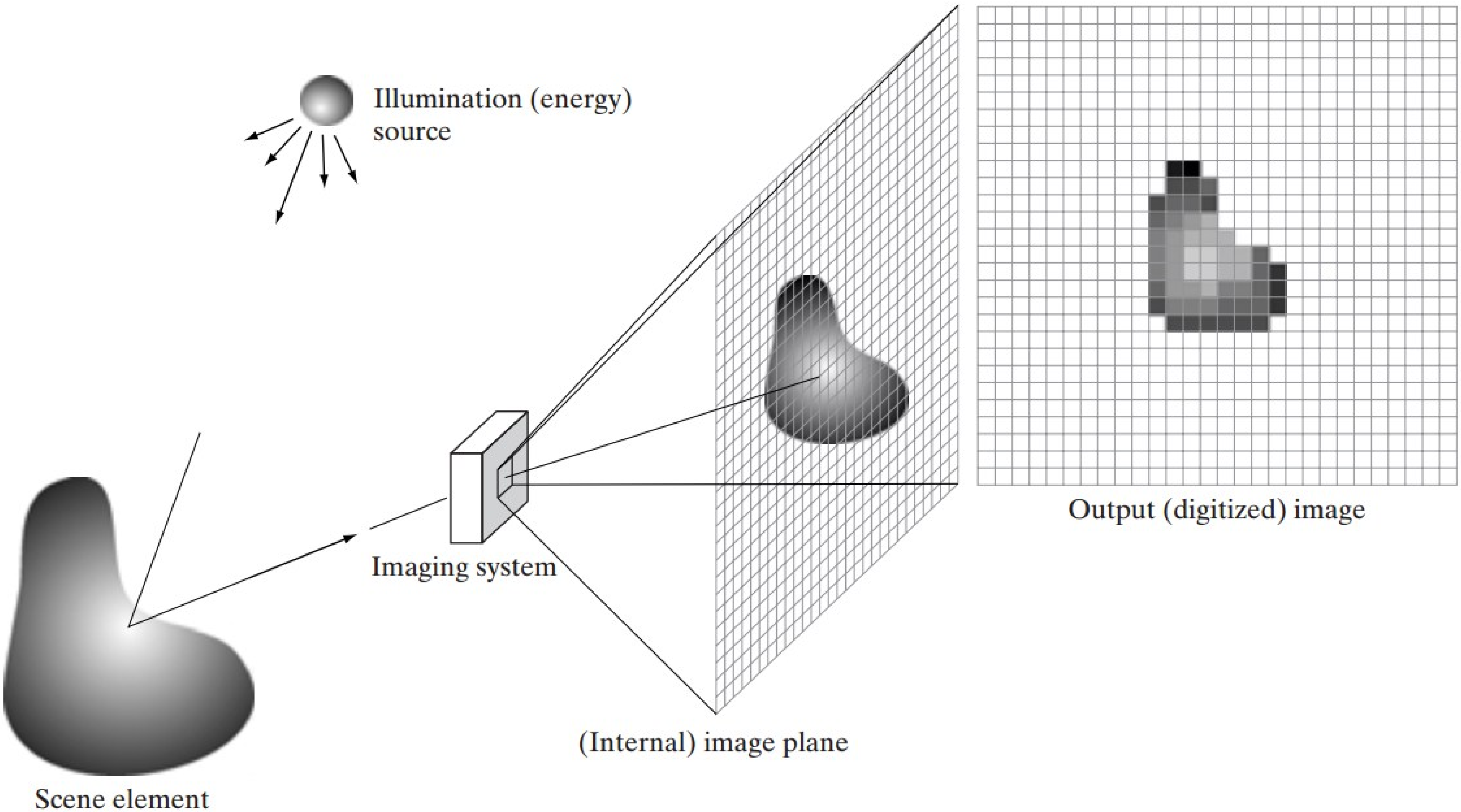
Moment（矩）

Def. Let X, Y be two R.V.s, and k, l be two positive integers,

$E(X^k)$	the k-th order raw moment (原点矩)
$E\{[X-E(X)]^k\}$	the k-th order central moment (中心矩)
$E(X^kY^l)$	the k+l-th order mixed (混合) raw moment
$E\{[X-E(X)]^k[Y-E(Y)]^l\}$	the k+l-th order mixed central moment

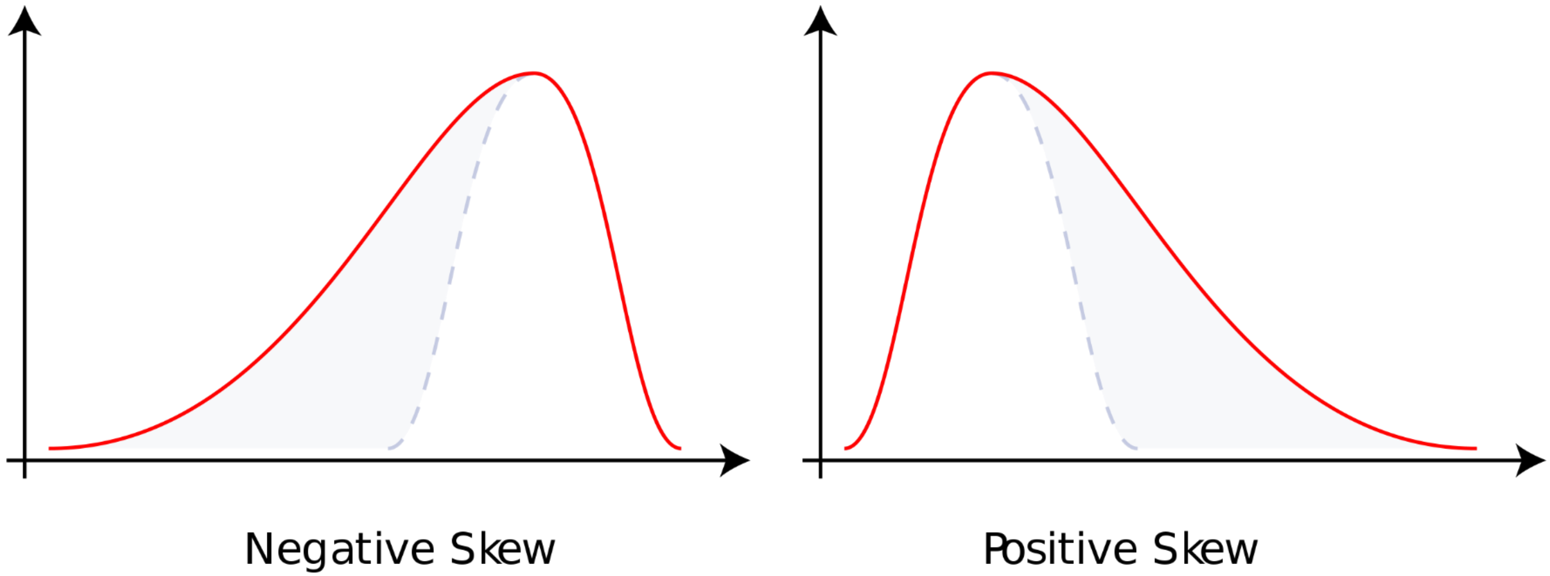
1. $E(X)$ is the first order raw moment of X.
2. $D(X)$ is the second order central moment.
3. $\mathrm{Cov}(X, Y)$ is the second order mixed moment of X and Y.
4. useful in moment estimator (to be covered in a few weeks).

Moment in a digital image



Moment ordinal	Moment		
	Raw	Central	Standardized
1	Mean	0	0
2	–	Variance	1
3	–	–	Skewness
4	–	–	(Non-excess or historical) kurtosis
5	–	–	Hyperskewness
6	–	–	Hypertailedness
7+	–	–	–

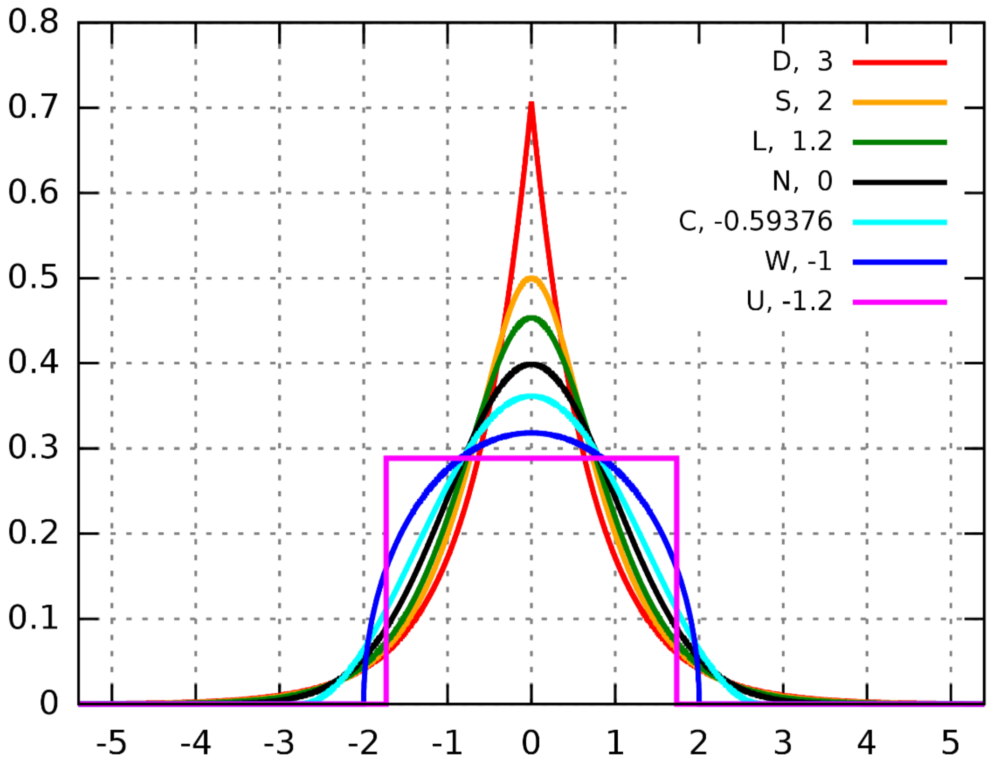
Significance of moments, Wikipedia



$$\text{skew}(X) = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean.

[Skewness, Wikipedia](https://en.wikipedia.org/wiki/Skewness)



$$\mathrm{Kurt}(X) = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$

Kurtosis is a measure of the **tailedness** of the prob. distribution of a real-valued random variable.

- D: **Laplace distribution**, also known as the double exponential distribution, red curve (two straight lines in the log-scale plot), excess kurtosis = 3
- S: hyperbolic secant distribution, orange curve, excess kurtosis = 2
- L: logistic distribution, green curve, excess kurtosis = 1.2
- N: **normal distribution**, black curve (inverted parabola in the log-scale plot), excess kurtosis = 0
- C: raised cosine distribution, cyan curve, excess kurtosis = −0.593762...
- W: Wigner semicircle distribution, blue curve, excess kurtosis = −1
- U: **uniform distribution**, magenta curve (shown for clarity as a rectangle in both images), excess kurtosis = −1.2.

[Kurtosis, Wikipedia]