Homework 1

Simple Linear Regression

1

- **(a)** One of the target variable could be the amount of assets
- **(b)** It’s continuous
- **(c)** The grades candidates provideed in the material
- **(d)** I don’t think it would be a linear model, instead it could be a simulated by a curve

2

In [16]:
```
import numpy as np
from scipy.optimize import leastsq
import matplotlib.pyplot as plt
Xi=np.array([0,1,2,3,4])
Yi=np.array([0,2,3,8,17])
```

In [39]:
```
xbar=np.mean(Xi)
ybar=np.mean(Yi)
print ('The answer of question (a) is')
print('xbar={0:1.2f},y={1:1.2f}'.format(xbar,ybar))
```

```
The answer of question (a) is
xbar=2.00,y=6.00
```

In [40]:
```
syy = np.mean((Yi-ybar)**2)
syx = np.mean((Yi-ybar)*(Xi-xbar))
sxx = np.mean((Xi-xbar)**2)
print ('The answer of question (b) is')
print ('Syy={0:1.2f},Syx={1:1.2f},Sxx={2:1.2f}'.format(syy,syx,sxx))
```

```
The answer of question (b) is
Syy=37.20,Syx=8.00,Sxx=2.00
```

In [41]:
```
beta1 = syx/sxx
beta0 = ybar - beta1*xbar
print('The answer of question (c) is')
print('beta1={0:1.2f},beta0={1:1.2f}'.format(beta1,beta0))
```

```
The answer of question (c) is
beta1=4.00,beta0=-2.00
```

In [35]:
```
x=2.5
y=beta1*x+beta0
print('The answer of question (d) is')
print('y=',y)
```

```
The answer of question (d) is
y= 8.0
```

3. (a) $z(t) = z_0 e^{-at}$

Let $y(t) = \ln(z(t))$.

$\therefore$ $y(t) = \ln z_0 - at$. which is a linear model.

4. (b). $y(t) = \ln z(t)$

$\bar{t} = \frac{1}{N}\sum t_i \qquad \bar{y} = \frac{1}{N}\sum y_i$

$S_t^2 = \frac{1}{N}\sum(t_i - \bar{t})^2, \quad S_y^2 = \frac{1}{N}\sum(y_i - \bar{y})^2.$

$S_{ty} = \frac{1}{N}\sum(y_i - \bar{y})(t_i - \bar{t}).$

$\beta_1 = S_{ty}/S_t. \qquad \beta_0 = \bar{y} - \beta_1 \bar{t}.$

$\Rightarrow \beta_1 = -a. \qquad \beta_0 = \ln z_0.$

$z_0 = e^{\beta_0} \qquad a = -\beta_1.$

In [ ]:

```
x=t
y=np.log(z)
xm = np.mean(x)
ym = np.mean(y)
syy = np.mean((y-ym)**2)
syx = np.mean((y-ym)*(x-xm))
sxx = np.mean((x-xm)**2)
beta1 = syx/sxx
beta0 = ym - beta1*xm
alpha=-b1
z0=exp(b0)
```

4. (a) $RSS(\beta) = \sum_{i=1}^{N} (y_i - \beta x_i)^2$.

(b) $\frac{dRSS(\beta)}{d\beta} = \sum_{i=1}^{N} 2(y_i - \beta x_i)(-x_i) = 0$

$$\sum_{i=1}^{N} -2y_i x_i + 2\beta x_i^2 = 0.$$

$$\beta \sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} y_i x_i$$

$$\beta = \frac{\sum_i x_i y_i}{\sum_i x_i^2}.$$