Introduction to Biocomputing Tutorial

Week 9

Debartolo 319

Review: Estimating statistical model parameters

- Linear $$y = \beta_0 + \beta_1 x + \epsilon$$

- x – independent variable
- y – dependent variable
- β – linear deterministic components of the regression
- ϵ - error

Review: Estimating statistical model parameters

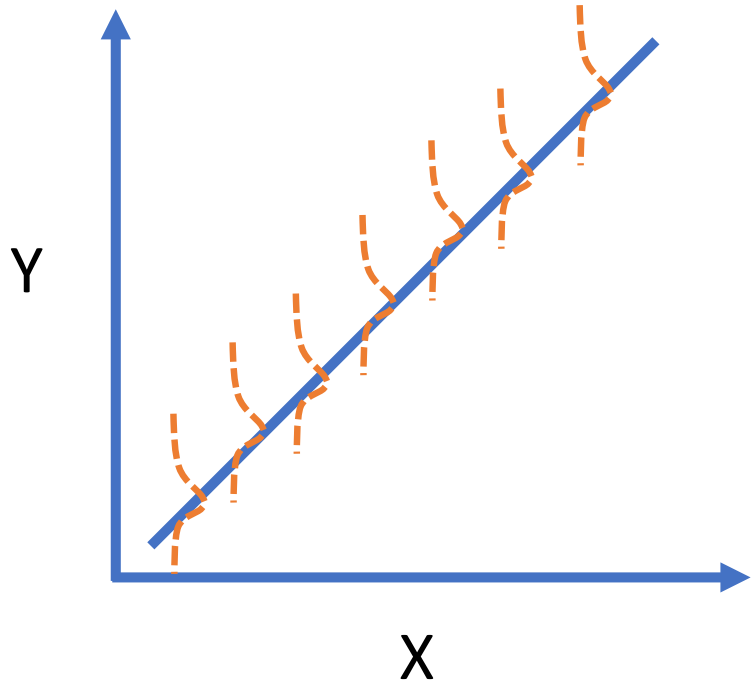
- t-test

$$y = \beta_0 + \beta_1 x + \epsilon$$

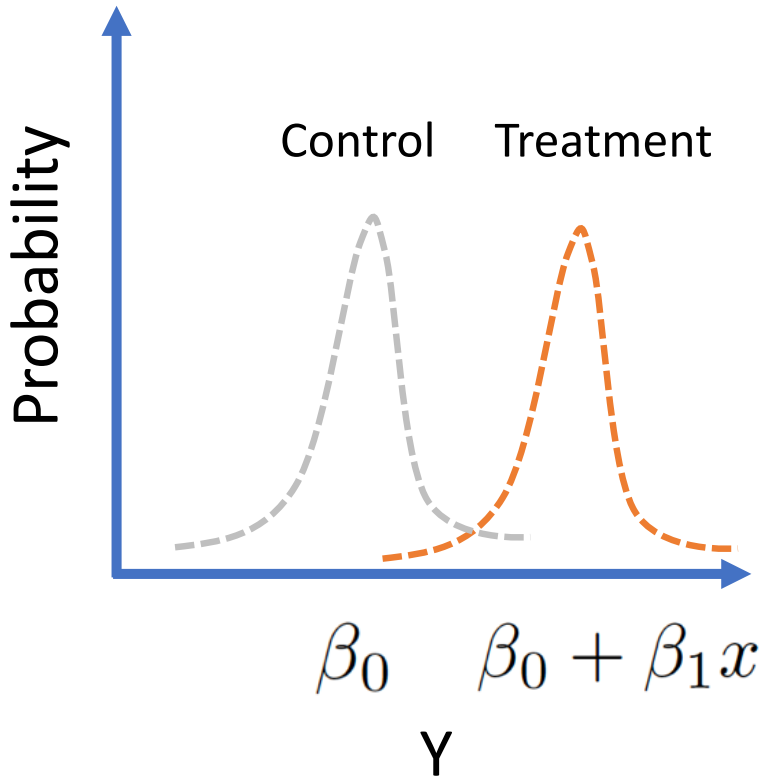
- x – independent variable discrete
- y – dependent variable
- β – linear deterministic components of the regression
- ϵ - error

Linear regression vs. t-test

- linear regression



t-test



The recipe

- load data
- create a function

```
def nllike(p,obs):
    B0=p[0]
    B1=p[1]
    sigma=p[2]

    expected=B0+B1*obs.x
    nll=-1*norm(expected,sigma).logpdf(obs.y).sum()
    return nll
```

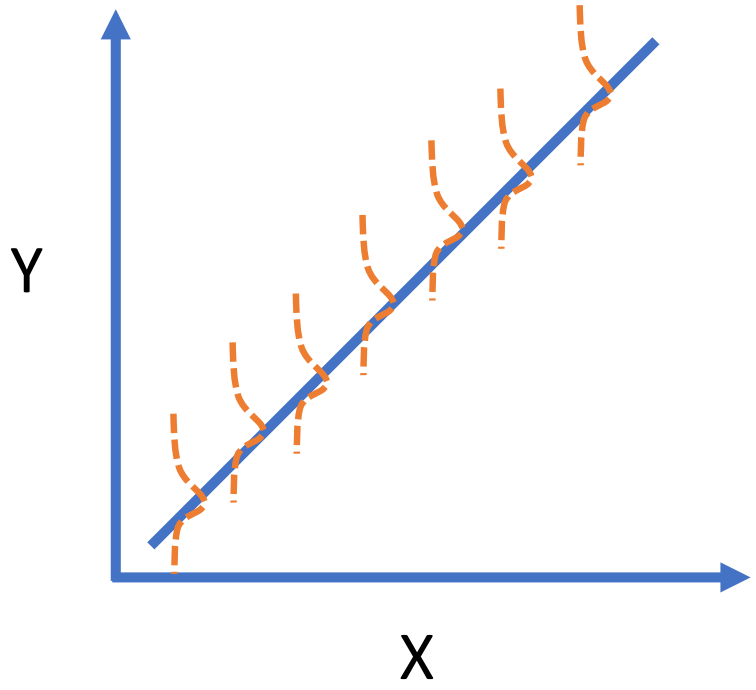
The recipe

- load data
- create a function
- estimate the maximum likelihood values

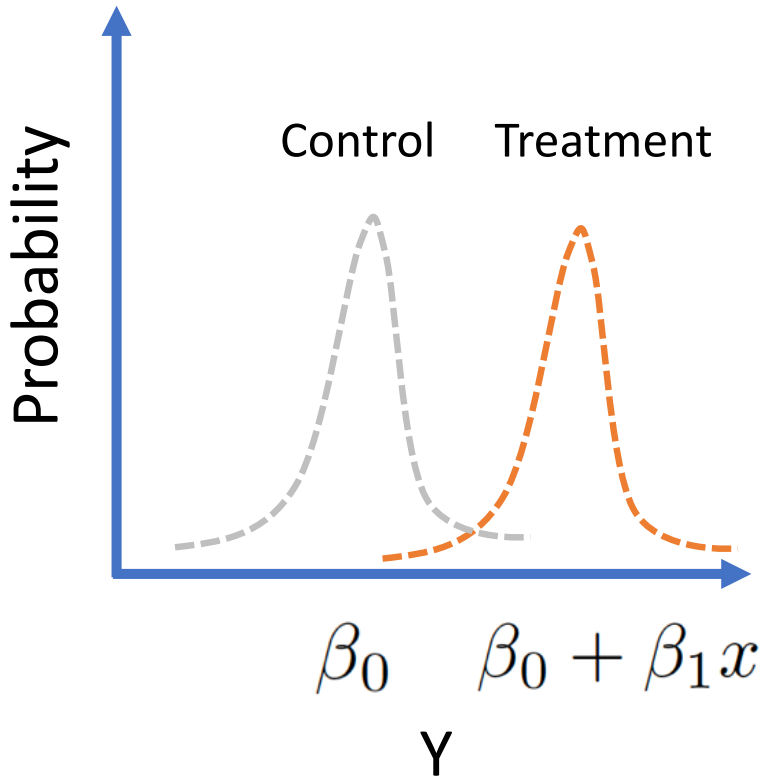
```
### estimate parameters by minimizing the negative log likelihood
initialGuess=numpy.array([1,1,1])
fit=minimize(nllike,initialGuess,method="Nelder-Mead",options={'disp': True},args=df)
```

Linear regression vs. t-test

- linear regression



t-test



For t-test

- Null hypothesis

```
def null (p,obs):
    B0=p[0]
    sigma=p[1]

    expected=B0
    nll=-1*norm(expected,sigma).logpdf(obs.y).sum()
    return nll
```

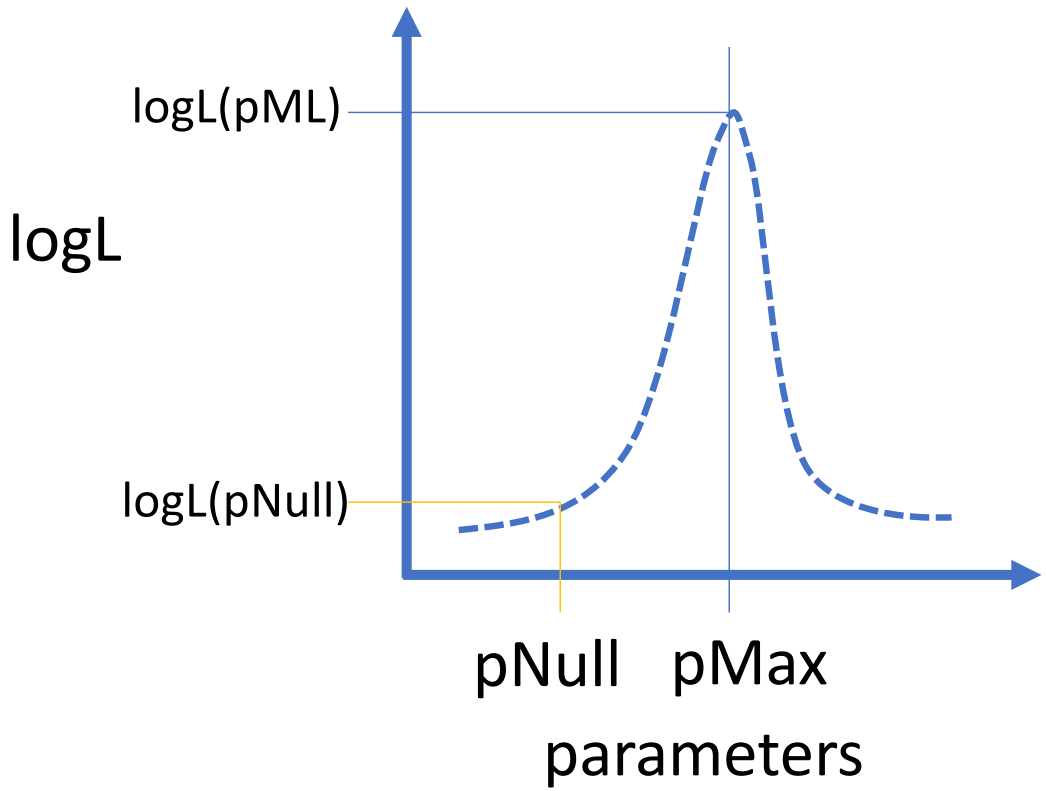
- Alternative hypothesis

```
def alter (p,obs):
    B0=p[0]
    B1=p[1]
    sigma=p[2]

    expected=B0+B1*obs.x
    nll=-1*norm(expected,sigma).logpdf(obs.y).sum()
    return nll
```

Chi2 test

- Estimate maximum likelihood for both Null and Alternative models
- LR = 2(logL(pML) – logL(pNull)) ~ chi2 with df



H0: p=pNull
H1: p>pNull

LR = large: reject H0
LR = small: accept H0

For t-test

- Estimate maximum likelihood for both Null and Alternative models
- Chi2 test

```
initialGuess=numpy.array([1,1,1])
fitNull=minimize(fun1a,initialGuess,method="Nelder-Mead",options={'disp':True},args=q1)
fitAlter=minimize(fun1b,initialGuess,method="Nelder-Mead",options={'disp':True},args=q1)

D=2*(fitNull.fun-fitAlter.fun)
1-chi2.cdf(x=D,df=1)
```