

## 模拟老虎机：

收益为0~100，老虎机*i*在第*t*轮的收益从一个均值 $\mu_{i,t}$ ，标准差 $\sigma_o = 4$ 的高斯分布中，随机生成一个整数。

每一轮的 $\mu_{i,t}$ 的变化为包含decay的**高斯四随机游走**：

$$\mu_{i,t+1} = \lambda\mu_{i,t} + (1 - \lambda)\theta + v$$

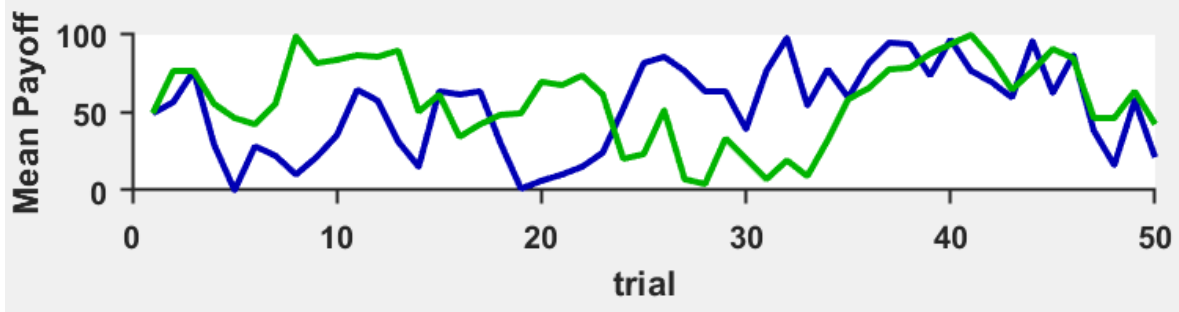
decay参数 $\lambda = 0.9836$ ，decay中心 $\theta = 50$ ，随机噪音*v*服从均值为0、标准差 $\sigma_{noise} = 25$ 的高斯分布，较大的随机噪音目的是为了产生较大的prediction error，如果超出取值范围则进行自动矫正

```
%% 老虎机参数设置
clear;clc; close all
trialNum=50;%总轮数
noisePay=4;%payoff随机噪音的标准差
initPay_A=50;initPay_B=50;%起始payment
lambda=0.9836;%decay参数
theta=50;%decay的中心
noise=25;%游走过程的随机噪音标准差
up=100;down=0;%payment的边界
A_mu=zeros(trialNum,1);B_mu=zeros(trialNum,1);AR_Box=zeros(trialNum,1);RB_Box=zeros(trialNum,1);%存储生成的数据
B_mu(1)=initPay_B;A_mu(1)=initPay_A;AR_Box(1)=A_mu(1);RB_Box(1)=B_mu(1);

for t=2:length(A_mu)
    a_mu=lambda*A_mu(t-1)+(1-lambda)*theta+normrnd(0,noise);
    b_mu=lambda*B_mu(t-1)+(1-lambda)*theta+normrnd(0,noise);
    disp(1)
    A_mu(t)=a_mu;B_mu(t)=b_mu;%mean payoff
    a_pay=round(normrnd(a_mu,noisePay));b_pay=round(normrnd(b_mu,noisePay));
    %如果老虎机的金额小于0，则取相反数（关于0对称），例如-10变成10
    %如果大于100，则将这个数关于100取对称，例如140变为60，
    %反复上述过程直到取值介于0~100之间
    while ~(a_pay>=down && a_pay<=up)
        if a_pay>=up
            a_pay=2*up-a_pay;
        end
        if a_pay<=down
            a_pay=2*down-a_pay;
        end
    end
    while ~(b_pay>=down && b_pay<=up)
        if b_pay>=up
            b_pay=2*up-b_pay;
        end
        if b_pay<=down
            b_pay=2*down-b_pay;
        end
    end
    AR_Box(t)=a_pay;RB_Box(t)=b_pay;%actual payoff
end
```

绘制老虎机的payoff变化形式画图：

```
figure
subplot(3,1,1)
hold on
plot(AR_Box, '-', 'Color', [0 0 180]./255, 'Linewidth', 2)%A蓝
plot(RB_Box, '-', 'Color', [0 180 0]./255, 'Linewidth', 2)%B绿
ax=gca;ax.Linewidth=1;ax.FontName='TimesNewRoman';ax.FontWeight='bold';ax.Box='off';ax.TickDir = 'out';
ylabel('Mean Payoff');xlabel('trial');xlim([0 trialNum]);ylim([down up])
```



## 模拟被试:

经典强化学习模型:

$$Q_{i,t+1} = Q_{i,t} + \alpha(R - Q_{i,t})$$

SoftMax函数:

$$P_t(i) = \frac{\exp(Q_{i,t}/\beta)}{\sum_{j=1}^n \exp(Q_{j,t}/\beta)}$$

两个参数: 学习率(learning rate)  $\alpha$ , inverse temperature  $1/\beta$

设  $Q_{i,0} = 50$

```
%% 模拟被试
alpha=0.5;%固定学习率为0.5;
beta=0.5;%SoftMax参数
QL=@(q,r)q+alpha.*(r-q);%经典强化学习, learning rate不变
Pi=@(Qi,Qj)exp(Qi./beta)/(exp(Qi./beta)+exp(Qj./beta));%SoftMax模拟被试选择
qa_Box=zeros(size(AR_Box));qb_Box=zeros(size(RB_Box));PE_Box=zeros(size(AR_Box));
;
Choice_Box=zeros(size(AR_Box));qa=50;q=50;
QA=[];QB=[];PEBOX=[];
for i=1:trialNum %模拟强化学习
    Pca=Pi(qa,qb);
    Choice=binornd(1,Pca);Choice_Box(i)=Choice;

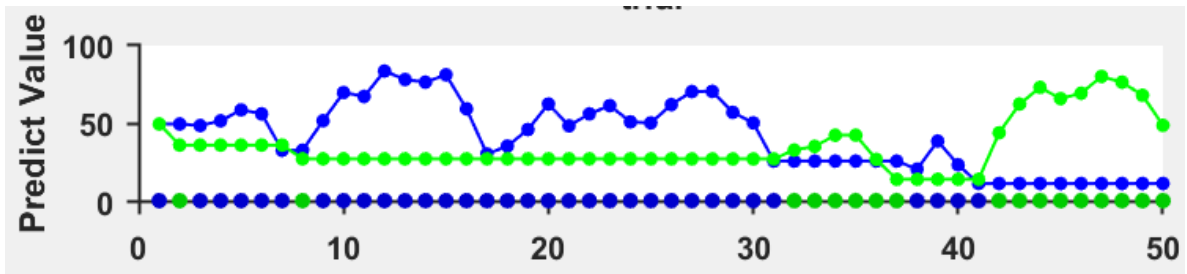
    if Choice==1
        rQ=RA_Box(i);
        PE_Box(i)=rQ-qa;
        qa=QL(qa,rQ);
    else
        rQ=RB_Box(i);
        PE_Box(i)=rQ-q;
        qb=QL(qb,rQ);
    end
end
```

```

        qa_Box(i)=qa;
        qb_Box(i)=qb;
    end
    QA=[QA;qa_Box];
    QB=[QB;qb_Box];

%% 画图
subplot(3,1,2)
hold on
dotColor=repmat([0 0 0.8],trialNum,1).*Choice_Box+repmat([0 0.8 0],trialNum,1).*(1-Choice_Box);
dotSize=25;
scatter(1:trialNum,ones(1,trialNum).*1.1,dotSize,dotColor,'filled');
plot(qa_Box,'.-b','MarkerSize',15,'Linewidth',1)
plot(qb_Box,'.-g','MarkerSize',15,'Linewidth',1)
ax=gca;ax.Linewidth=1;ax.FontName='TimesNewRoman';ax.FontWeight='bold';ax.Box='off';ax.TickDir = 'out';
ylabel('Predict Value');xlabel('trial');xlim([0 trialNum]);

```

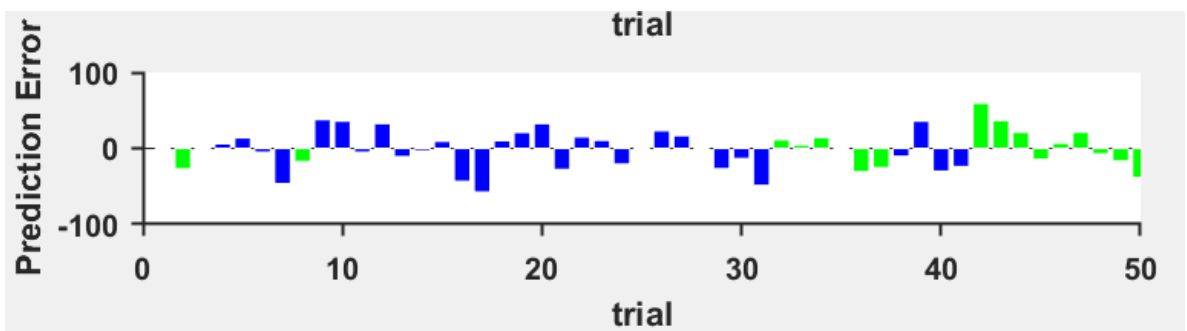


根据参数，计算每一轮对应的prediction error

```

%% PE
subplot(3,1,3)
hold on
dotColor=repmat([0 0 1],trialNum,1).*Choice_Box+repmat([0 1 0],trialNum,1).*(1-Choice_Box);
for i=1:trialNum
    bar(i,PE_Box(i),'FaceColor',dotColor(i,:), 'EdgeColor',[1 1 1]);
end
ax=gca;ax.Linewidth=1;ax.FontName='TimesNewRoman';ax.FontWeight='bold';ax.Box='off';ax.TickDir = 'out';
ylabel('Prediction Error');xlabel('trial');xlim([0 trialNum]);

```



## 根据被试选择反推其参数：最大似然估计拟合

利用fminsearchbnd.m函数、强化学习模型似然估计函数RL.m进行拟合，

RL的参数一共包括4个： $\alpha$ 、 $\beta$ 、对老虎机a和b的起始值估计：qa和qb

```
%% 根据选择拟合被试数据
tot=20; %重复多次相同的拟合操作，选择其中最大似然值最大一次的拟合结果
options = optimset('MaxFunEvals',100000, 'MaxIter', 10000);

alphaBox=zeros(1,tot);betaBox=zeros(1,tot);
qaBox=zeros(1,tot);qbBox=zeros(1,tot);
LLBox=zeros(1,tot);
N=ones(trialNum,1);
for k=1:tot
    fprintf('%1.0f...',k)
    LB = [0 0 0 0];%参数下界
    UB = [1 inf 100 100];%参数上界
    x0 = [rand rand];
    [paramsEst, minuslli, ~] = ...
        fminsearchbnd(@(params)RL(params,RA_Box,RB_Box,trialNum,N,Choice_Box),
    x0, LB, UB,options);%并利用fminsearchbnd函数获得最大似然的值
    alphaBox(k) = paramsEst(1);betaBox(k) = paramsEst(2);qaBox(k) =
    paramsEst(3);qbBox(k) = paramsEst(4); LLBox(k)= minuslli;
end
indx=find(LLBox==max(LLBox));indx=indx(1);
alpha_new=alphaBox(:,indx);beta_new=betaBox(:,indx);
qa_new=qaBox(:,indx);qb_new=qbBox(:,indx); %记录拟合的新参数
```

利用拟合得到的参数建立新的被试模型

```
QL_new=@(q,r)q+alpha_new.*(r-q);%新的经典强化学习
P_choose_i_new=@(Qi,Qj)exp(Qi./beta_new)/(exp(Qi./beta_new)+exp(Qj./beta_new));%
新的SoftMax模拟被试选择
```

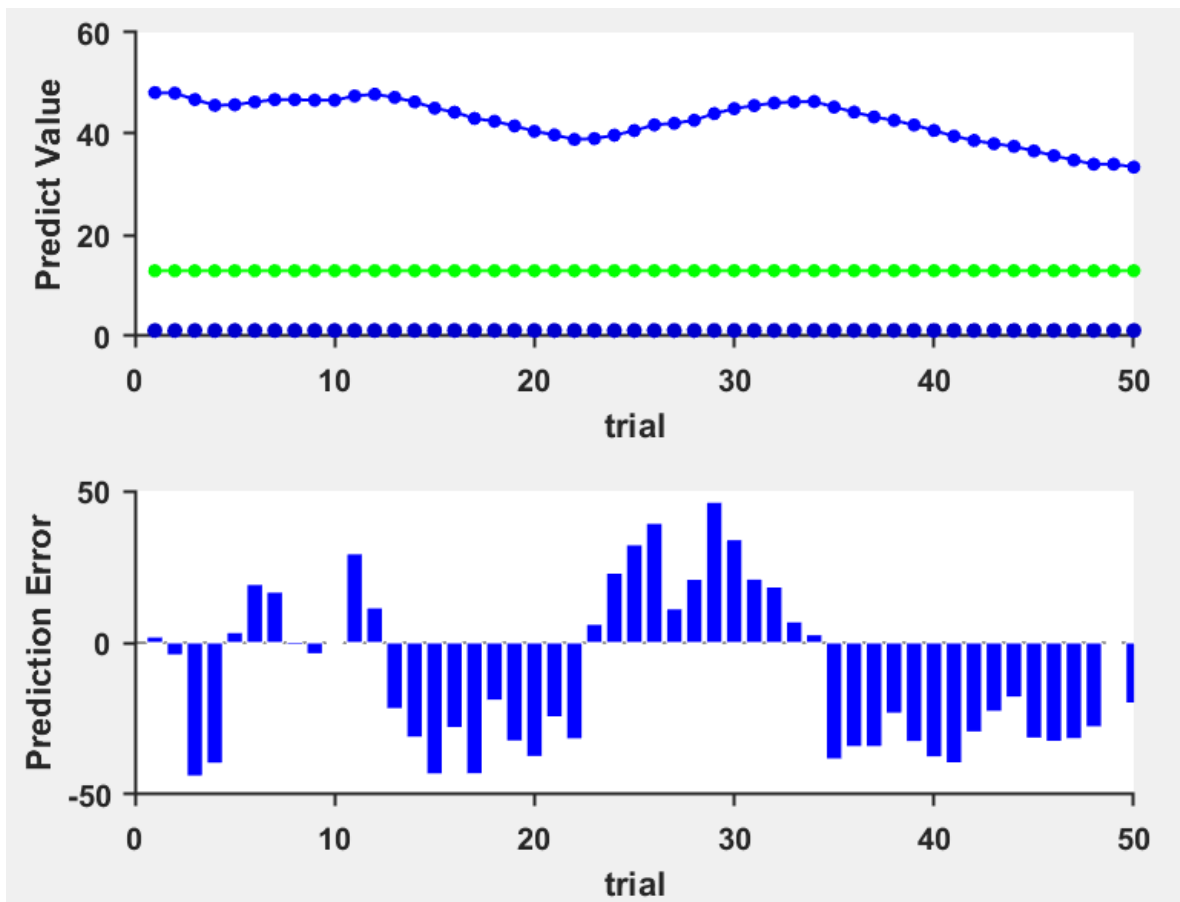
将新的被试模型用于相同老虎机情景中：

```
qa_Box2=zeros(size(RA_Box));qb_Box2=zeros(size(RB_Box));PE_Box2=zeros(size(RA_Box));
Choice_Box=zeros(size(RA_Box));
qa=qa_new;q=qb_new;%新的起始点
for i=1:trialNum %模拟强化学习
    Pca=P_choose_i_new(qa,qb);
    Choice=binornd(1,Pca);Choice_Box(i)=Choice;
    if Choice==1
        rQ=RA_Box(i);
        PE_Box2(i)=rQ-qa;
        qa=QL_new(qa,rQ);
    else
        rQ=RB_Box(i);
        PE_Box2(i)=rQ-qb;
        qb=QL_new(qb,rQ);
    end
    qa_Box2(i)=qa;
    qb_Box2(i)=qb;
end
```

```

figure
subplot(2,1,1)
hold on
dotColor=repmat([0 0 0.8],trialNum,1).*Choice_Box+repmat([0 0.8 0],trialNum,1).*(1-Choice_Box);
dotSize=25;
scatter(1:trialNum,ones(1,trialNum).*1.1,dotSize,dotColor,'filled');
plot(qa_Box2,'.-b','MarkerSize',15,'Linewidth',1)
plot(qb_Box2,'.-g','MarkerSize',15,'Linewidth',1)
ax=gca;ax.Linewidth=1;ax.FontName='TimesNewRoman';ax.FontWeight='bold';ax.Box='off';ax.TickDir = 'out';
ylabel('Predict Value');xlabel('trial');xlim([0 trialNum]);
%% PE
subplot(2,1,2)
hold on
dotColor=repmat([0 0 1],trialNum,1).*Choice_Box+repmat([0 1 0],trialNum,1).*(1-Choice_Box);
for i=1:trialNum
    bar(i,PE_Box2(i),'FaceColor',dotColor(i,:), 'EdgeColor',[1 1 1]);
end
ax=gca;ax.Linewidth=1;ax.FontName='TimesNewRoman';ax.FontWeight='bold';ax.Box='off';ax.TickDir = 'out';
ylabel('Prediction Error');xlabel('trial');xlim([0 trialNum]);

```



(这次估计把qb的起始值估计过低导致)

最后将被试在原始参属下的PE与拟合结果下的PE做相关

%% 画两次PE的散点图

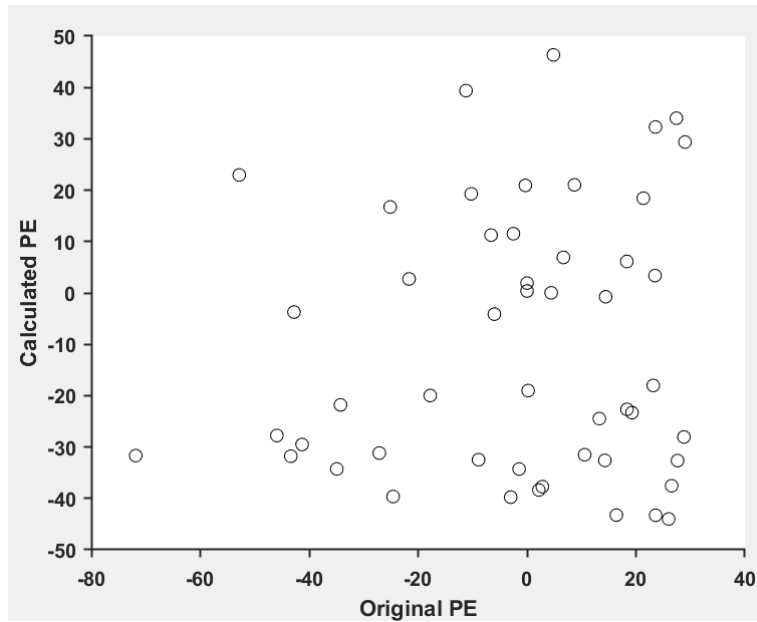
figure

scatter(PE\_Box,PE\_Box2,40,[0 0 0])

ax=gca;ax.Linewidth=1;ax.FontName='TimesNewRoman';ax.Fontweight='bold';ax.Box='off';ax.TickDir = 'out';

ylabel('Calculated PE');xlabel('Original PE');

r=corr(PE\_Box,PE\_Box2)



( $r=0.09$ , 这次结果不靠谱, 原因在于之前的估计都不包括对老虎机a和b的起始值估计:  $q_a$ 和 $q_b$ , 只估计了 $\alpha$ 、 $\beta$ )

上次的结果记录: (只估计了 $\alpha$ 和 $\beta$ )

