Dynamic Discrete Choice Models

Thierry Kamionka

2021-2022

Part I : Simutlated EM - MLE and MSM

The idea consists to specify the problem solved by the individual in order to estimate the parameters of the structural model. This approach was introduced by Rust (1987). This example is used here to illustrate the approach.

Rust considers the optimal replacement of **GMC bus engines.**

The **state variable** is denoted x_t. It is the accumulated mileage of the GMC bus engines since the last replacement.
Rust (1987) assumes that, during a period, the mileage traveled, namely δx, is distributed as an exponential random variable with pdf $\theta_2 \exp(-\theta_2 \, \delta x)$, where θ_2 is a parameter.

The utility function u is

$$u(x_t, i_t, \theta_1) = \begin{cases} -c(x_t, \theta_1) & \text{if} \quad i_t = 0, \\ -[\bar{P} - \underline{P} + c(0, \theta_1)] & \text{if} \quad i_t = 1, \end{cases}$$

where $\bar{P}$ is a cost of a new engine, $\underline{P}$ is a scrap value for the old engine and $c(x, \theta_1)$ is the **operating costs** when the accumulated mileage is x. i_t is the decision.

The stochastic process governing the evolution of x_t is defined by the transition probability

$$p(x_{t+1}|x_t,i_t,\theta_2)=\begin{cases}\theta_2\exp(-\theta_2(x_{t+1}-x_t)) & \text{if } i_t=0 \text{ and } x_{t+1}\geq x_t,\\ \theta_2\exp(-\theta_2 x_{t+1}) & \text{if } i_t=1 \text{ and } x_{t+1}\geq 0,\\ 0 & \text{otherwise},\end{cases}$$

The **value function** is denoted $V_\theta(x_t)$ and it is the unique solution to the Bellman's equation

$$V_\theta(x_t) = \max_{i_t\in\{0,1\}}[u(x_t,i_t,\theta_1)+\beta EV_\theta(x_t,i_t)] \tag{1}$$

where

$$EV_\theta(x_t,i_t) = \int_0^\infty V_\theta(y)\,p(dy \mid x_t,i_t,\theta_2),\ \text{Rust (1987) notation}$$

$$EV_\theta(x_t,i_t) = \int_0^\infty V_\theta(y)\,p(y \mid x_t,i_t,\theta_2)\,dy, \text{Using an other notation}$$

Using Bellman's equation, there exists an optimal Markovian replacement policy such that

$$i_t = f(x_t,\theta) = \begin{cases} 1 & \textit{if } x_t > \gamma(\theta_1,\theta_2), \\ 0 & \textit{if } x_t \leq \gamma(\theta_1,\theta_2), \end{cases}$$

where $\gamma(\theta_1,\theta_2)$ is the unique solution to the equation

$$(\bar{P} - \underline{P})(1-\beta) = \int_0^{\gamma(\theta_1,\theta_2)} [1 - \beta \exp(-\theta_2(1-\beta)y)] \frac{\partial c(y,\theta_1)}{\partial y} dy$$

Assuming that monthly mile age and replacement decisions are independently distributed across buses, the likelihood function is

$$L(\theta_1,\theta_2) = \prod_{i=1}^{n} L_i(\theta_1,\theta_2)$$

where $L_i(\theta_1,\theta_2)$ is the contribution of bus i to the likelihood function.

For instance, the contribution of a bus such that we observe $(i_1, \ldots, i_T, x_1, \ldots, x_T)$, where T is the number of periods is

$$L_i(\theta) = \prod_{t=1}^{T} \mathbb{1}[i_{t-1}=0; x_{t-1} \leq \gamma(\theta_1, \theta_2); x_t \geq x_{t-1}] \, \theta_2 \exp(-\theta_2(x_t - x_{t-1}))$$

$$+ \mathbb{1}[i_{t-1}=1; x_{t-1} > \gamma(\theta_1, \theta_2); x_t \geq 0] \, \theta_2 \exp(-\theta_2 \, x_t)$$

where $i_0 = 1$ and $x_0 = 0$ (in order to simplify the notations).

Rust (1987) shows, moreover, that the optimal stopping time is Poisson with parameter $\theta_2 \gamma(\theta_1, \theta_2)$.

In such a **structural model**, the likelihood function is obtained from the solution to the optimization problem and the parametric specification comes from the utility function $u(.)$ and the transition probability $p(.)$.

. Rust (1987) underlines several limits to such a specification.
. The exponential distribution constraints the mean and the standard deviation of the monthly mileage to be the same. On the data, the standard deviation is only one third of the average.
. He tries to use more realistic distribution for the mileage but he cannot obtain an explicit solution for the stochastic control problem. (Remark : why not consider that the parameter of the exponential is $\theta_2\nu$, where ν is known from the decider but not for the econometrician ?)
.On data, the mileage at replacement varies a lot for a bus to the other. This variation is too large to be consistent with single fixed optimal stopping barrier γ.
.The econometrician may assume that the decider - **Harold Zurcher** - might base the replacement decisions on other information than the mileage x_t. Such an information may be unobserved by econometrician. Rust attempts to reformulate the model in this direction but it leads models with no analytical solution.

The approach Rust (1987) considers does not require a closed form solution for stochastic control problem of the decider. It incorporates unobservables variables ϵ_t in the model.

The objects of the model in the general case:

$C(x_t)$: **Choice set** when the state variable is x_t. The **decision** $i_t \in C(x_t)$. (Harold Zurcher : $C(x_t) = \{0, 1\}$).

$\epsilon_t = \{\epsilon_t(i) \mid i \in C(x_t)\}$: where $\epsilon_t(i)$ is a component of the utility associated to alternative i and period t. It is known by **decider** (a vector of state variables) and not by **econometrician** (unobserved heterogeneity).

$U(x_t, i, \theta_1) + \epsilon_t(i)$: Utility of the alternative i when the **state variable** is (x_t, ϵ_t) and θ_1 is a vector of parameters.

$x_t = (x_t(1), \ldots, x_t(K))$ a vector of state variables observed by decider and econometrician.

$p(x_{t+1}, \epsilon_{t+1} \mid x_t, \epsilon_t, i_t, \theta_2, \theta_3)$: the Markov transition density of the states variables (x_t, ϵ_t) if the agent decide to choose alternative i_t. θ_2 is a vector of parameters associated to the marginal density of x_{t+1} and θ_3 is the vector of parameters specific to the conditional density of ϵ_{t+1} given x_{t+1} .

$\theta = (\beta, \theta_1, \theta_2, \theta_3)$: is the vector of parameters (β is a **discounting factor**).

Let $V_\theta(x_t, \epsilon_t)$ denote the **expected discounted utility over an infinite horizon for the optimal decision** given states variables.

Under regularity conditions, Rust (1987) shows that the solution of the program of optimization of the expected discounted utility over an infinite horizon is given by a **stationary decision rule**

$$i_t = f(x_t, \epsilon_t, \theta)$$

It is the **optimal decision** of the agent given the state variables are (x_t, ϵ_t).

Remark: a finite horizon would induce a non stationary optimal decision. In the application for optimal bus replacement decision, the period t is a month.

The optimal value $V_\theta(x_t, \epsilon_t)$ is the unique solution of the **Bellman's equation**

$$V_\theta(x_t, \epsilon_t) = \max_{i \in C(x_t)} [u(x_t, i, \theta_1) + \epsilon_t(i) + \beta EV_\theta(x_t, \epsilon_t, i)] \qquad (2)$$

where

$$EV_\theta(x_t, \epsilon_t, i) = \int_y \int_\eta V_\theta(y, \eta)\, p(dy, d\eta \mid x_t, \epsilon_t, i, \theta_2, \theta_3)$$

and the **optimal control** f is

$$f(x_t, \epsilon_t, \theta) = \underset{i \in C(x_t)}{arg\,max} [u(x_t, i, \theta) + \epsilon_t(i) + \beta EV_\theta(x_t, \epsilon_t, i)]$$

Thus V_θ is the fixed point of the Bellman's equation.

$$\mathbf{V}_\theta(x_t,\epsilon_t) = \max_{i\in C(x_t)} [u(x_t,i,\theta_1)+\epsilon_t(i) + \beta \int_y \int_\eta \mathbf{V}_\theta(y,\eta)\, p(dy,d\eta|x_t,\epsilon_t,i,\theta_2,\theta_3)]$$

The dimensionality of the problem is (too) large. The conditional distribution at time t depend on ϵ_{t-1} that is unobserved by the econometrician.

Conditional Independence Assumption (CIA) : The transition density is such that

$$p(x_{t+1},\epsilon_{t+1}|x_t,\epsilon_t,i,\theta_2,\theta_3) = q(\epsilon_{t+1}|x_{t+1},\theta_2)p(x_{t+1}|x_t,i,\theta_3) \quad (3)$$

Assumptions (CIA) implies that x_{t+1} is a sufficient statistics for ϵ_{t+1} and the density of x_{t+1} depend on x_t and not on ϵ_t. What is important is the evolution of the **state variable** x_t (mileage of the bus) and ϵ_t appears as a **kind of noise.**

Let $q(\epsilon \mid x, \theta_2)$ denote the pdf of the multivariate extreme value distribution

$$q(\epsilon \mid x, \theta_2) = \prod_{j \in C(x)} \exp\{-\epsilon(j) + \theta_2\} \exp\{-\exp\{-\epsilon(j) + \theta_2\}\} \quad (4)$$

where $\theta_2 = 0.577216$ (Euler's constant).

$P(i \mid x; \theta)$ denote the conditional probability the agent chooses the alternative $i \in C(x)$ given state variable x.

The conditional choice probabilities can be computed using the formulas of the static model of discrete choice (see McFadden, 1981) adding the term $\beta EV_\theta(x, i)$ to the utility function $u(x, i, \theta_1)$.

When $p(x_{t+1} \mid x_t, i, \theta_3)$ does not depend on i (myopic agent), the **static model** of discrete choice is a special case of the **dynamic discrete choice model**.

Let EV_θ denote the unique solution to the functional equation

$$EV_\theta(x,i) = \int_y \ln\left(\sum_{j\in C(x)} \exp[u(y,j,\theta_1) + \beta EV_\theta(y,j)]\right) p(dy \mid x,i,\theta_3).$$

Theorem

Le us assume that (CIA) holds and $q(\epsilon \mid x;\theta_2)$ is the multivariate extreme value distribution. The $P(i \mid x;\theta)$ is given by

$$P(i \mid x;\theta) = \frac{\exp\{u(x,i,\theta_1) + \beta EV_\theta(x,i)\}}{\sum_{j\in C(x)} \exp\{u(x,j,\theta_1) + \beta EV_\theta(x,j)\}} \tag{5}$$

The conditional probability that the agent choose alternative $i \in C(x)$ is given by the multinomial formula.

Remark : If $P(x_{t+1} \mid x_t, i, \theta_3)$ does not depend on i, $EV_\theta(x, i)$ is independent of i and

$$P(i \mid x; \theta) = \frac{\exp\{u(x, i, \theta_1)\}}{\sum_{j \in C(x)} \exp\{u(x, j, \theta_1)\}}$$

It is the **usual static choice formula** for the choice probabilities. In other words, if the current decision do not have any impact on the evolution of the state variables $(x_{t+1}, \epsilon_{t+1})$, this decision has no consequence in the future. In this case, it is optimal that the agent behave **myopically** for each period of time and choose the alternative i that maximizes $u(x_t, i, \theta_1) + \epsilon_t(i)$ without considering the future.

If the current decision have an impact on the future, $EV_\theta(x, i)$ must be added to the utility to correctly describe the behavior of the agent. $EV_\theta(x, i)$ can be interpreted as a **"shadow price"** (or cost) of the future consequences of current decisions.

Remark : Rust (1987) gives the expression of the conditional probabilities $P(i \mid x;\theta)$ under the (CIA) assumption **relaxing the** multivariate extreme value distribution **assumption.**

The data observed for a given individual:
$\{(i_0, x_0), (i_1, x_1), \dots (i_T, x_T)\}$, where T is the number of observation times.
Under the asumption (CIA) and if $p(\epsilon \mid x, \theta_2)$ is given by a the multivariate extreme value distribution (4), a contribution to the likelihood function is

$$L(x_1, \dots, x_T, i_1, \dots, i_T \mid x_0, i_0) = \prod_{t=1}^{T} P(i_t \mid x_t, \theta)\, p(x_t \mid x_{t-1}, i_{t-1}, \theta_3), \tag{6}$$

where $P(i_t \mid x_t, \theta)$ is given by equation (5).
Under the assumption the sample is i.i.d. the likelihood function is the product of the contributions of the individuals.

If the number of individuals tends to infinity, the **MLE** of θ is consistent and asymptotically normal.

The bus engine replacement:

The choice set is $C(x_t) = \{0, 1\}$.

The unobservable state variables $(\epsilon_t(0), \epsilon_t(1))$ are distributed as a bivariate extreme value process, with mean $(0, 0)$ and variance $(\pi^2/6, \pi^2/6)$.

The expected cost of a replacement bus engine is $RC = \bar{P} - \underline{P}$.

The utility function is

$$u(x_t, i, \theta_1) + \epsilon_t(i) = \begin{cases} -RC - c(0, \theta_1) + \epsilon_t(1), & \text{if } i = 1, \\ -c(x_t, \theta_1) + \epsilon_t(0), & \text{if } i = 0, \end{cases}$$

The transition density is

$$p(x_{t+1} \mid x_t, i, \theta_3) = \begin{cases} g(x_{t+1} - x_t, \theta_3), & \text{if } i_t = 0, \\ g(x_{t+1} - 0, \theta_3), & \text{if } i_t = 1, \end{cases}$$

where g is parametric density function on $\boldsymbol{R}^{+}$ (for instance the exponential distribution with parameters θ_3).

Some functional forms for c

i) polynomial : $c(x, \theta_1) = \theta_{11}\, x + \theta_{12}\, x^2 + \theta_{13}\, x^3$,

ii) exponential : $c(x, \theta_1) = \theta_{11} \exp(\theta_{12}\, x)$,

iii) hyperbolic : $c(x, \theta_1) = \theta_{11}/(91 - x)$,

iv) square root : $c(x, \theta_1) = \theta_{11}\sqrt{x}$,

with the normalization $c(0, \theta_1) = 0$ (if $i = 1$).

The parameters of the model $\theta = (\beta, \theta_1, RC, \theta_3)$ are estimated by maximizing the likelihood function.

Let us define the following partial likelihood functions

$$\ell^1(x_1, \dots, x_T, i_1, \dots, i_T \mid x_0, i_0, \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, i_{t-1}, \theta_3),$$

and

$$\ell^2(x_1, \dots, x_T, i_1, \dots, i_T \mid \theta) = \prod_{t=1}^{T} p(i_t \mid x_t, \theta),$$

The estimation procedure:

The **first stage** consists to estimate θ_3 using the likelihood ℓ^1. This partial likelihood estimator is consistent and asymptotically normal.

The **second stage** consists to use these estimates of θ_3 in order to estimate the remaining parameters β, RC and θ_1 using the partial likelihood function ℓ^2.

In the second stage of the estimation method, one need to calculate the fixed point EV_θ each time it is necessary to evaluate the partial likelihood function ℓ^2.

For the **third stage**, we use the estimation of θ resulting from stages 1 and 2 as initial value and to maximize the likelihood function (6). We maximize the full likelihood function and we obtain **consistent and efficient estimates** of the parameters. We can obtain also a consistent estimation of the variance-covariance matrix of the parameters.

A fixed point algorithm allows to evaluate EV_θ for each value of θ. This can be done using a Newton-Kantorovich algorithm.

We can use a Newton-Raphson algorithm in order to maximize the log likelihood function with respect to θ.

For each period t ($t = 1, 2, 3$) the individual makes a decision $d_t \in \{0, 1\}$. d_t is the decision to complete period of college t.

Example 1: 1) Individual decides to pass the SAT (standardized exam) and obtains the SAT score. 2) He decides to apply to college. 3) He decides to attend college and receives an earnings.

Example 2: 1) Individual decides to attend university for Bachelor level. 2) He applies and obtains a Master level. 3) He complete a PhD and receives earnings.

If the individual decides $d_t = 0$, then for all $t' > t$, $d_{t'} = 0$.

C is a binary variable that indicate whether the individual has **completed college** ($C = d_1 \, d_2 \, d_3$).

After he takes a decisions for period t, he receives a realization on state variables (for instance, decides to apply to college and receives a financial aid). The model is similar to Cameron and Heckman (2001).

After the last decision, he receives earnings denoted Y

$$\ln(Y) = [X \;\; C \;\; Type]\, \gamma + \eta \tag{7}$$

where X is a vector of characteristics (known by the individual even before the final decision), *Type* is the **type of the individual** (**unknown** to the econometrician).
$Type \in \{1, \ldots, K\}$.

η is the part of the earnings that are unknown to the individual until all educational decisions are taken.

We assume η is distributed as a $N(0, \sigma_Y^2)$ and is independent of X, C and *Type*. γ is a vector of parameters to be estimated. We are particularly interested by γ_C (**returns to college education**).

Each action to take to make $C = 1$ is costly (these costs vary from an individual to the other).

Let us consider the beginning of period $t = 3$. The expected present value of lifetime utility to choose $d_3 = 1$ (given $d_1 = d_2 = 1$) is

$$v_3(d_3{=}1) = [Z_3 \;\; Type \;\; (E_3[Y|C{=}1] - E_3[Y|C{=}0])] \, \alpha_3 + \epsilon_3 \qquad (8)$$

where Z_3 is a vector of characteristics. Some of the elements of Z_3 may be known by the individual only at period 3. When he makes earlier decisions (d_1 and d_2), he faces uncertainty for these variables.

ϵ_3 is the unobserved preference to choose $d_3 = 1$ and it is assumed to be distributed as a **logistic random variable** ($d_3 \in \{0, 1\}$). ϵ_3 is unknown to the individual before period 3.

The utility $v_3(d_3{=}0) = 0$ and the individual decides to choose $d_3 = 1$ iff $v_3(d_3{=}1) > 0$.

The expected present value of lifetime utility to choose $d_t = 1$ for $t = 1, 2$ (the decision d_t is described given $d_{t-1} = 1$) is

$$v_t(d_t{=}1) = [Z_t \quad Type \quad E_t[V_{t+1}|d_t{=}1]]\ \alpha_t + \epsilon_t \tag{9}$$

where Z_t is a vector of individual characteristics. ϵ_t is the unobserved preference to choose $d_t = 1$ and it is assumed to be distributed as a **logistic random variable.** ϵ_t is unknown to the individual before period t. We assume ϵ_t are independent across time (but types are unobserved).

The utility $v_t(d_t{=}0) = 0$ and the individual decides to choose $d_t = 1$ iff $v_t(d_t{=}1) > 0$. In order to evaluate (9), the individual has to take expectation over future preferences (ϵ'_t, $t' > t$) and future state variables (Z'_t, $t' > t$).

Using the assumptions on the unobserved preferences terms (ϵ_t) and integrating with respect of the distribution of the future state variables, Rust (1987) has shown that (9) can be written

$$v_t(d_t{=}1) = [Z_t \;\; Type \;\; E_t[V_{t+1}|d_t{=}1]]\, \alpha_t + \epsilon_t$$

where

$$E_t[V_{t+1}|d_t{=}1] = \int \ln(\exp\{\bar{v}_{t+1}(d_{t+1}{=}1|Z_{t+1})\}+1)\, \pi_t(Z_{t+1}) dZ_{t+1}$$

$E_t[V_{t+1}|d_t{=}1]$ is an expected value associated to lifetime utility and π_t is the density of the vector Z_{t+1}. $\bar{v}_{t+1}$ is the value to choose d_{t+1} **net of preferences** ($\epsilon_{t+1} = 0$).

Arcidiacono and Jones (2003) approximate the distribution of unknown state variables ($t = 1$ and $t = 2$) using discrete distributions.

In practice, in a MC experiment, they assume there is only one unknown state variable per period and that these variables are $N(0, \sigma_s^2)$, independent of all other variables from preceding periods and then approximate these distributions for estimation by fixed discrete distributions (10 points discrete distributions, in the spirit of Heckman and Singer, 1984).

Suggestion: replace the expectation by a simulator. Then, $E_t[V_{t+1}|d_t{=}1] = \frac{1}{H}\sum_{h=1}^{H} \ln(\exp\{\bar{v}_{t+1}(d_{t+1}{=}1|z_h)\}{+}1)$ where $z_h = \sigma_s \Phi^{-1}(u_h)$ and u_h are iid. uniform random draws.

$\Phi(.)$ is the cumulative distribution function of the standard normal distribution.

The random draws u_h, $h = 1, \ldots, H$, should be specific to the individual. These states variables are unknown for the individual before the period of their realization but they are observed for the econometrician. It then possible to estimate σ_s using MLE from the incomplete observation (consistent by not efficient).

The probability that the individual choose $d_t = 1$ is (see Rust, 1987, (4.13))

$$Prob_t(d_t = 1 \mid X, Z, Type; \gamma, \alpha) = \frac{\exp(\bar{v}_t(d_t = 1))}{\exp(\bar{v}_t(d_t = 1)) + 1}$$

Rust (1987) assume the error term ϵ has a multivariate extreme value distribution.
The likelihood function is (hypothetical case such that Type is observed)

$$L(\gamma, \alpha) = \prod_{t=1}^{3} Prob_t(d_t \mid X, Z, Type; \gamma, \alpha)\ f(Y \mid X, C, Type; \gamma)$$

where f is the conditional probability density function of earnings Y.

Remark: If the Type of the individual were observed, the contribution to the log likelihood function would be

$$\ln(L(\gamma,\alpha)) = \ln(L_4(\gamma)) + \ln(L_3(\gamma,\alpha_3)) + \ln(L_2(\gamma,\alpha_3,\alpha_2)) + \ln(L_1(\gamma,\alpha_3,\alpha_2,\alpha_1))$$

Remark : The simulations of the expected lifetime utility appear inside the logit probabilities. Consequently, the situation is different of the one corresponding to a mixed logit model (see Ruud 2007, McFadden and Train, 2000).

$\mathbf{L_4}$ is the conditional log-likelihood function of earnings, $\mathbf{L_t}$, $t \in \{1,2,3\}$, is the log-likelihood function corresponding to the choice d_t at the beginning of period t.

Consequently, the estimation of the parameters can be made, in such a situation, **sequentially**.

In such a context

- A consistent estimation of γ can be obtained maximizing $\ln(L_4(\gamma))$,
- A consistent estimation of α_3 can be obtained maximizing $\ln(L_3(\hat{\gamma}, \alpha_3))$,
- A consistent estimation of α_2 can be obtained maximizing $\ln(L_2(\hat{\gamma}, \hat{\alpha}_3, \alpha_2))$,
- A consistent estimation of α_1 can be obtained maximizing $\ln(L_1(\hat{\gamma}, \hat{\alpha}_3, \hat{\alpha}_2, \alpha_1))$,

Let us consider now the context such that the **type of the individual is unobserved** by econometrician. A contribution to the log likelihood function is

$$\ln(L(\gamma, \alpha, p)) = \ln(\sum_{k=1}^{K} p_k \, L_{4k}(\gamma) L_{3k}(\gamma, \alpha_3) L_{2k}(\gamma, \alpha_3, \alpha_2) L_{1k}(\gamma, \alpha))$$

where p_k is the probability to belong to type k $(\alpha = (\alpha_1, \alpha_2, \alpha_3)')$.

Consequently, in such a context, the log-likelihood function is not separable.

Expectation Sequential Maximization (**ESM algorithm**):

The maximization step of the algorithm, the contribution of a given individual to the objective function is

$$\ln(L(\gamma, \alpha, p)) = \sum_{k=1}^{K} Prob[k|Y, X, Z, C; \gamma, \alpha, p] \, [\ln(L_{4k}(\gamma))$$

$$+\ln(L_{3k}(\gamma, \alpha_3)) + \ln(L_{2k}(\gamma, \alpha_3, \alpha_2)) + \ln(L_{1k}(\gamma, \alpha))]$$

where p_k is the probability to belong to type k and $Prob[k|Y, X, Z, C; \gamma, \alpha, p]$ is the conditional probability that the individual belongs to type k.

Using $(\tilde{\gamma}, \tilde{\alpha}, \tilde{p})$ obtained from the previous iteration, one can maximize (sequentially)

$$\tilde{\tilde{\gamma}} = \underset{\gamma}{argmax} \sum_{i=1}^{n} \sum_{k=1}^{K} Prob[k | Y, X, Z, C; \tilde{\gamma}, \tilde{\alpha}, \tilde{p}] \ \ln(L_{i4k}(\gamma))$$

$$\tilde{\tilde{\alpha}}_3 = \underset{\alpha_3}{argmax} \sum_{i=1}^{n} \sum_{k=1}^{K} Prob[k | Y, X, Z, C; \tilde{\gamma}, \tilde{\alpha}, \tilde{p}] \ \ln(L_{i3k}(\tilde{\tilde{\gamma}}, \alpha_3))$$

$$\tilde{\tilde{\alpha}}_2 = \underset{\alpha_2}{argmax} \sum_{i=1}^{n} \sum_{k=1}^{K} Prob[k | Y, X, Z, C; \tilde{\gamma}, \tilde{\alpha}, \tilde{p}] \ \ln(L_{i2k}(\tilde{\tilde{\gamma}}, \tilde{\tilde{\alpha}}_3, \alpha_2))$$

$$\tilde{\tilde{\alpha}}_1 = \underset{\alpha_1}{argmax} \sum_{i=1}^{n} \sum_{k=1}^{K} Prob[k | Y, X, Z, C; \tilde{\gamma}, \tilde{\alpha}, \tilde{p}] \ \ln(L_{i1k}(\tilde{\tilde{\gamma}}, \tilde{\tilde{\alpha}}_3, \tilde{\tilde{\alpha}}_2, \alpha_1))$$

Then we update the conditional probabilities

$$Prob[k|Y,X,Z,C;\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}},\tilde{p}] =$$

$$= \frac{\tilde{p}_k \; L_{4k}(\tilde{\tilde{\gamma}})L_{3k}\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}}_3)L_{2k}(\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}}_3,\tilde{\tilde{\alpha}}_2)L_{1k}(\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}}_3,\tilde{\tilde{\alpha}}_2,\tilde{\tilde{\alpha}}_1)}{\sum_{k=1}^{K}\tilde{p}_k \; L_{4k}(\tilde{\tilde{\gamma}})L_{3k}(\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}}_3)L_{2k}(\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}}_3,\tilde{\tilde{\alpha}}_2)L_{1k}(\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}}_3,\tilde{\tilde{\alpha}}_2,\tilde{\tilde{\alpha}}_1)}$$

And we update the unconditional probabilities

$$\tilde{\tilde{p}}_k = \frac{1}{n}\sum_{i=1}^{n} Prob[k|Y_i,X_i,Z_i,C_i;\tilde{\tilde{\gamma}},\tilde{\tilde{\alpha}},\tilde{p}]$$

Then we iterate to the maximization step with the last values of the parameters.

We consider the estimation of discrete choice models. Two approches are often used : MLE and Method of Simulated Moments (MSM). Simulation based estimation methods are particularly adapted to estimate structural models because these methods are computationally easier to implement.

The MSM estimation consists to compare a vector of moments calculated on the observed sample with a vector of moments resulting from the theoretical model and draws for the error terms of the model.

We are going to consider the two methods of estimation using a simple **dynamic discrete choice model** of schooling. It is a simple model - but realistic - in term of number of alternatives, number of periods. Agent have to take only a **binary decision** for each period of time. In this case, the MLE is still tractable (more generally, one will have to use simulations or interpolation techniques).

Let us consider an illustration of the sequence of decisions the agent has to take.

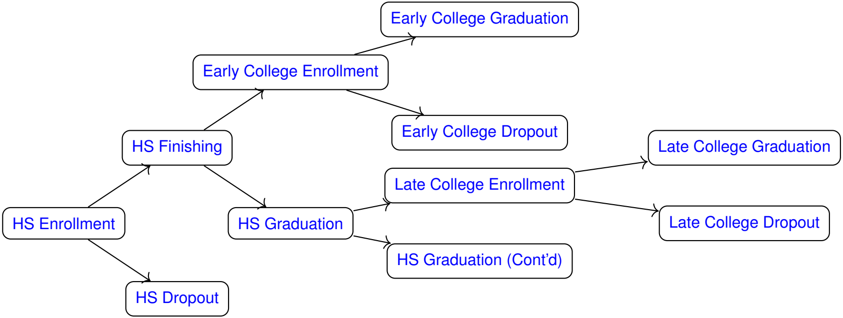

Figure: Decision tree (HS: High School)

The current state is denoted s, $s \in S = \{1 \ldots, K\}$. Let $\Omega(s)$ denote the choice set of an agent in state s: it is the list of feasible states that the agent can decide to move into from state s. The agent is assumed to have only binaries choice - at most - at each decision node ($Card(\Omega(s)) \leq 2$).

Ex post, an agent who decide to move from state s to state s', obtains per period rewards

$$Y(s') - C(s, s')$$

where $Y(s')$ is the per period **earnings** and $C(s, s')$ is the **costs** associated to moving from state s to state s'. These costs can include monetary costs (housing, tuition, travel expenses) and psychic costs (stress). The costs associated to one of the exit per node are set to **zero**. In what follow, we are going to distinguish a state with non zero cost $\dot{s}'$ with a state with zero cost $\bar{\mathbf{s}}'$ (we use a bold letter).

Let us consider an agent in state s. Let us denote $d(s)$ the number of periods spent by the agent in that state.

For instance, for an individual who spent 4 years in College, **the duration** of the college enrollment state is then $d(s) = 4$ years.

The duration of the counterfactual state is fixed to the median of the durations among all agents who transit by that state. The duration of the terminal state is calculated up to age 65.

Let $Y(s, t)$ denote **the earnings** of an agent in state s at time t. All the values $Y(s, t)$ corresponding to state s for an agent are collapsed into the average

$$Y(s) = \frac{\sum_{t=1}^{d(s)} \frac{Y(s,t)}{(1+r)^{t-1}}}{\sum_{t=1}^{d(s)} \frac{1}{(1+r)^{t-1}}}$$

where the **discount factor** $r = 0.04$. A similar operation is done for time varying covariates.

Per period earnings are

$$Y(s) = \mu_s(x(s)) + u_Y(s),$$

where $x(s)$ is a vector of observable covariates.
The moving costs when the agent is going from state s to state $\dot{s}'$ are

$$C(s, \dot{s}') = K_{s,\dot{s}'}(Q(s, \dot{s}')) + u_C(s, \dot{s}'),$$

where $Q(s, \dot{s}')$ is a vector of observable characteristics.
The error term have the following structure

$$\begin{aligned} u_Y(s) \quad &= \quad \theta'\alpha_s + \epsilon_s, \\ u_C(s, \dot{s}') &= \quad \theta'\varphi_{s,\dot{s}'} + \eta_{s,\dot{s}'}, \end{aligned}$$

where θ is a vector of individual **latent factors**. These latent factors are unknown by the econometrician but known by the agent.

The idiosyncratic terms ϵ_s and $\eta_{s,\dot{s}'}$ are **known by the agent only at the time of the decision** to move from state s to state $\dot{s}'$, i.e. only for the ongoing period of time. The error term $\epsilon_{\dot{s}'}$ is assumed not to be known by the agent when he decides whether to move from state s to state $\dot{s}'$. These terms are unknown for the econometrician and are assumed to be independent.

α_s and $\varphi_{s,\dot{s}'}$ are vectors of parameters. These vectors are considered as **factor loadings**. The presence of latent factors allows to correlate earnings and costs via these unobservable variables. These latent factors can be interpreted as **individual specific traits**.

We assume the existence of J **individuals measures** that allow to **proxy** the latent factors θ

$$M_j = \mu_j(X(K+j)) + \theta'\gamma_j + \nu_j,$$

where $J \in M = \{1, \ldots, J\}$.

A dynamic discrete choice model (Eisenhauer, Heckman and Mosso, 2015)

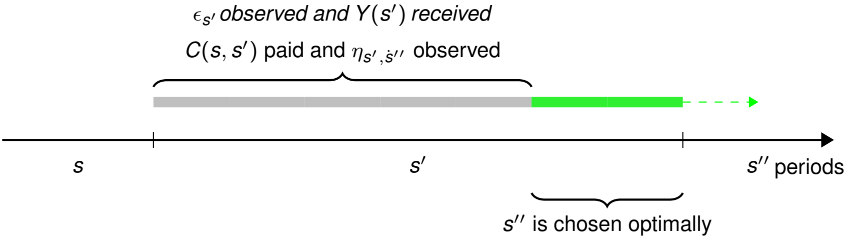

Figure: Timing of the arrival of information

Let us assume that the agent is **risk neutral** and maximizes the discounted lifetime rewards.

Let us denote $I(s)$ the information set of the agent in state s. $I(s)$ includes the following informations

- Error terms $\eta_{s,\dot{s}'}$ and ϵ_s for all s already visited and for all $\dot{s}'$ that can be reached from s,
- Vectors of explanatory variables $X(s)$ and $Q(s, \dot{s}')$ for all state s,
- Vectors of characteristics $X(K + j)$ for all $j = 1, \ldots, J$
- The latent factors θ.

Remark : The error term of the earnings equation for the next period $\epsilon_{\dot{s}'}$ is not included in $I(s)$ as it is not known by the agent (for all states $\dot{s}'$ that can be visited from state s).

Let $F_{\epsilon_s}(\epsilon_s)$ denote the cumulative distribution function of the **earning shocks** ϵ_s and let $F_{\eta_{s,\dot{s}'}}(\eta_{s,\dot{s}'})$ denote the cdf of the transition costs $\eta_{s,\dot{s}'}$.

The value of **discounted lifetime earnings** given the information available in state s is

$$V(s) = Y(s) + \max_{s' \in \Omega(s)} \left\{ \frac{1}{1+r}(-C(s,s') + E[V(s') \mid I(s)]) \right\}$$

Let us define the **continuation value** associated to state s as follow

$$CV(s) = \max_{s' \in \Omega(s)} \left\{ \frac{1}{1+r}(-C(s,s') + E[V(s') \mid I(s)]) \right\}$$

An agent who occupies state s decide to move into the state s' such that

$$s' = \begin{cases} \dot{s}' & \text{if } E[V(\dot{s}') \mid I(s)]) - C(s, \dot{s}') > E[V(\bar{\mathbf{s}}') \mid I(s)] - 0, \\ \bar{\mathbf{s}}' & \text{otherwise.} \end{cases}$$

Let us consider the following functional forms

$$\begin{aligned} M_j &= x(K+j)'\kappa_j + \theta'\gamma_j + \nu_j, \\ Y(s) &= x(s)'\beta_s + \theta'\alpha_s + \epsilon_s, \\ C(s, \dot{s}') &= Q(s, \dot{s}')'\delta_{s,\dot{s}'} + \theta'\varphi_{s,\dot{s}'} + \eta_{s,\dot{s}'}, \end{aligned}$$

for all $s \in S$ and for all $j \in M$.

Let us consider the following assumptions for the unobservables of the model

$$\begin{array}{rr} \nu_j \sim N(0, \sigma_{\nu_j}), & \epsilon_s \sim N(0, \sigma_{\epsilon_s}) \\ \theta_k \sim N(0, \sigma_\theta), & \eta_{s,\dot{s}'} \sim N(0, \sigma_{\eta_{s,\dot{s}'}}) \end{array}$$

for all $s \in S$ and for all $j = K+1, \ldots, K+J$ and $k = 1, 2$ (**two factors**). K is the number of states.

The unobservable variables ϵ_s, $\eta_{s,\dot{s}'}$ and ν_j are independent across measures (j) and states (s). The two factors are independent from each others and from other unobservables.

We can observe correlations across outcomes. For instance

$$cov(Y(s), Y(s') \mid X) = cov(\theta' \alpha_s, \theta' \alpha_{s'} \mid X) = (\alpha_{s,1}\alpha_{s',1} + \alpha_{s,2}\alpha_{s',2})\sigma_\theta^2$$

Consequently,

$$Corr(Y(s), Y(s') \mid X) = \frac{(\alpha_{s,1}\alpha_{s',1}\sigma_\theta^2 + \alpha_{s,2}\alpha_{s',2}\sigma_\theta^2)}{\sqrt{\alpha_{s,1}^2\sigma_\theta^2 + \alpha_{s,2}^2\sigma_\theta^2}\sqrt{\alpha_{s',1}^2\sigma_\theta^2 + \alpha_{s',2}^2\sigma_\theta^2}}$$

Using the functional forms, the **agent chooses state** $\dot{s}'$ iff

$$Y(\dot{s}') + CV(\dot{s}') - Q(s, \dot{s}')'\delta_{s,\dot{s}'} - \theta'\varphi_{s,\dot{s}'} - \eta_{s,\dot{s}'} > Y(\bar{\mathbf{s}}') + CV(\bar{\mathbf{s}}'),$$

and, then

$$x(\dot{s}')'\beta_{\dot{s}'} + \theta'\alpha_{\dot{s}'} + E[\epsilon_{\dot{s}}|I(s)] + CV(\dot{s}') - Q(s, \dot{s}')'\delta_{s,\dot{s}'} - \theta'\varphi_{s,\dot{s}'} - \eta_{s,\dot{s}'}$$

$$> x(\bar{\mathbf{s}}')'\beta_{\bar{\mathbf{s}}'} + \theta'\alpha_{\bar{\mathbf{s}}'} + E[\epsilon_{\bar{\mathbf{s}}}|I(s)] + CV(\bar{\mathbf{s}}'),$$

where $E[\epsilon_{\dot{s}}|I(s)] = E[\epsilon_{\bar{\mathbf{s}}}|I(s)] = 0$.

Let ψ denote the vector of parameters of the structural model and $\mathcal{S}$ the subset of the **states visited** by the agent. Let G_s denote a **binary variable** that is equal to 1 if the agent has visited state s ($G_s = 0$ otherwise). Let D dente the vector of the **observed characteristics** of the agent.

A **contribution** of a given agent to the likelihood function is

$$\int_{\Theta} \prod_{j \in M} f(M_j | D, \theta, \psi) \prod_{s \in S} \left[f(Y(s) | D, \theta; \psi) P[G_s{=}1 | D, \theta; \psi] \right]^{\mathbf{1}[s \in \mathcal{S}]} dF(\theta)$$

where Θ is the support of the distribution of θ ($\Theta = \boldsymbol{R}^2$).

Within the integral, the first term represents the conditional density of the **measurements** given the **factors** and observable explanatory variables. The second term is the conditional density of the **earnings**. The last term is the conditional probability that the individual occupies **state** s. Le log of the likelihood function is the sum of the logarithm of the contributions of the agents.

The conditional density function for the **measurement variable** M_j is

$$f(M_j|D,\theta,\psi) = \phi_{\sigma_{\nu_j}}(M_j - x(K+j)'\kappa_j - \theta'\gamma_j),$$

where $j \in M$ and $\phi_\sigma(.)$ denote the pdf of the normal distribution with mean 0 and variance $\sigma^2_{\nu_j}$.

The conditional density function for **the earnings** $Y(s)$ is

$$f(Y(s)|D,\theta,\psi) = \phi_{\sigma_{\epsilon_s}}(Y(s) - x(s)'\beta_s - \theta'\alpha_s),$$

where $s \in S$.

The pdf of the distribution of the **unobserved factors** is

$$f(\theta) = \prod_{k=1}^{2} \phi_{\sigma_\theta}(\theta_k)$$

where $\phi_{\sigma_\theta}(\theta_k) = \frac{1}{\sigma_{\sigma_\theta}\sqrt{2\pi}} \exp(-\frac{\theta_k^2}{2\sigma_\theta^2})$.

An agent who occupies state s has to decide whether he moves to the state $\dot{s}'$ (the transition is costly) or to move to state $\bar{\mathbf{s}}'$ (no cost for the transition).

Its ex ante **valuation** $T(s,s')$ takes into account the **expectation of transition costs**, the expectation of the earnings and the continuation value $CV(s')$. The ex ante **value of state** s' has the expression

$$T(s,s') =$$

$$\begin{cases} x(\dot{s}')'\beta_{\dot{s}'} + \theta'\alpha_{\dot{s}'} + CV(\dot{s}') - Q(s,\dot{s}')'\delta_{s,\dot{s}'} - \theta'\varphi_{s,\dot{s}'} & \text{if } s' = \dot{s}', \\ x(\bar{\mathbf{s}}')'\beta_{\bar{\mathbf{s}}'} + \theta'\alpha_{\bar{\mathbf{s}}'} + CV(\bar{\mathbf{s}}') & \text{if } s' = \bar{\mathbf{s}}'. \end{cases}$$

Let us remark that we have assumed that the agent, for each node, has only **two alternatives** denoted $\dot{s}'$ and $\bar{\mathbf{s}}'$.

The conditional transition probability from state s to state s' is

$$P[G_{s'} = 1 \mid D, \theta; \psi] = \begin{cases} \Phi_{\sigma_{\eta_{s,\dot{s}'}}}(T(s,\dot{s}') - T(s,\bar{\mathbf{s}}')) & \text{if } s' = \dot{s}', \\ 1 - \Phi_{\sigma_{\eta_{s,\dot{s}'}}}(T(s,\dot{s}') - T(s,\bar{\mathbf{s}}')) & \text{if } s' = \bar{\mathbf{s}}'. \end{cases}$$

where Φ_σ is the cdf of the normal distribution with mean 0 and variance σ^2.

The **continuation value** of state s is

$$CV(s) = \int_{-\infty}^{T(s,\dot{s}') - T(s,\bar{\mathbf{s}}')} [T(s,\dot{s}') - \eta]\, \phi_{\sigma_{\eta_{s,\dot{s}'}}}(\eta)\, d\eta$$
$$+ [1 - \Phi_{\sigma_{\eta_{s,\dot{s}'}}}(T(s,\dot{s}') - T(s,\bar{\mathbf{s}}'))]\, T(s,\bar{\mathbf{s}}')$$

Let us remark that the agent decides **to move** to the costly state $\dot{s}'$ **if and only if** $T(s,\dot{s}') - \eta_{s,\dot{s}'} > T(s,\bar{\mathbf{s}}')$

In order to maximize the **likelihood function**, the integrals can be evaluated using Gaussian quadrature (approximation errors). The log-likelihood function can be maximized using an optimization algorithm such as BFGS.

The **Method of Simulated Moments** approach:

The parameters of the model are chosen such as to minimize the weighted distance between **a collection of moments** obtained from the observed sample and a sample simulated using the model.

The objective function is

$$\varphi(\psi) = (\check{f} - \hat{f}(\psi))' W^{-1} (\check{f} - \hat{f}(\psi))$$

where $\check{f}$ is a vector of moments obtained from the observed data. $\hat{f}(\psi)$ is the vector of the average moments calculated from R **simulated data sets obtained from the model**. W is a positive definite symmetric weighting matrix.

The simulation is repeated R times. Consequently, we generate R samples of size $n = 5000$. Let $u_{r,i} = (\epsilon_{s,r,i}, \eta_{s,\dot{s}',r,i},\ \textit{for all } s \in S)$ a set of unobserved components of the model for a given agent i (θ can be taken into account by estimating a vector of factors).

For instance, **for the example** (see decision tree), as we have 11 **states** and 5 **transitions costs**. $u_{r,i}$ determines the agent's earnings and choices. $u_{r,i}$ is a fonction of the parameters of the model as

$$\epsilon_s = \sigma_{\epsilon_s}\ v_s, \text{ and } \eta_{s,\dot{s}'} = \sigma_{\eta_{s,\dot{s}'}}\ v_{s,\dot{s}'}$$

where $v_s \overset{iid}{\sim} N(0,1)$ and $v_{s,\dot{s}'} \overset{iid}{\sim} N(0,1)$.
Let $u_{r,i} \in \boldsymbol{R}^{16}$. Let $u_r = (u_i, i = 1, \ldots, n)$ denote a vector of unobserved components for a given replication. The number of replications R can be fixed to 30.

The random drawings v_s and $v_{s,\dot{s}'}$ **are fixed** throughout the optimization of the objective function (do not depend on ψ).

Let us now consider the expression of the average of the moment on the simulated samples

$$\hat{f}(\psi) = \frac{1}{R}\sum_{r=1}^{R}\hat{f}(u_r;\psi)$$

In practice, as the drawings $(v_s, v_{s,\dot{s}'}, s \in S)$ are fixed at the beginning, the objective function $\varphi(\psi)$ varies only via the vector of parameters ψ.

u_r is a collection of 5000 **"education careers"** that are used to estimate the vector of moments. The model is solved by **backward induction** for each simulated sample and each agent.
θ is replaced by an estimated vector of factors obtained from $M_1, \ldots, M_J$ (see Bartlett, 1937). These estimated vector of factors is then treated as other explanatory variables in the estimation of the parameters.

An alternative way to do (see Del Boca, Flinn and Wiswall, 2014) is to **simulate the latent factors** taking into account the assumptions made on their data generating process.

The identification of the model is considered by Eisenhauer, Heckman and Mosso (2015). **All the parameters** of the models **are identified** including the discount rate.

Singularly, they note that the measurement system is not necessary to identify the factors' distributions if number of states plus the number of transitions is greater than $2\times$ (numbers of factors)+1. However, the **measurement system** allows to increase the **precision**. Identification of linear measurement models has been examined by Williams (2016).

Let I denote the number of moments. The weighting matrix W is a diagonal matrix with the variance of these moments on the main diagonal and zero otherwise. These variances are estimated by resampling the observed data 200 times.

The objective function is then

$$\varphi(\psi) = \sum_{i=1}^{I} \left(\frac{\check{f}_i - \hat{f}_i(\psi)}{\hat{\sigma}_i} \right)^2$$

where f_i is the moment i and σ_i is the estimation of the standard deviation of moment i obtained from bootstrap.

This objective function is **not smooth** with respect to the parameters of the model.

Indeed, a small modification of the parameters can induce **some agents** to modify their choices.

Finally, a small modification of the parameters can induce a discrete jump of the moments.

One cannot use gradient based method to minimize the objective function such as BFGS algorithm.

One should use a derivative free algorithm like the Practical Optimization Using No Derivatives for sums of Squares algorithm (POUNDerS algorithm).

It is an algorithm for solving the least squares problem when the vector of derivatives of the objective function is not available.

POUNDerS exploit the characteristics of the non linear least squares problem to use derivative-free trust-region framework in order to converge to local minimizers of the objective function (see Munson et al., 2012).

Let us consider a more general set of notations. In dynamic programming problems, it usual to distinguish the **state variable** z_t, and a **decision variable** d_t (or "control variable"). For Eisenhaueur, Heckman and Mosso (2015), the state variables are the unobserved components of earnings and costs. The decision d_t is taken at the end of period $t-1$ (at this time z_t is unknown).

The agent has a **utility** (or profit) denoted $U(z_1, d_1, \dots, z_T, d_T)$ where T is the horizon.

The $I_{t-1} = (z_1, d_1, \dots, z_{t-1}, d_{t-1})$ denote the **history** at time $t-1$. Let $p_t(z_t \mid I_{t-1})$ denote the decision-dependent conditional probabilities. **Let us assume** that the agent decide to choose d_t after observing z_{t-1}. For Eisenhaueur, Heckman and Mosso (2015), the **decision** at the end of period $t-1$ is $d_t = s_t$ who is the state occupied during period t (see the timing of decisions) and the original state is not an element that the agent has to decide as they are all enrolled in high school.

The agent's decision d_t is restricted to the set $S_t(I_{t-1})$.
Let us assume that the agent maximizes its expected utility.

The agent considers history dependent decision rules $(\delta_1, \ldots, \delta_T)$. For each time t, the decision is a function of all information the agent has. Finally, at time t the **decision** is $d_t = \delta_t(I_{t-1})$ (decision taken at the end of period $t-1$).

We assume that $I_0 = \emptyset$. For, Eisenhaueur et al. 2015, all agent starts their educational career in the same way and later decisions are taken given this information that is the same for all agents.

A **feasible decision** is such that $\delta_t(I_{t-1}) \in S_t(I_{t-1})$ for all I_{t-1}.
An optimal decision rule $\delta^* = (\delta_1^*, \ldots, \delta_T^*)$ is such that

$$\delta^* = \underset{\delta \in \mathcal{F}}{argmax}\, E[U(\{\tilde{z}_t, \tilde{d}_t\}_\delta)]$$

where $\mathcal{F}$ is the set of all dependent-history feasible decision rules.

$\{\tilde{z}_t, \tilde{d}_t\}_\delta$ denote the stochastic process associated to the decision rule δ (indeed $p_t(z_t \mid I_{t-1})$ depend of the past decisions.

Backward induction :

We start the resolution at the last period T. For each I_{T-1} we calculate the value function V_T and decision rule δ_T.

$$V_T(I_{T-1}) = \max_{d_T \in S_T(I_{T-1})} E[U(I_{T-1}, \tilde{z}_T, d_T) \mid I_{T-1}]$$

$$\delta_T(I_{T-1}) = \underset{d_T \in S_T(I_{T-1})}{argmax} E[U(I_{T-1}, \tilde{z}_T, d_T) \mid I_{T-1}]$$

Consequently, the optimal decision for period T is a function of the history I_{t-1}.

Remark: For Eisenhauer et al. (2015), the horizon is $T = 5$. d_T is the state occupied during period T, d_{T-1} is the state occupied at period $T - 1$, ..., d_1 is the state occupied at period 1 and this initial state is fixed ($d_1 =$"High School Enrollement").

If we consider the decision tree for this example, at period $t = 4$, the decision we consider is taken given the individual occupies state "Late college enrollment". In this case, $S_T(I_{T-1}, z_T) = \{"\text{Late College Graduation}", "\text{Late College Dropout}"\}$.

At period $t = 3$ we have two decisions to consider:

- One decision given the agent occupies state "Early college enrollment" (he has two feasible alternatives "Early college graduation" and "Early college dropout")
- One decision given he occupies state "High school graduation" (the agent has two feasible decisions "Late college enrollment" and "High School Graduation C'd").

Then, we **move backward** to period $T - 1$.

For each I_{T-2} we calculate the value function V_{T-1} and decision rule δ_{T-1}.

$$V_{T-1}(I_{T-2}) = \max_{d_{T-1} \in S_{T-1}(I_{T-2})} E[V_T(I_{T-2}, \tilde{z}_{T-1}, d_{T-1}) | I_{T-2}]$$

$$\delta_{T-1}(I_{T-2}) = \mathop{argmax}_{d_{T-1} \in S_{T-1}(I_{T-2})} E[V_T(I_{T-2}, \tilde{z}_{T-1}, d_{T-1}) | I_{T-2}]$$

where the expectation is taken with respect to the density of z_{T-1} (z_{T-1} **unknown at the end of period** $T-2$). Let us remark that $I_{T-1} = (I_{T-2}, z_{T-1}, d_{T-1})$.

We continue the backward induction for $T-2$, ..., $T=1$.

For the period t, the **value function** V_t and decision rule δ_t are solultion of

$$V_t(I_{t-1}) = \max_{d_t \in S_t(I_{t-1})} E[V_{t+1}(I_{t-1}, \tilde{z}_t, d_t) | I_{t-1}]$$

$$\delta_t(I_{t-1}) = \underset{d_t \in S_t(I_{t-1})}{argmax} E[V_{t+1}(I_{t-1}, \tilde{z}_t, d_t) | I_{t-1}]$$

where the equation for the value V_t is called the **Bellman equation**.

We continue the backward induction for $T - 2, \ldots, T = 1$.

The **backward induction generates an optimal decision rule**.

In many economic application, the utility function is assumed to be time-separable

$$U(z_1, d_1, \ldots, z_T, d_T) = \sum_{t=1}^{T} \beta^t u_t(z_t, d_t)$$

where β is a discount factor ($\beta \in (0, 1)$).

Arcidiacono, P., Jones, J.B., 2003. Finite mixture distributions, sequential likelihood and the EM algorithm. Econometrica, 71(3), 933-946.

Bartlett, M.S, 1937. The statistical conception of mental factors, British Journal of Psychology,1, 97-104.

Blevins, J.R., 2016. Sequential Monte Carlo methods for estimating dynamic microeconomics models, Journal of Applied Econometrics, 31, 773-804.

Cameron, S.V., Heckman, J.J., 2001. The dynamics of educational attainment for black, hispanic, and white males. Journal of Political Economy, 109(3), 455-499.

Del Boca, D., Flinn, C., Wiswall, M, 2014. Household choices and child development. Review of Economic Studies, 81, 137-185.

Heckman, J.J., Singer, B., 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica, 52(2), 271-320.

Eisenhauer, P., Heckman, J.J., Mosso, S., 2015, Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. International Economic Review, 56(2), 331-357.

McFadden, D., Train, K., 2000. Mixed MNL models of discrete response. Journal of Applied Econometrics,15, 447?470.

Rust, J., 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. Econometrica, 55(5), 999-1033.

Rust, J., 2008. Dynamic Programming. in The New Palgrave Dictionary of Economics, 2nd edition. Edited by S. N. Durlauf and L. E. Blume.

Ruud, P.A., 2007. Estimating mixtures of discrete choice Models. University of California, Berkeley.

Williams, B., 2016. Identification of the Linear Factor Model. RPF Working Paper No. 2018-002.