

# Odhad s kategorickým modelem

**Úloha:** Odhad parametrů modelu pro jeho další použití.

- vstup i výstup jsou binární (mají hodnoty 1 a 2)
- model s pamětí a řízením  $f(z_t|u_t, z_{t-1})$

Jako data se využívá buď simulace z datového souboru `dataT12.dat` (generovaného souborem `T12SimCat`) nebo reálných dat `dataCat.dat` - diskretizované hodnoty plyn a moment motoru. Reálná rada se generují souborem `generDatCat.sce`, kde lze získat podrobnější informace o datech. Oba datové soubory jsou uloženy v adresáři `_data`. Nic ale nebrání tomu, vyrobit si vlastní data - buď přímo v souboru `T12SimCat` nebo si data uložit na disk a natáhnout si je tak, jak je to připravené.

Vlastní odhad se provádí postupným přepočtem statistiky  $V$  (matice  $4 \times 2$ ), na začátku nulová nebo s apriorními statistikami. Přepočet je

$j = \text{psi2row}([u(t), z(t-1)], [2, 2])$  - číslo řádku tabulky

$V(j, z(t)) = V(j, z(t)) + 1$  - přepočet aktuálního prvku

**Předpoklady:** diskrétní vstup i výstup, matice  $V$  má obecně počet řádků roven součinu počtu hodnot veličin z regresního vektoru a počet sloupců roven počtu hodnot výstupu.

## Poznámka

*Apriorní hodnoty statistiky lze určit předběžným stanovením poměru koeficientů v jednotlivých řádcích statistiky, sílu počáteční informace zadáme absolutní velikosti čísel v apriorní statistice.*

## Značení

1.  $z$  -  $z_t$  (diskrétní výstup)
2.  $u$  -  $u_t$ ,
3.  $\theta$  -  $\theta$
4.  $\theta_E$  - odhadnuté parametry

## Volitelné parametry

- `I_dataSim` - data: simulovaná, reálná,
- (ostatní parametry se přebírají ze simulace)

## Doporučené experimenty

1. Testujte simulovaná data pro různé simulované soustavy. Zkuste porovnat odhad soustavy s hodně/málo neurčitosti.
2. Z reálných dat jsou k dispozici ještě "spotřeba" a "rychlost". Zkoušejte odhadovat model pro různé kombinace těchto veličin v roli výstupu a vstupu. Porovnejte výsledky.

## Program

```
// Estimation of categorical model  $f(y(t)|u(t),y(t-1))$  with  $y,u=1,2$ 
// --- on-line update of statistics, off-line point estimates
[u,t,n]=file(); // find working directory
chdir(dirname(n(1))); // set working directory
clear("u","t","n") // clear auxiliary data
exec("ScIntro.sce",-1),mode(0) // intro to sesion

I_dataSim=1; // get data: 1=simulated (in T12SimCat.sce)
// 2=real (from file dataCat.dat)

if I_dataSim==1
    load _data/dataT12.dat Sim // load of simulated data
    zt=Sim.Cz.zt;
    ut=Sim.Cz.ut;
    th=Sim.Cz.th;
else
    load _data/dataCat.dat sp co ga mo
    zt=mo; // volba diskretních
    ut=ga; // veličin (tady mo, ga)
end

nd=length(zt); // number of data
V=zeros(4,2); // statistics

for t=2:nd // time loop
    j=psi2row([ut(t),zt(t-1)],[2,2]); // row in parameter table
    V(j,zt(t))=V(j,zt(t))+1; // generation of y
end
thE=fnorm(V,2); // point estimates
Est.Cz.V=V;
Est.Cz.th=thE;

// results
if I_dataSim==1
    bar([th(:,1) thE(:,1)])
    title('Simulated and estimated parameters (first column of the model table)')
    legend('simulated','estimated');
else
    bar(thE(:,1))
    title('Estimated parameters - real data (first column)')
end

save _data/dataT22.dat Est
```