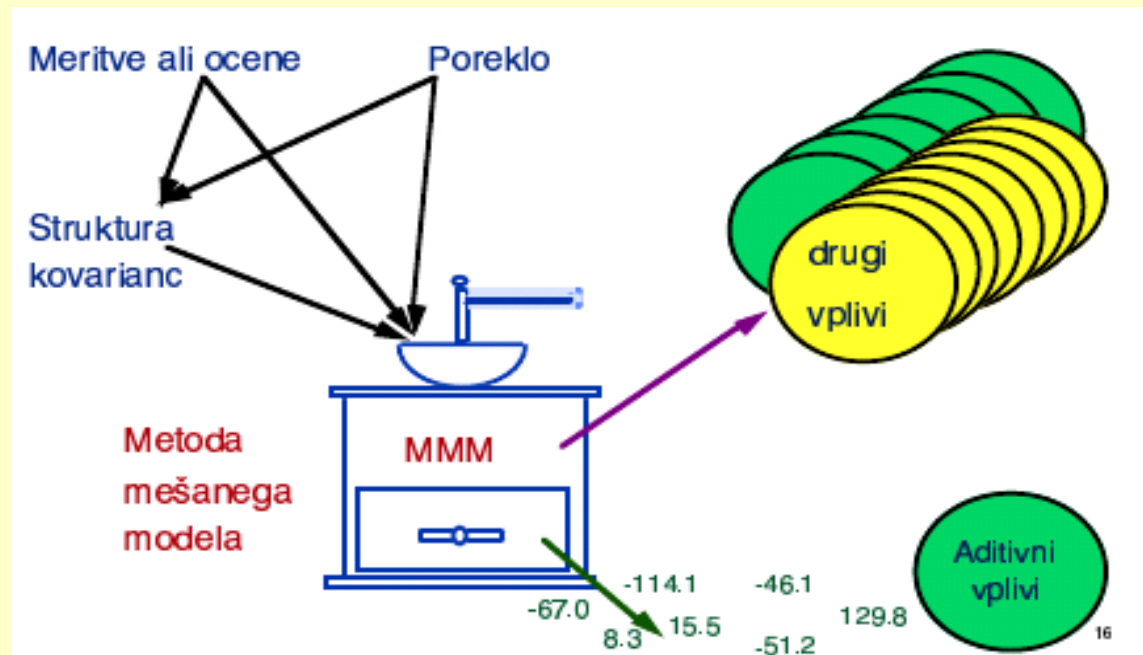


# Viri, kakovost in količina podatkov za točnost napovedi (G)PV



Milena KOVAČ, Špela MALOVRH  
Univerza v Ljubljani

# Napoved plemenske vrednosti in zanesljivost/točnost napovedi

$$\begin{bmatrix} \hat{\beta} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}$$

Variance napake napovedi  
(prediction error variance)

- Število opazovanj
- Sorodstvo
- Varianca za ostanek in heritabiliteta
- Korelacije med lastnostmi
- **PREDPOGOJ**: da so podatki pravilni!

# Viri podatkov

- Lastnost, merjena na živalih in pogoji, v katerih je nastajala
- Poreklo oz. rodovnik
- Korelirane lastnosti in komponente variance
- Genomske informacije
  - Geni z velikim učinkov
  - Označevalci
  - Snp-ji

# Kakovost meritev

- Identifikacija živali
- Način preizkusa
  - Preizkus lastne proizvodnosti
  - Preizkus sorodnikov (katerih in koliko)
  - Preizkus v pogojih reje
- Opravljanje meritev, pravilni odvzem vzorcev
- Zajemanje podatkov

Poskusi izboljšanja uvrstitev (“preferential treatment”)  
Površnost enega je težava za vse več let

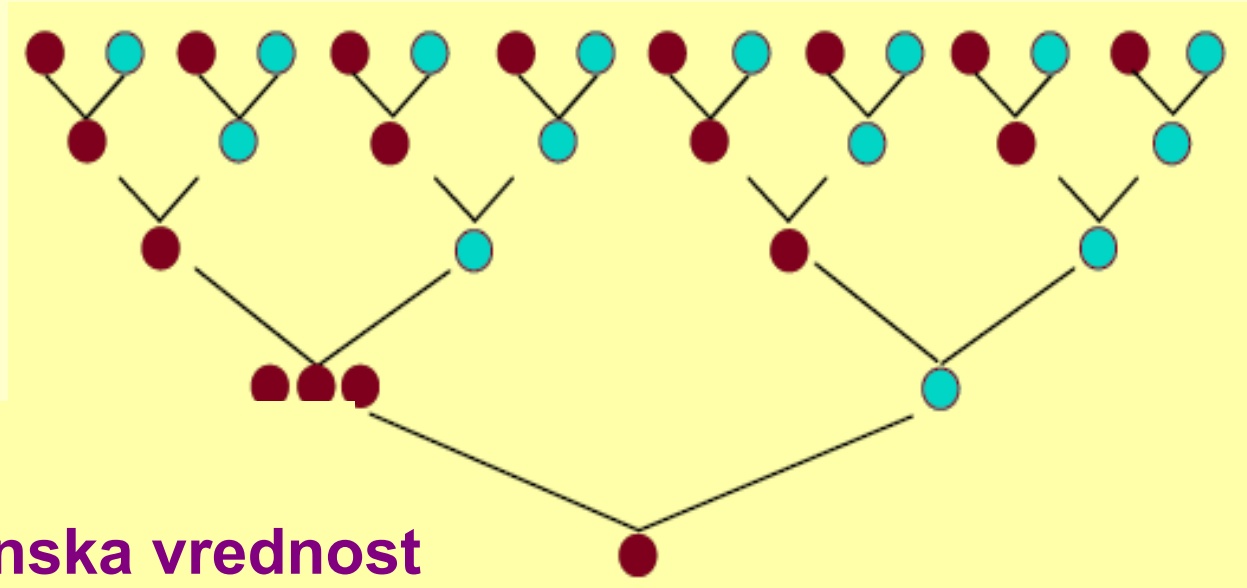
# Identifikacija

- Življenjska številka pri govedu, drobnici
- Podeljevanje rodovniških števil
- Označitev živali iz tujih populacij
- Ušesna številka pri plemenskih prašičih

ZP	U – štev.
6	xx-1396-81
	xx-1396-82
	xx-1396-83
7	xx-1396-79
	xx-1396-80
	xx-1396-81
	xx-1396-82
	xx-1396-83
	xx-1396-84
	xx-1396-85

Živali s podvojenimi številkami se zavrnejo kot plemenske živali.

# Uporaba informacij sorodnikov



Plemenska vrednost

-----

Lastni test

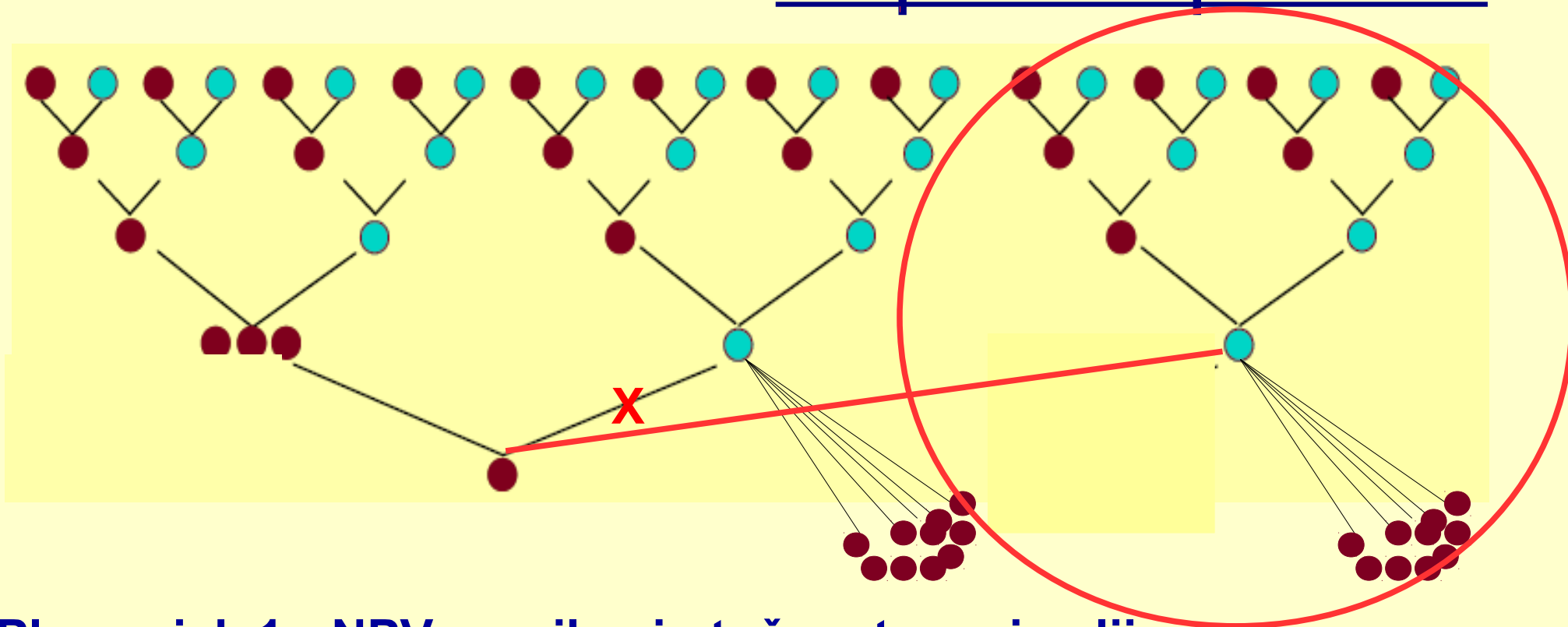
Test sorodnikov

Test prednikov

-----

Test potomcev

# Napake v poreklu



**Plemenjak 1 – NPV pravilna in točnost sprejemljiva**

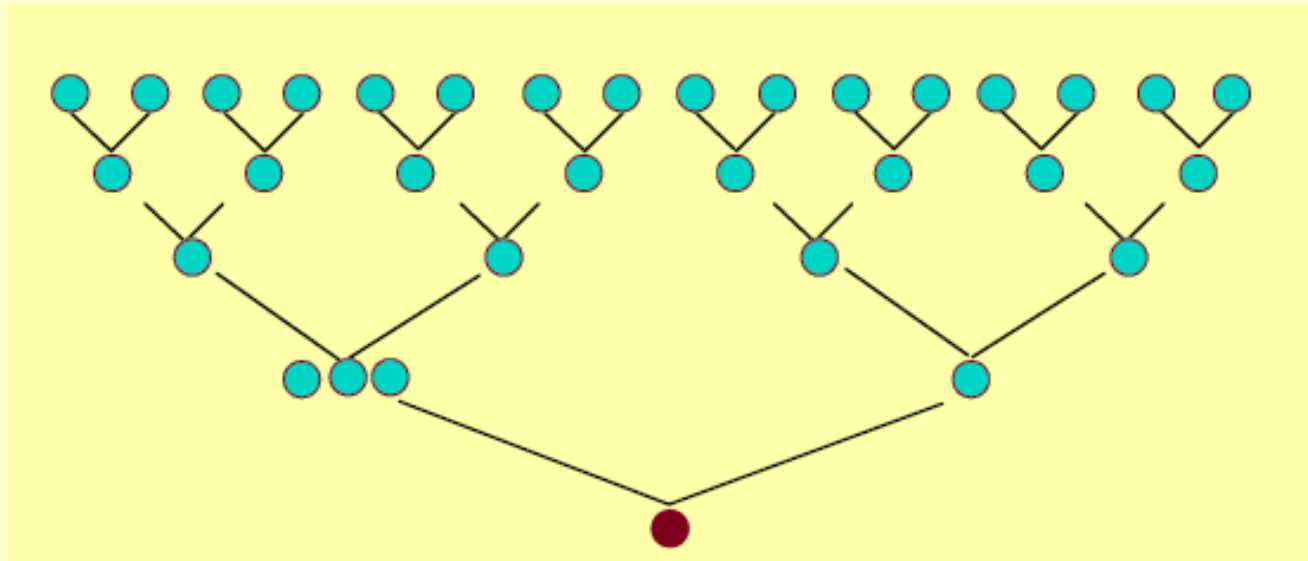
**Plemenjak 2 – NPV napačna in točnost sprejemljiva**

**Neznan – NPV pravilna, a točnost manjša  
(manj informacij)**

**Sumljivo starševstvo !!!!!!!!!!!**

- preverimo z genskimi preveritvami
- ali neznan in potomci niso plemenska žival

# Zloraba informacij sorodnikov

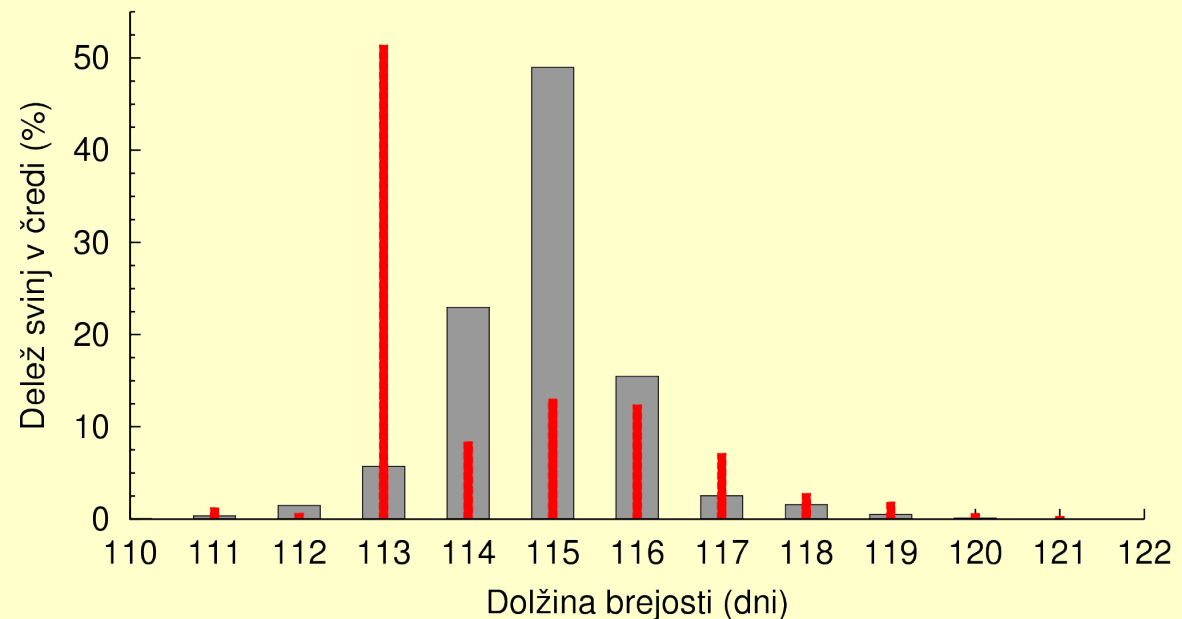


**Plemenska vrednost za vse sorodnike!!  
Napačni predniki!**

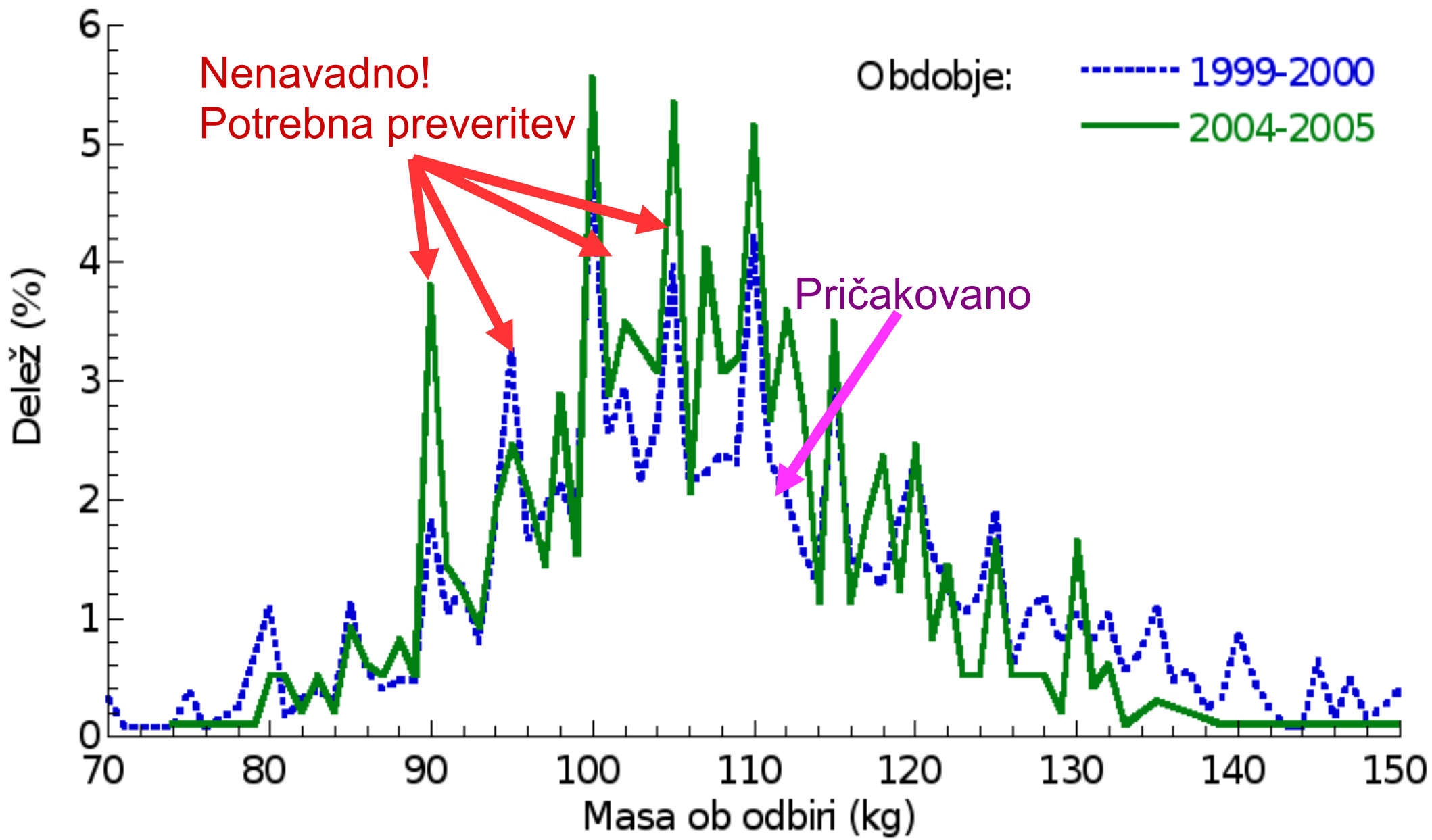


# Preizkus v pogojih reje

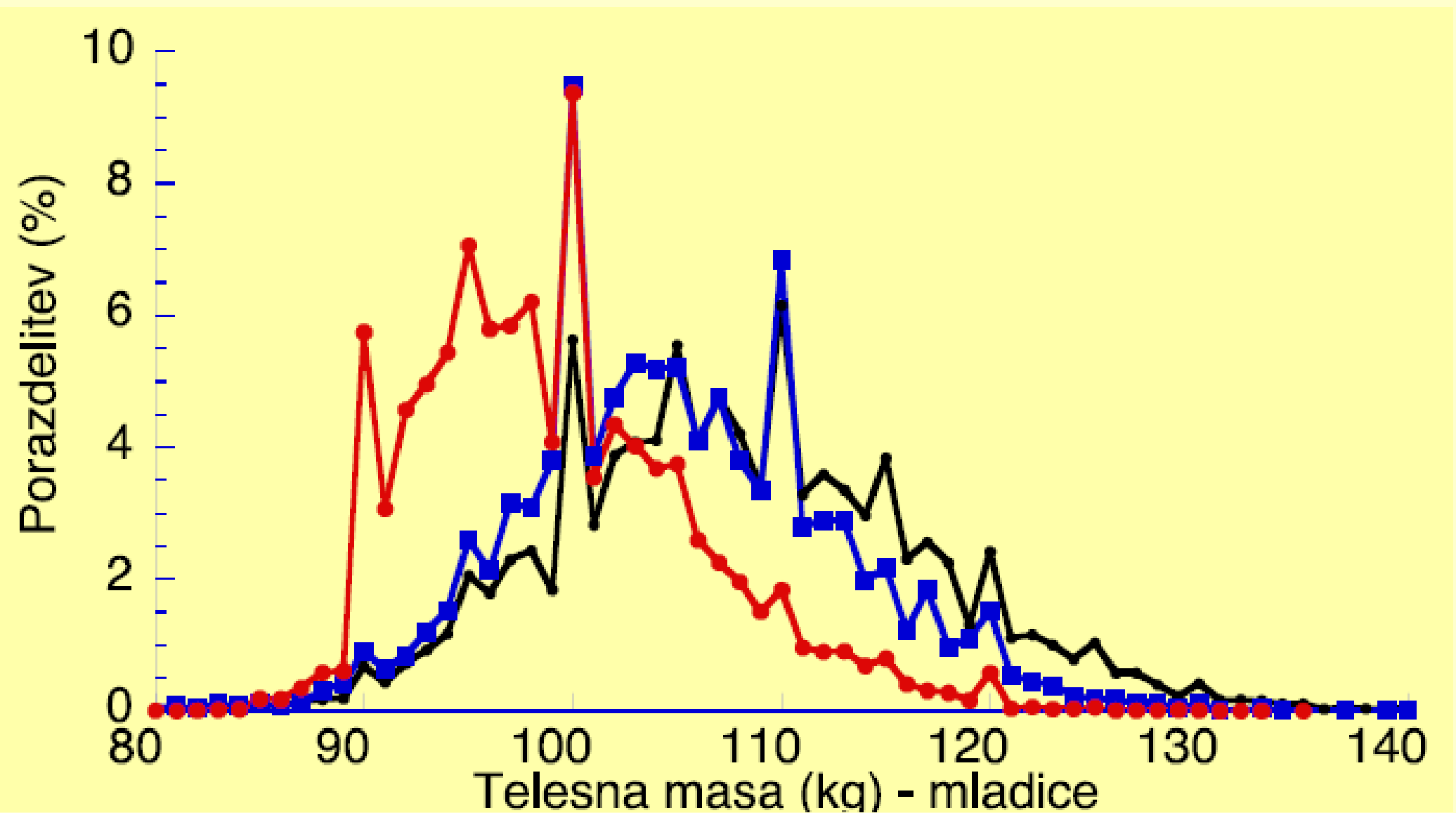
- Lahko je cenejši
- Lahko zberemo veliko podatkov
- Okolje je manj standardizirano kot na testnih postajah
- Več pomot
  - Slučajne napake – niso problem
  - Dopolnitve podatkov po logiki – katastrofa!
- Občasno želje po »posebni« obravnavi
- Usklajevanje ocenjevalcev



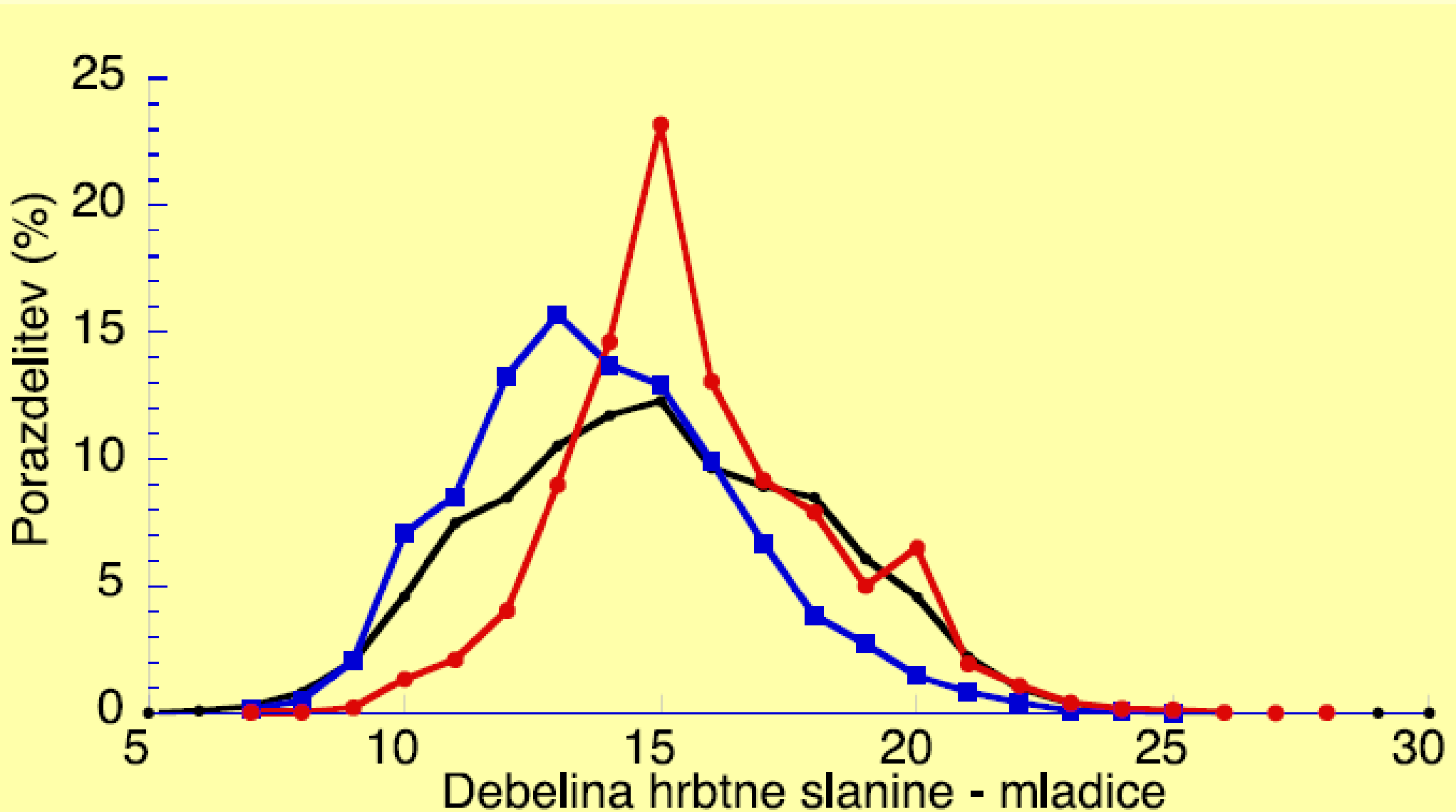
# Porazdelitev mase ob preizkusu mladice



# Porazdelitev mase pri preizkusu mladice po zavodih

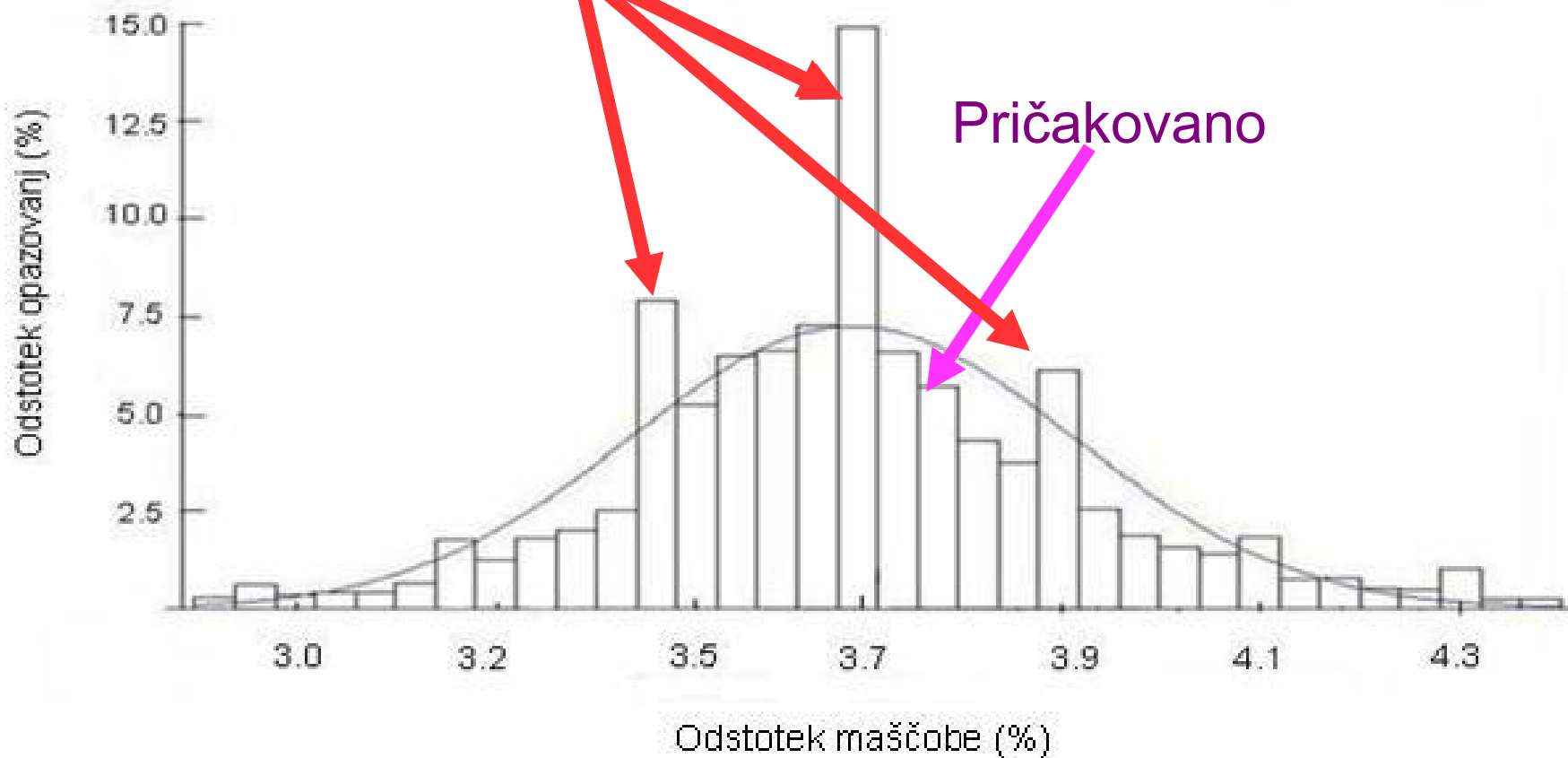


# Preizkus mladice - DHS

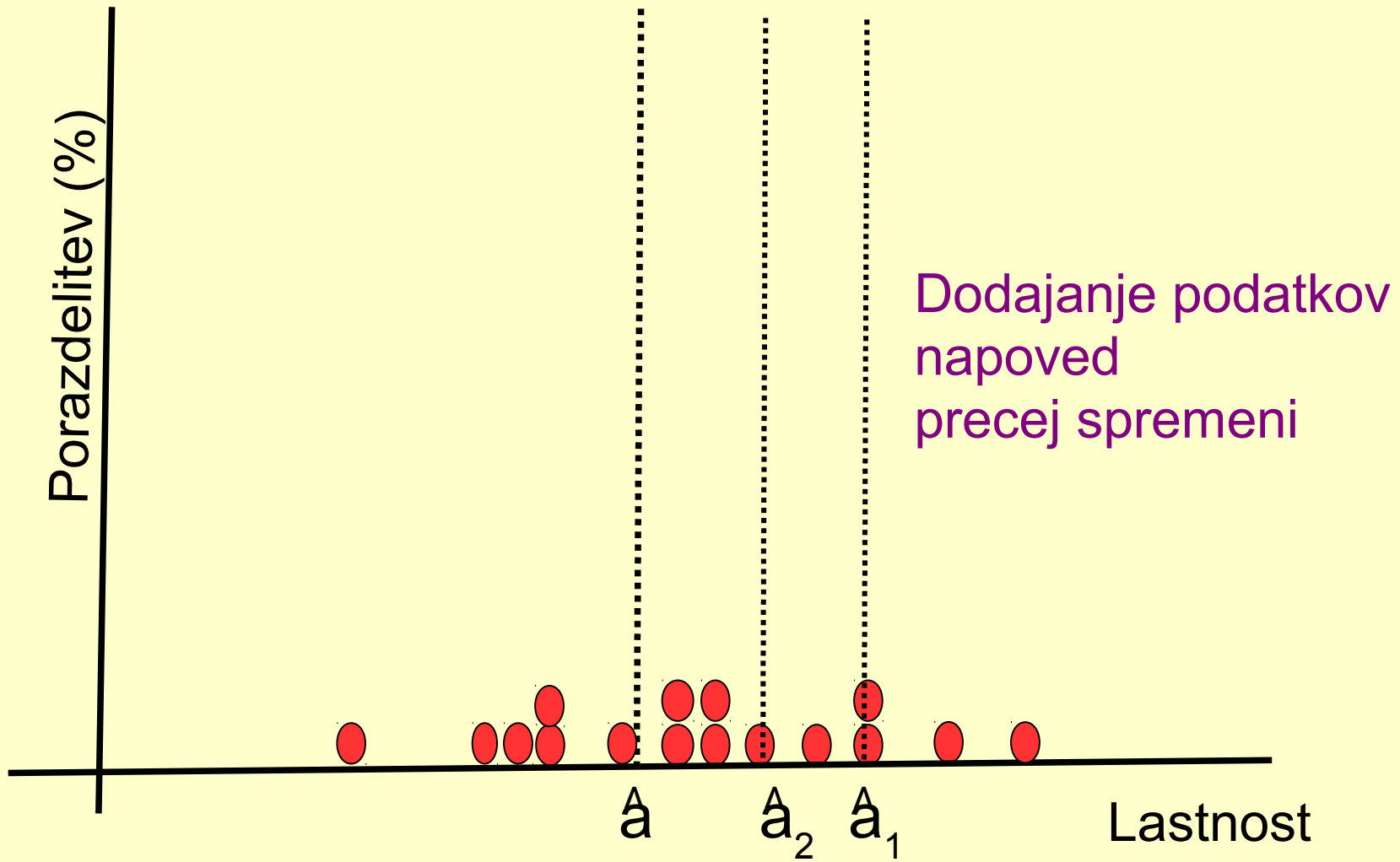


# Porazdelitev odstotka maščobe v mleku

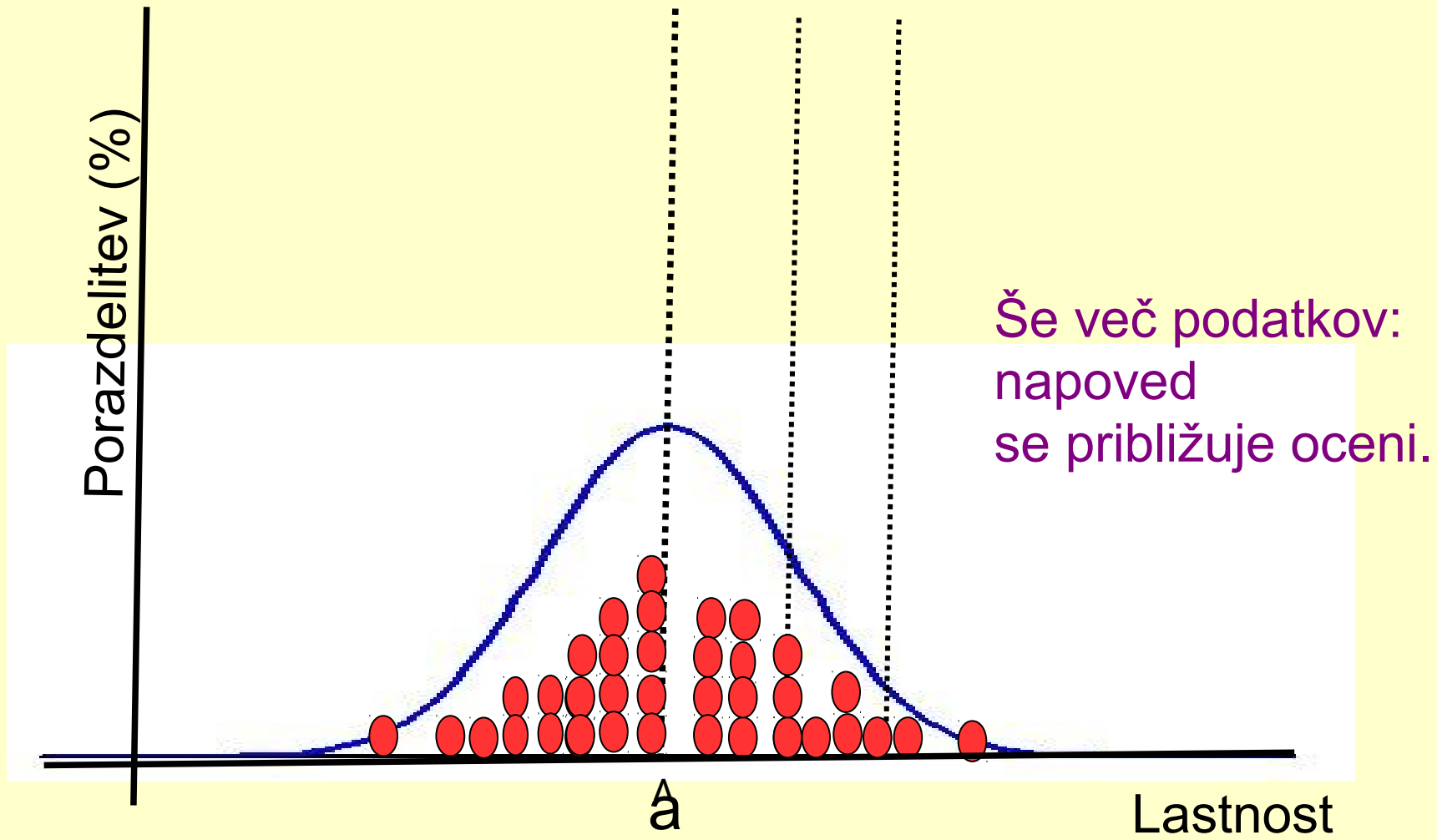
Nenavadno!  
Potrebna preveritev



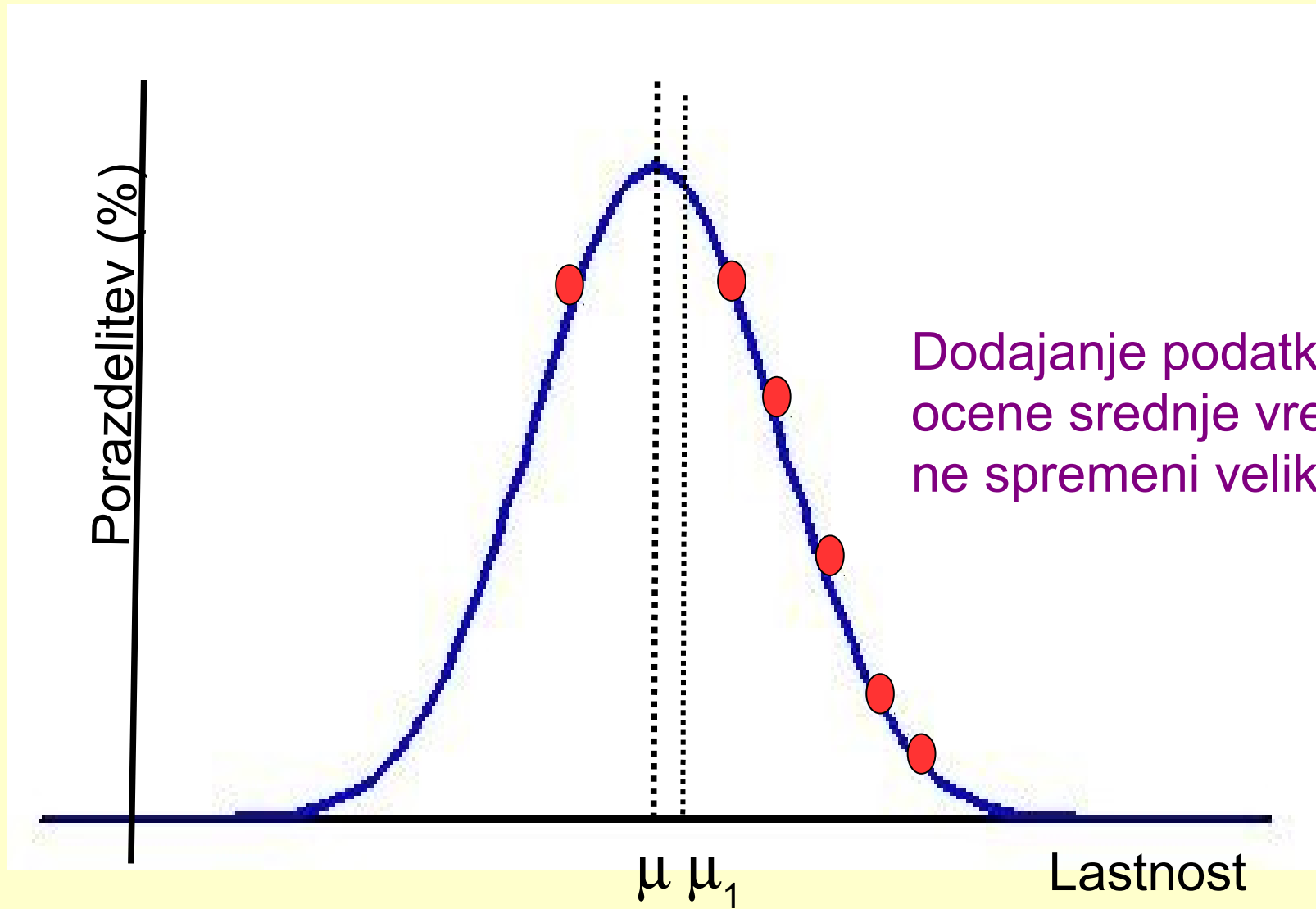
# Napoved



# Napoved plemenskih vrednosti



# Ocena





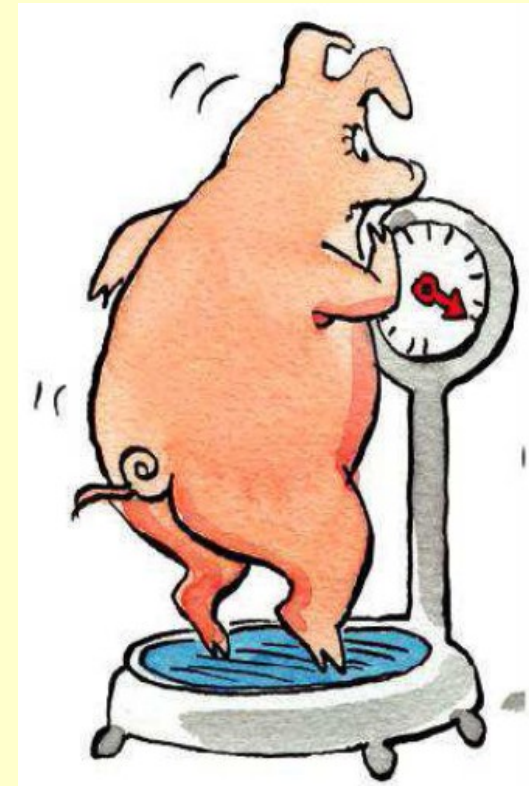
# Pasti pri preizkusih v pogojih reje

- Nepričakovane porazdelitve
  - Zaokrožene vrednosti
  - Odrezane porazdelitve (selekcija pred meritvijo)
  - Manjši standardni odkloni (selekcija na obeh straneh)
- Nestandardizirani pogoji med čredami
- Majhne črede in majhne primerjalne skupine
- Potrebno obdržati skoraj vse živali

Manj natančne meritve,  
večja varianca za ostanek,  
manjša točnost ...

# Zbiranje podatkov

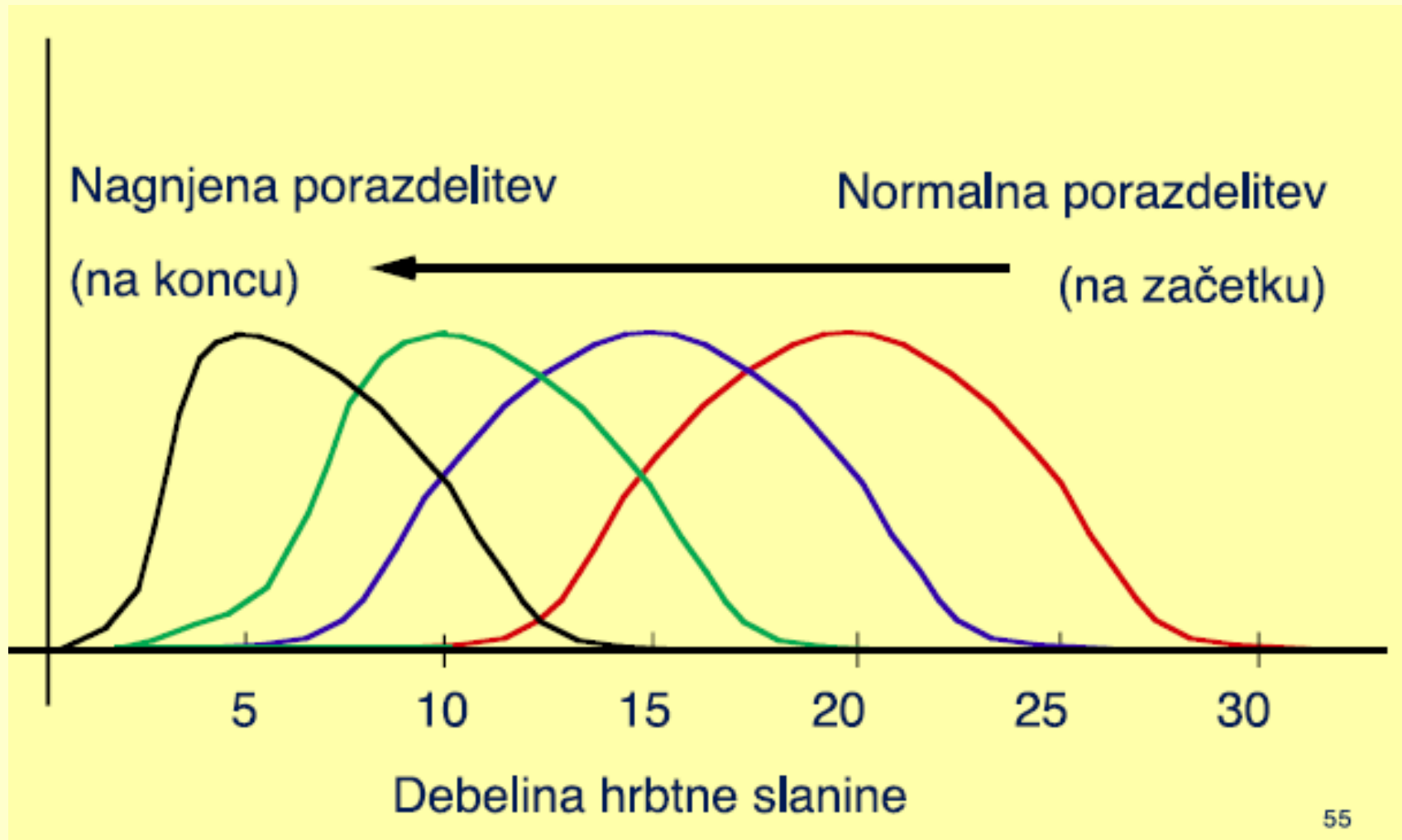
- Zajemanje podatkov na **mestu** nastanka v **realnem** času
  - Fizična in logična preveritev
  - Krajevni in časovni odmik zmanjša možnost
- Poenostavitev in poenotenje
  - Definicij (način merjenja ... ICAR)
  - Dokumentacije
- Posodobitev zbiranja podatkov
  - Uporaba novih tehnologij zajemanja podatkov
  - Podatkovne zbirke
  - Cenejši postopki
- Uporaba za več namenov: rejec, svetovalci, veterinar, selekcija



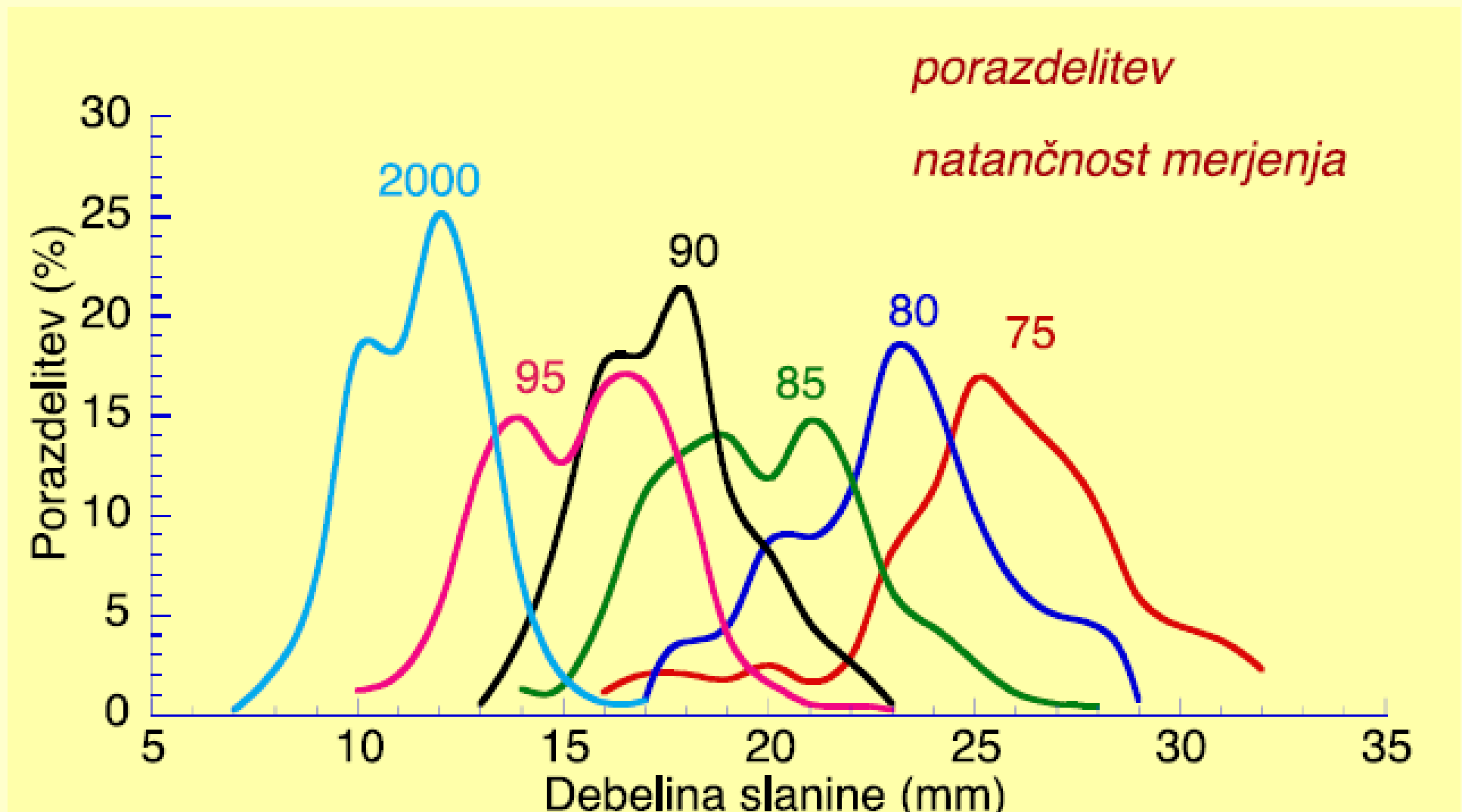
# Čas prejema podatkov

- Čimprej po nastanku
- Kdaj podatke potrebujemo (roki za pošiljanje v obdelavo)?
- Uporaba podatkov v druge namene
  - Spremljanje prireje

# Spreminjanje porazdelitve



# DHS pri merjascih



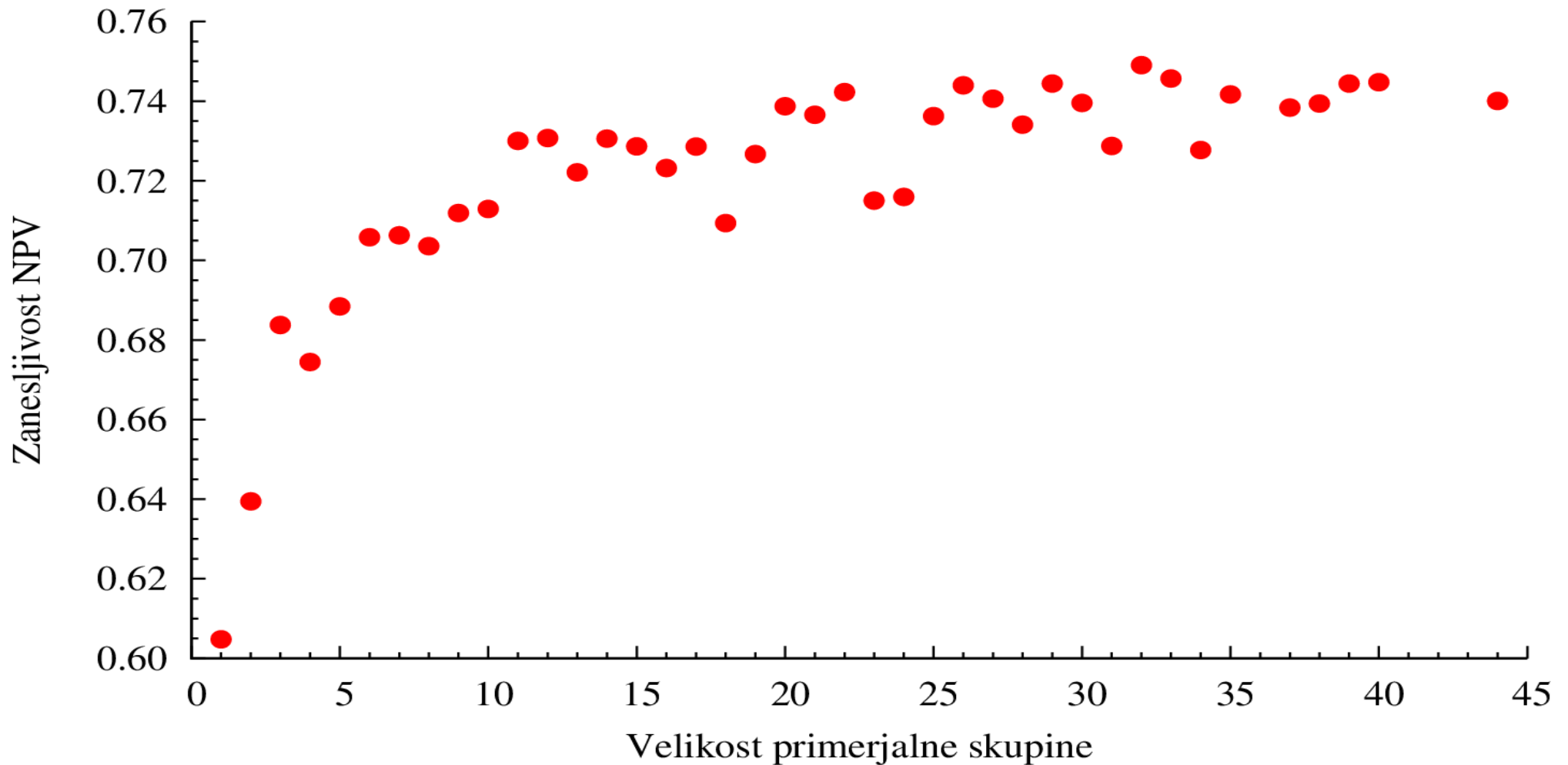
# Problematične lastnosti

- Lastnosti, ki dosežejo limito
- Merimo bolj natančno
- Izberemo druge ali dodatne meritve

# Statistični model

- Izbor sistematskih vplivov
  - Izpuščanje značilnih vplivov
  - Oblikovanje primerjalnih skupin (čreda-sezona...)
  - Pasma, križanje ali delež genov tuje populacije
    - Prisotnost neaditivnih genetskih komponent
  - Genetske skupine pri nakupih genetskega materiala iz tujih populacij
- Izbor naključnih vplivov
  - Skupno okolje v čredi, gnezdu ...
  - Permamentno okolje
  - Dominanca
  - Interakcija genotip x okolje

# Vpliv primerjalne skupine na točnost napovedi





# Statistični model

- Lahko zmanjša varianco za ostanek in se lahko poveča točnost ocene
- Ne more odpraviti hujše zavestne ali nezavestne napake
  - Ustvarjanje ali prirejanje podatkov
  - Nekorektno vodenje rodovništva
- Posledica:
  - Nekaj časa lahko nekomu ugajamo
  - Dolgoročno se zmanjšuje genetski trend

# Večlastnostni model

$$y_{tijkkl} = \mu_t + L_{ti} + b_{tI} (x_{tijk} - 10) + b_{tII} (x_{tijk} - 10)^2 + r_{tj} + g_{tijk} + a_{tijkkl} + e_{tijkkl}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_r \mathbf{u}_r + \mathbf{Z}_g \mathbf{u}_g + \mathbf{Z}_a \mathbf{u}_a + \mathbf{e}$$

$$\mathbf{G}_{r0} = \begin{bmatrix} 0.588 & 0.84 & -0.70 \\ & 6.00 & 4.75 \\ \text{sim.} & & 38.44 \end{bmatrix},$$

$$\mathbf{G}_r = \mathbf{I}_r \otimes \mathbf{G}_{r0}$$

$$\mathbf{G}_{g0} = \begin{bmatrix} 0.196 & 0.22 & -0.40 \\ & 1.00 & 0.06 \\ \text{sim.} & & 1.44 \end{bmatrix},$$

$$\mathbf{G}_g = \mathbf{I}_g \otimes \mathbf{G}_{g0}$$

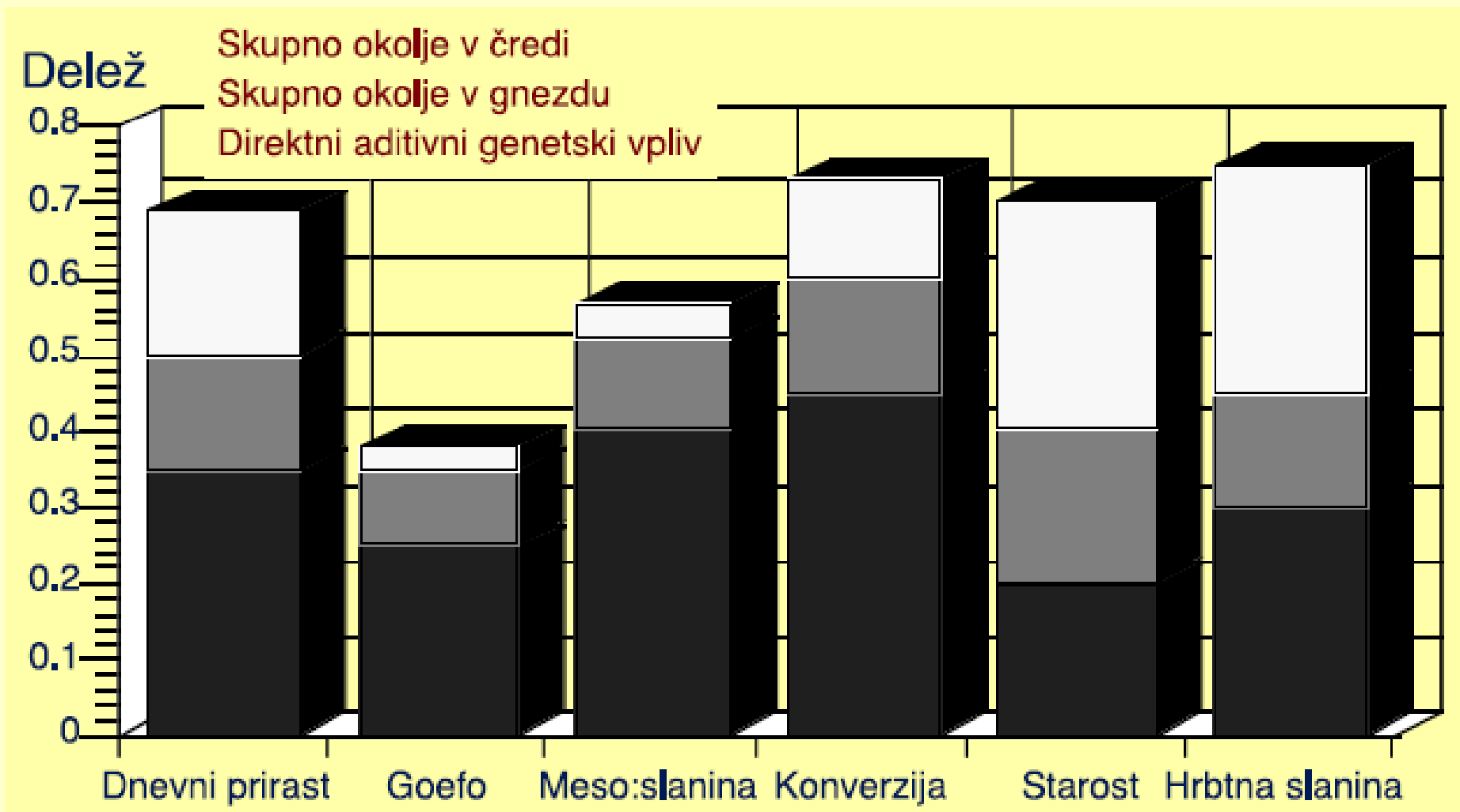
$$\mathbf{G}_{a0} = \begin{bmatrix} 0.490 & 1.12 & -4.21 \\ & 7.00 & -0.24 \\ \text{sim.} & & 94.09 \end{bmatrix},$$

$$\mathbf{G}_a = \mathbf{A} \otimes \mathbf{G}_{a0}$$

$$\mathbf{R}_0 = \begin{bmatrix} 0.686 & 2.45 & 17.12 \\ & 11.00 & 2.78 \\ \text{sim.} & & 156.25 \end{bmatrix},$$

$$\mathbf{R} = \mathbf{I} \otimes \mathbf{R}_0$$

# (Ne)poznavanje komponent (ko)variance



# Genomska selekcija

- Tudi možne napake pri vzorčenju in laboratoriju
- NPV odvisna od
  - Velikost »trening populacija«
  - Sorodstva z živalmi, ki jim napovedujemo plemensko vrednost
- Ne izboljša podatkov, če so fenotipski slabi