

Bayesove metode (bayesovsko sklepanje)

Pri sklepanju (inferenci) oziroma predlaganju konsistentnih (logičnih) modelov na osnovi eksperimentalnih podatkov se v biologiji, posebej še pri študijah na ravni celotnega genoma, praviloma srečujemo s situacijo, kjer so podatki nepopolni (npr. poznamo izražanje genov in koncentracije proteinov, ne poznamo pa loka(liza)cije proteinov), kjer v modele vgrajujemo novo pridobljene podatke, ne da bi pri tem celoten model gradili znova, in kjer kombiniramo eksperimentalna opažanja več modelnih sistemov, ne da bi poznali stopnjo paralelnosti/ortogonalnosti med njimi (t.j. ne vemo, kateri modeli so si v danem kontekstu podobni/enaki in kateri različni).

Za sklepanje v tem kontekstu so uporabne metode, ki jih široko imenujemo Bayesove metode (po Bayesovem pravilu $p(A|B)=p(B|A)\times p(A)/p(B)$). Večina 'klasičnih' statističnih metod omogoča opis problemov oblike $p(\text{podatki}|\text{hipoteza})$, t.j. kakšna je verjetnost opažanja podatkov pri dani hipotezi. Zaradi zgoraj omenjenih lastnosti podatkovnih nizov pa nas v biologiji pogosteje zanima verjetnost $p(\text{hipoteza}|\text{podatki})$ – torej, kako verjetna je postavljena hipoteza (model) pri danih podatkih.

Za podrobnejše seznanjanje s področjem Bayesovih metod v biologiji priporočam študij po predavanji Boba O'Hara z univerze v Helsinkih, ki so prosto dostopna na spletu (<http://www.rni.helsinki.fi/~boh/Teaching/Bayes/Bayes.html>).