|  |  |
| --- | --- |
| **Which statistical model should I estimate?** | |
| *Form of Data* | *Name of Model* |
| Continuous outcome | Standard linear regression model |
| Binary outcome (0,1) | Logit or probit model |
| Multiple categorical outcome | Multinomial logit model or multinomial probit |
| Ordered categorical outcome | Continuation ratio model, proportional odds model or stereotype logistic model |
| Count data | Poisson model |
| Count data (with over-dispersion) | Negative binomial model |
| Count data (with excessive zeros) | Zero-inflated poisson model |
| Continuous outcome (censored) | Tobit model |
| Multiway table | Loglinear model |
| Continuous outcome (experimental conditions) | Analysis of variance model |
| Correlation matrix | Factor analysis model |
| Hierarchical or multilevel data | Fixed or random effects model |
| Panel data continuous outcome | Fixed or random effects model |
| Panel data (with explicit time dimension) | Growth curve model |
| Panel data binary outcome (0,1) | Random effects logit model or random effect probit model |
| Panel count data | Random effects poisson model |
| Panel count data (with over-dispersion) | Random effects negative binomial model |
| Panel data (policy related) | Generalized estimating equation model |
| Duration data (no censored data) | Accelerated life model |
| Duration data (censored data) | Cox model |
| Duration data with more than one outcome | Competing risks model |
| Matched case control | Conditional (or fixed effects) logit |
| Continuous outcome with sample selection | Heckman model |
| Two binary outcomes (0,1) | Bivariate probit model |
| Two binary outcomes with sample selection | Heckprob model |
| Multiple discrete observed variables indicating continuous latent outcome | Latent trait model |
| Multiple discrete observed variables indicating categorical latent outcomes | Latent class model |
| Multiple continuous outcomes (experimental conditions) | Multivariate analysis of variance |
| Multiple continuous outcomes | Multivariate regression |
| Truncated continuous outcome | Truncated regression |
| Seemingly unrelated outcomes | Seemingly unrelated regression |
| Linear outcome with endogenous explanatory variables | Instrumental variable regression model |
| Binary outcome (skewed) | Scobit model |
| Binary outcomes (nested) | Nlogit model |
| Binary outcomes (with outliers) | Robit model |
| Grouped binary data | Glogit model |
| Blocked binary data | Blogit model |
|  | |
|  | |
| More exotic models… | |
|  | |
|  | |
| Panel data (continuous outcome) | Allison’s hybrid model |
| Count data (with over-dispersion and excessive zeros) | Zero-inflated negative binomial model |
| Count data (two processes one for 0 and another for the count) | Hurdle model |
| Repeated measures of one observed variable indicating latent development | Latent trajectory model |
| Dynamic panel data (continuous outcome) | Arellano-Bond model |
| Dynamic panel data (binary outcome) | Redprob (Marks Stewart’s Heckman estimator) |
| Paired comparisons (e.g. contests) | Bradley-Terry model |
| Hierarchical or multilevel data (with cross-classified structure) | Cross-classified random effects model |
| Panel data binary outcome (0,1) | Alternative-specific conditional logit (McFadden's choice) model |
| Tabular data | Goodman’s RC model,  Unidiff Model |
| Repeated data on transitions | Markov chain models |
| Repeated data on transitions with multiple indicators | Latent transitions model |