Ch 6. Characterizing uncertainty

Amy Hurford
Memorial University

Traditional statistical assumptions

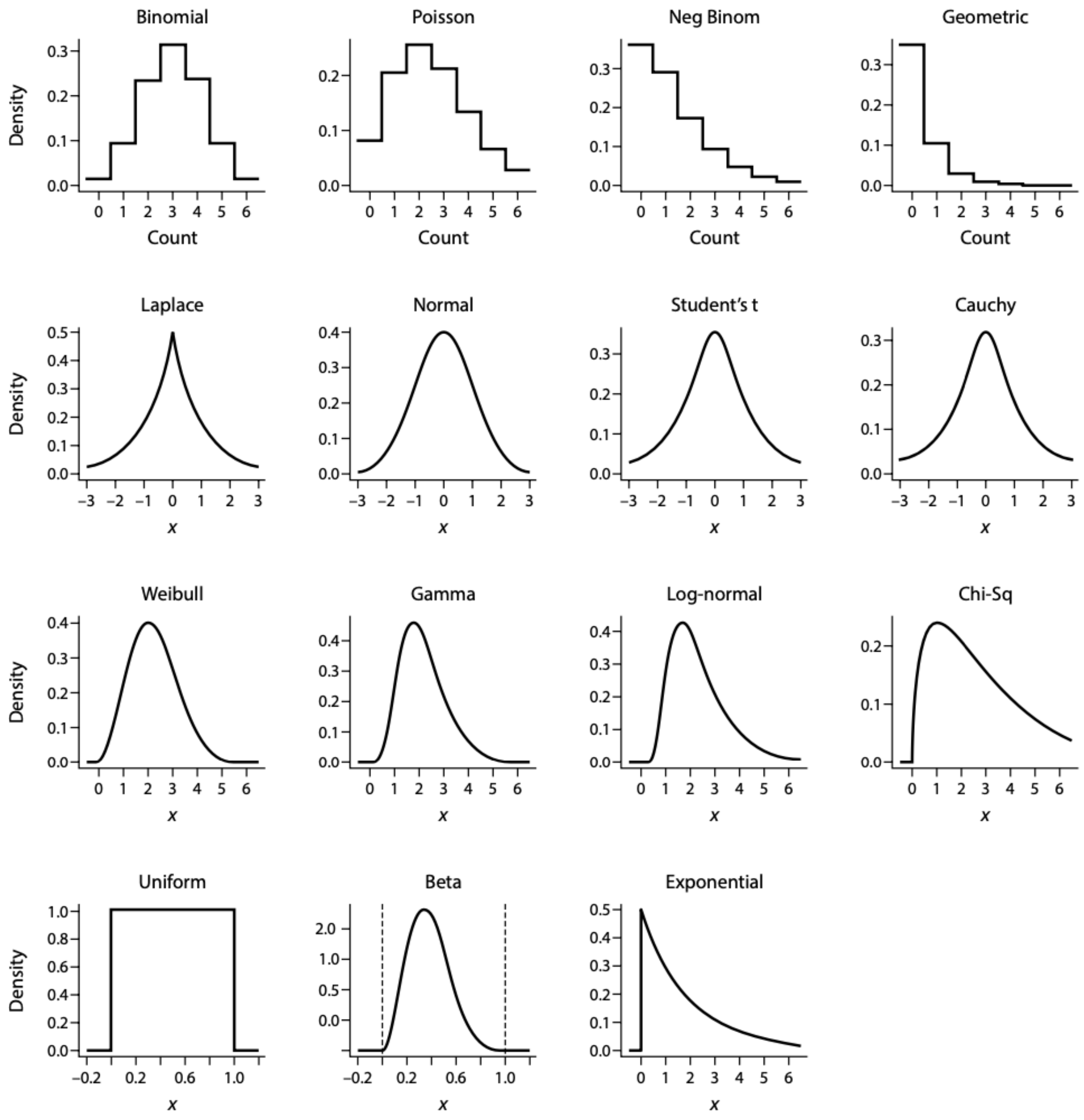
Figure

Probability distributions

Discrete			
Bernoulli(p)	$p^x(1-p)^{1-x}$	[0,1]	Success ($x = 1$) with probability p. Binomial with $n = 1$.
Binomial(n,p)	$\binom{n}{x}p^x(1-p)^{n-x}$	$[0 \ldots N]$	Number of successes, with probability p, given n trials.
Poisson(λ)	$\frac{\lambda^x}{x!}e^{-\lambda}$	$0,\infty$	Number of events, occurring at rate λ, to occur over a fixed interval.
Negative Binomial(n,p)	$\binom{x+n-1}{x}p^n(1-p)^x$	$0,\infty$	Number of trials, with probability p, before n successes occur. Also a Poisson-Gamma mixture.
Geometric(p)	$p(1-p)^x$	$0,\infty$	Number of trials needed before a success occurs. Special case of negative binomial with $n = 1$.

Dietz 2017. Ecological forecasting.

Probability distributions



Dietz 2017. Ecological forecasting.

Generalized linear models

Logistic regression:

$y_i = Bern(p_i)$ Data model

Boolean: TRUE/FALSE or 0/1

$p_i = logit(\theta_i) = \frac{1}{1 + e^{-\theta_i}}$ Link function

Converts real number to domain of the distributions mean

$\theta_i = \beta_0 + \beta_1 x_i$ Linear model

Dietz 2017. Ecological forecasting.

Observation error

A discussion on modeling observation error is straightforward if we believe that all of the observed residual error is measurement error and that there is no uncertainty in the model structure, covariates, or drivers.

$y_{i,obs} \sim g(y_i)$	Data model
$y_i = \beta_0 + \beta_1 x + \varepsilon_{i,add}$	Process model
$\varepsilon_{i,add} \sim N(0,\tau_{add})$	Process error

Random effects

- A fixed factor has categories, or levels, that we set to certain values in an experiment or levels that we choose in an observational study. We infer only to those levels.
- A random factor has categories that we have not chosen, or that vary even after we make them as uniform as possible.
- The same factor might be treated as fixed or random depending on the population of inference.

Fixed or random effect?

- treated versus untreated (control) units
- before versus after treatment of an experimental unit.
- day versus night
- habitat types
- insect stages (larval, adult)
- tanks in aquaculture
- plots in agriculture
- individual organisms

Hierarchical models

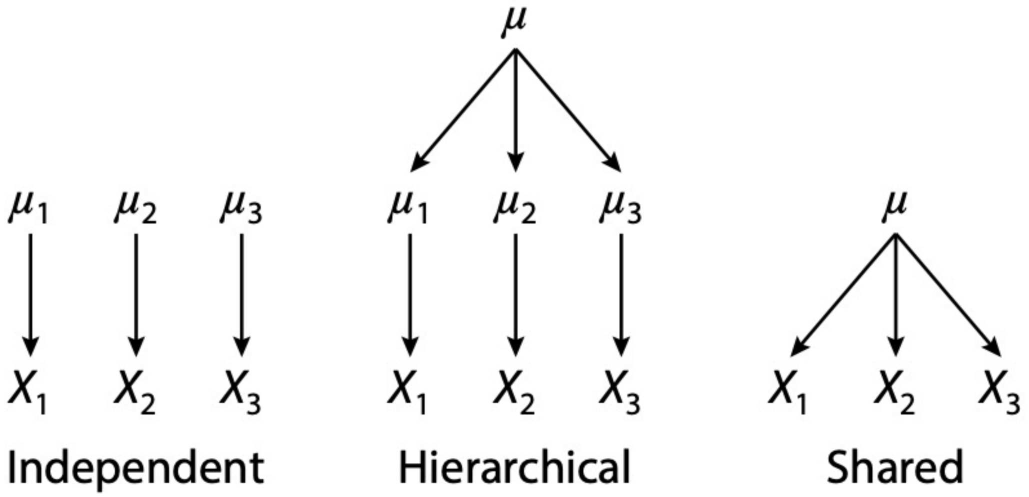


FIGURE 6.6. Hierarchical models represent the continuum of possibilities between treating data sets as independent versus treating them as identical. As a result, they partition process variability between the different levels of the hierarchy.

Powerful tool for estimating and partitioning process error