## Review Likelihood

A way of measuring how "likely" a set of parameters are, given some data. Useful for comparing several possible options for parameters. However, adding parameters always will increase the likelihood. Can use likelihood ratio test to test nested models; can also use an information criteria, like AIC or BIC.

## AIC and BIC

AIC and BIC: ways of penalizing the likelihood when additional parameters are added.

AIC: $-2 \times$ log-likelihood $+ \; 2 \times$ number of parameters

BIC: $-2 \times$ log-likelihood $+ \log$(Number of observations) $\times$ number of parameters

AIC usually keeps more parameters than BIC.

## Procedure for Making Inferences

Two things we want to make inferences about: 1) fixed effects and 2) random effects/covariance structure. Hard to test both at the same time; better to concentrate on one aspect of the model at a time. Usually:

1. look at pictures, guess at reasonable fixed effects; ok to include too much
2. compare various covariance structures using those fixed effects
3. using that covariance structure, compare fixed effects

## REML vs ML

REML gives "better estimates" but likelihood values are only comparable between models with the same covariance structure

- To test fixed effects: use ML
- To test random effects: use REML
- To get parameter estimates from a chosen model: use REML

### Methods of Inference

How to do? Why use? Why not?

- ANOVA-like: for single fixed effects only within the context of a single model
- LRT: nested models only, good for inference on specific parameters; can test multiple fixed effects simultaneously
- IC: best for choosing a model, not for testing specific parameters, especially useful for covariates and prediction

Also…

- Stepwise and all subsets procedures
- Bayesian procedures
- simulation-type methods

And don't forget

- Knowledge about the problem
- Desire for an interpretable model

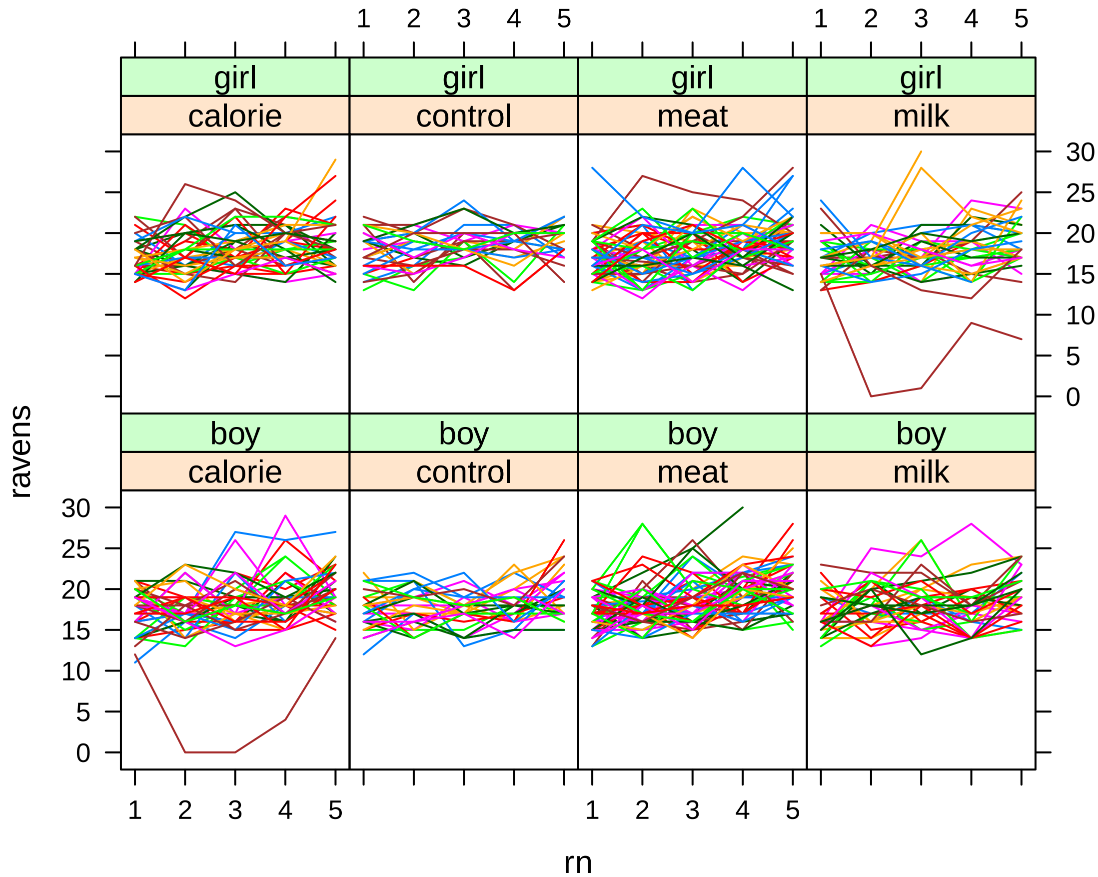
## An Example

Read in data, select only columns we'll use, and get "complete cases"

```
> d <- read.delim("http://rem.ph.ucla.edu/rob/mld/data/tabdelimiteddata/cognitive.txt")
> ds <- subset(d, select = c("id", "ravens", "rn", "treatment",
+     "sex", "age_at_time0", "height", "weight", "yrsofsch"))
> ds <- subset(ds, complete.cases(ds))
```

Make a picture

```
> p <- xyplot(ravens ~ rn | treatment * sex, group = id, type = "l",
+     data = ds)
> plot(p)
```




Compare some covariance structures: Note that R gives you output you can't use!

```
> m0 <- gls(ravens ~ rn * treatment + sex, data = ds)
> m1 <- lme(ravens ~ rn * treatment + sex, random = ~1 | id, data = ds)
> m2 <- lme(ravens ~ rn * treatment + sex, random = ~rn | id, data = ds)
> m3 <- gls(ravens ~ rn * treatment + sex, correlation = corAR1(form = ~1 |
+     id), data = ds)
> anova(m0, m1, m2, m3)
```

```
   Model df      AIC      BIC    logLik   Test   L.Ratio p-value
m0     1 10 7953.278 8007.031 -3966.639
m1     2 11 7793.578 7852.706 -3885.789 1 vs 2 161.69999  <.0001
m2     3 13 7790.028 7859.907 -3882.014 2 vs 3   7.54991  0.0229
m3     4 11 7803.806 7862.934 -3890.903 3 vs 4  17.77737  0.0001
```

We'll choose the correlation structure with random intercept and slope.

Now for fixed effects:
Remember to use ML instead of REML:

```
> f2 <- lme(ravens ~ rn * treatment + sex, random = ~rn | id, data = ds,
+     method = "ML")
> logLik(m2)

'log Lik.' -3882.014 (df=13)

> logLik(f2)

'log Lik.' -3872.784 (df=13)

> anova(f2)

             numDF denDF   F-value p-value
(Intercept)      1  1273 31499.875  <.0001
rn               1  1273   145.435  <.0001
treatment        3   323     1.605  0.1880
sex              1   323     1.523  0.2181
rn:treatment     3  1273     1.780  0.1492
```

Here;s a bunch of models

```
> f0 <- lme(ravens ~ 1, random = ~rn | id, data = ds, method = "ML")
> f1 <- lme(ravens ~ rn, random = ~rn | id, data = ds, method = "ML")
> f2 <- lme(ravens ~ sex, random = ~rn | id, data = ds, method = "ML")
> f3 <- lme(ravens ~ rn + sex, random = ~rn | id, data = ds, method = "ML")
> f4 <- lme(ravens ~ rn * treatment + sex, random = ~rn | id, data = ds,
+     method = "ML")
```

And one way to get AIC/BIC (but disregard the LRT tests)

```
> anova(f0, f1, f2, f3, f4)

   Model df      AIC      BIC    logLik   Test   L.Ratio p-value
f0     1  5 7884.964 7911.868 -3937.482
f1     2  6 7769.133 7801.418 -3878.567 1 vs 2 117.83088  <.0001
f2     3  6 7885.180 7917.465 -3936.590
f3     4  7 7769.580 7807.246 -3877.790 3 vs 4 117.60039  <.0001
f4     5 13 7771.569 7841.520 -3872.784 4 vs 5  10.01057  0.1242
```

Better LRT tests are

```
> anova(f3, f4)

   Model df      AIC      BIC    logLik   Test  L.Ratio p-value
f3     1  7 7769.580 7807.246 -3877.790
f4     2 13 7771.569 7841.520 -3872.784 1 vs 2 10.01057  0.1242

> anova(f1, f0)

   Model df      AIC      BIC    logLik   Test  L.Ratio p-value
f1     1  6 7769.133 7801.418 -3878.567
f0     2  5 7884.964 7911.868 -3937.482 1 vs 2 117.8309  <.0001

> anova(f2, f0)

   Model df      AIC      BIC    logLik   Test L.Ratio p-value
f2     1  6 7885.180 7917.465 -3936.590
f0     2  5 7884.964 7911.868 -3937.482 1 vs 2 1.78411  0.1816

> anova(f4, f0)

   Model df      AIC      BIC    logLik   Test  L.Ratio p-value
f4     1 13 7771.569 7841.520 -3872.784
f0     2  5 7884.964 7911.868 -3937.482 1 vs 2 129.3951  <.0001
```