Trying to clean up the mess: Bayes, Frequentism, NHST, Parameter estimation etc. etc.

Bob O'Hara
Senckenberg BiK-F
Frankfurt am Main
Germany
I'm an idiot. You're an idiot. Everyone is an idiot. Never forget.
P-values are evil, so be Bayesian instead

Many statisticians also advocate replacing the P value with methods that take advantage of Bayes' rule: an eighteenth-century theorem that describes how to think about probability as the plausibility of an outcome, rather than as the potential frequency of that outcome.


One possible replacement [of p-values] that might fit the bill is a rival approach of data analysis called Bayesianism.

But how, exactly does Bayes help?
But how, exactly does Bayes help?

It doesn't
### Everyone Can do Everything

<table>
<thead>
<tr>
<th></th>
<th>Frequentist</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis Testing</td>
<td>P-values (BIC?)</td>
<td>Bayesian p-values Bayes Factors</td>
</tr>
<tr>
<td>Model selection</td>
<td>AIC, stepwise selection</td>
<td>DIC, WAIC</td>
</tr>
<tr>
<td></td>
<td>LASSO etc</td>
<td>Shrinkage Indicator models</td>
</tr>
<tr>
<td>Parameter Estimation</td>
<td>MLE Confidence Intervals</td>
<td>Posterior mode etc Credible Intervals</td>
</tr>
</tbody>
</table>
The Problem with Hypothesis Tests

The null hypothesis is almost always false
A 2-sample t-test

\[ t = \sqrt{2n} \left( \bar{x}_1 - \bar{x}_2 \right) \]

Depends on \( n \): if you want a significant result, simply increase your sample size
Bayes Factors are no better

\[ BF = \left( \frac{1 + \frac{t^2}{\nu}}{1 + \frac{t^2}{\nu \left( 1 + n_\delta \sigma_\delta \right)}} \right)^{-\left(\nu + 1\right)/2} \left(1 + n_\delta \sigma_\delta \right)^{1/2} \]
The Solution

Estimate meaningful parameters

Regression coefficients

Variance components/proportions
Frequentists & Bayesians can both do this right

<table>
<thead>
<tr>
<th>Frequentist</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Likelihood Estimator</td>
<td>Posterior Mode (mean, median)</td>
</tr>
<tr>
<td>Confidence Interval</td>
<td>Credible Interval</td>
</tr>
</tbody>
</table>
What to do when?

Everything has its place

"Pleuronectes platessa" by © Hans Hillewaert. Licensed under CC BY-SA 4.0 via Commons - https://commons.wikimedia.org/wiki/File:Pleuronectes_platessa.jpg#/media/File:Pleuronectes_platessa.jpg
Hypothesis Tests

- Have to assume null hypothesis can be true
  - otherwise, test of whether have enough data
- Different questions:
  - Is $H_0$ true
  - Does it matter if $H_0$ is false?
Model Selection (AIC etc)

- No strong *a priori* hypothesis
- Question: which model is “best”
- More about removal of unimportant effects
  - improves rest of model (better estimates, simpler interpretation)
Parameter Estimation

- What we should be trying to do
- Answers the biological question
  - how much more intelligent are statisticians than psychologists?
- Interpretation should not be statistical
  - psychologists know how big an effect is
Finally

Frequentists vs Bayesians
If you're a purist...
If you're a pragmatist...
Use the approach that works best

Just make sure you understand it!
Both approaches have strengths and weaknesses

Bottom line: do what works for you
What works for me
For simpler stuff, frequentist methods are easier to fit and work with.

Once stuff gets difficult, the flexibility of Bayesian methods makes them preferable.
Face width & Fitness
Testosterone

Aggression

Survival in war

Number of offspring

Face width

Measured Face width

The mechanistic model
The Model to be fitted

Other Covariates

Survival in war

Number of offspring

Face width

Measured Face width
This would be straightforward

- Survival in war
  - Face width
  - Other Covariates

- Number of offspring
  - Face width
  - Other Covariates
But…..

No observation model

No use of information across models

Also have missing data
Wider faces → more children
The mess cleaned up?

“P-values are evil”

Subjective or objective probability
There's a real world out there