This foundation;
time: multiparameter
next time: problems

real: DSch.8

today I'll finish covering all the material in hwk 1

eCommons site now ready for you to upload your PDF of hwk 1, drop-dead due date for hwk 2 at eCommons set to 23 Mar, but please get your assignment posted to eCommons by next Tue, when hwk 2 will be assigned. Next Mon is a holiday, so no discussion sections next week.

Ingredients in the full Bayesian paradigm

problem \(\mathbf{P} = (\mathbf{A}, \mathbf{C})\) context

\(\mathbf{P} = (\mathbf{0}, \mathbf{D}, \mathbf{B})\)

unknown of \(A\) data set

principal interest II propositions (all T)
In the Bayesian story, to draw good inferences, make good predictions, and help people make good decisions under uncertainty, you need to build a probability model that relates unknown quantities \( \Theta \) to known quantities \( (D, B) \).

**Inference & Prediction:**

- \( p(\Theta | B) \) \( \vdash \) prior dist

- \( p(D | \Theta, B) \) \( \vdash \) sampling dist

\[
L(\Theta | D, B) = c \cdot p(D | \Theta, B)
\]

model for inf. &/or pred.

\[
M = \{ p(\Theta | B), L(\Theta | D, B) \}
\]

positive constant
Decision \( (a \mid \theta) \): action space

\[ \{ a_1, a_2, \ldots \} \]

Possible actions

\[ U(a, \theta \mid B) = \text{trade off costs & benefits of choosing action} \]

\[ \text{takes values on real line} \]

\[ \text{(who bij } u \text{ = better really were } \theta \text{)} \]

Ex: how fast should I drive on

IS? to LA

optimal

Pick action to maximize

\[ E(B) (u(a, \theta \mid B)) \]

The hard part in applied statistical work:
The mapping from $P$ to $M = \{ p(0,1|B), \ell(0,1|B) \}$
or $M_d = \{ p(0|B), \ell(0|B), \varphi(1|B), \psi(0,0|B) \}$
is rarely unique - typically there's a range of plausible sampling dist & a range of plausible priors - this is model uncertainty.

---

Venn wanted: if scientists $A$, $B$ have same data (set $D$), then $A$, $B$ draw same conclusions: to Venn, this is objectivity.

---

objectivity

in uncertainty

quantification

in uncertainty
However, A & B may have different amounts of info. external to S, & even if they have the same external info. they may dispute reasonably disagree on assumptions or judgments about how to use it to quantify uncertainty about θ. The kind of objectivity Van wanted is both impossible & undesirable. 

By far with high-dimensional, $p(\theta | y, B) = c \cdot p(\theta | B) \cdot q(\theta | DB)$
$\theta = (\theta_1, \ldots, \theta_p)$ for positive integer $p$ (could be big)

to normalize LH, need to compute

$$\int \int \int \cdots \int p(\theta | B) \cdot R(\theta | DB) \, d\theta_1 \cdots d\theta_p$$

$$(p=1,000,000): \text{fun? nooooo}$$

Technical challenge

Approximating high-dimensional integrals

This was only solved in the 1980s since. [End of page, for Week 2]
Inference [Ex. New data study: N810]

With we'll demonstrate it that multi-parameter data vector problems \( \mathbf{y} = (409, 400, \ldots, 400) \)
does not look like it can be modeled well by a normal (Gaussian)
sample distribution, but in phase I of the analysis we're going to use the Gaussian sample model anyway, to get practice with it.

Phase I

Sample model

\( (Y_1, \ldots, Y_n) \) IID

\( Y \sim N(\mu, \sigma^2) \)

vector of unknowns

mean variance

here is \( \Theta = (\mu, \sigma^2) \): dimension \( p = 2 \)
Fisher's algorithm

Step 0: marginal sampling distribution of $Z_i$

$Z_i$ is continuous, so this is a density (pdf):

$$p(l_i | \mu, \sigma^2, \beta) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2\sigma^2} (l_i - \mu)^2 \right]$$

Step 0: Joint sampling distribution is product of marginals:

$$p(l | \mu, \sigma^2, \beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2\sigma^2} (l_i - \mu)^2 \right]$$

$$= \sigma^{-n} (2\pi)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (l_i - \mu)^2 \right]$$

Step 0: Likelihood function:

$$L(\theta | \tau, \beta) = L(\mu, \sigma^2 | \tau, \beta)$$

$$= c \sigma^{-n} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (l_i - \mu)^2 \right]$$
Step 2.5  \[ \log \text{likelihood function} \]

\[ \ell(\theta | y, B) = \ell(\mu_0, \sigma | y, B) \]

\[ = -n \log \sigma - \frac{1}{2 \sigma^2} \sum_{i=1}^{n} (y_i - \mu_0)^2 \]

\[ \sim \]

\[ \theta \sim \text{Beta}(\frac{1}{2}, \frac{1}{2}) \]

\[ (y_i | \theta, B) \sim \text{Bernoulli}(\theta) \]

\[ (i = 1, \ldots, n) \]

\[ \theta \sim \text{Beta}(\alpha, \beta) \]

\[ \rightarrow E(\theta) = \frac{\alpha}{\alpha + \beta} \]

\[ \text{Beta}(\alpha + s, \beta + n - s) \]

\[ E(\theta | y, B) = \frac{s + \frac{1}{2}}{n + 1} \]

\[ \text{Beta}(\frac{1}{2}, \frac{1}{2}) : \text{Jeffreys prior} \]
sorted \( y_i \)

what the sorted \( y_i \) should look like if normality is ok assumption

\[
\text{standardized } (\text{mean } 0, \text{SD } 1)
\]

expected normal values