

Pedagogical goals: help; inspiration w/ your research projects.

~~Recall.~~
~~Question genre what makes 2 types of lang. different~~

lecture content ; pointers to resources;

~~individualized feedback~~

↑
generic to ~~class~~
everybody
in class

R pgs to add!

individualized feedback

I haven't been able to get to this on Piazza, so instead I'm going to allocate face-to-face time for next week.

Course webps: - ~~quick~~
▷ quick link to today
▷ ~~quick~~ link to ycsch

- ~~write~~ mandatory
- presumably will discuss ~~how you~~ when you are w.r.t. what you promised to get done in a week, but ~~that's not~~
- primary goal: ~~fast~~ general ; specific feedback from me.
- you pick time duration.

<as for generic stuff:>

↓
top of webps.
some todbits, alphabetized.

"what differentiates 2 types of lang?"

- last time: ~~no p start~~ Bayesian ID of distinguishing features >
 - inspire you to try it or improve on it.
- this time: case study of hypothesis-driven approach
 - interesting ; inspiring features;
 - adapt to yr own devices...

~~bars: Akentara: B~~

finishing up from last time:

- see if students can reconstruct!

this went pretty well
test stat needs θ 's, which we can estimate from
counts as MLE for the posterior.

... should get to: posterior $P(\bar{\theta} | \bar{c})$ ~~prop to \bar{c}~~ Dir w/

$$\alpha_i = d_i + c_i$$

\int integration as pseudocount

- we recalled many things, and discussed choice points.

- note: using the mode was easier, b/c you can talk about
"the multinomial" that this Dirichlet likes best.

ideas / variants:

- have enough data that you can just estimate the "null hypothesis" version of
your test statistic ~~by~~ from it? (bootstrap sampling?)

- if you just want the prior for smoothing, you could "just" smooth (e.g. by add- δ)
instead of going thru all this machinery.

- should you have a separate prior for "Dens" vs. "Reg" or just use one?
Is the "use concept of the same data" to get prior OK in terms of
really reinforcing a (non-previous \smile) prior that D's & R's use
the same language?

- instead of taking the $\hat{\theta}$ that is the MLE for the posterior, could you
integrate out?

- better estimates of the variance of the Log-odds - ratios

- how they picked $\alpha_0 = \sum \alpha_i$

case + study:

< Louis; Nenkava '13 > - hypothesis driven approach.

Pedagogical struggle:

(a) go thru paper line-by-line, ~~and~~ stick to ~~exp~~ studying the many interesting choices? ideas.

(b) use paper as springboard ~~to~~ ^{deeper} to, insert ^{deeper} asides on related interesting ideas [that may be ~~an little too complicated~~ of less ~~practical~~ direct applicability].

• Decision: you can; should read the paper on your own ~~so~~
~~I'm going to skim~~
* * it really bears line-by-line analysis * *

~~but~~ - and this allows me to take the more meandering approach.

[QUICK]: question set-up, data setup & idea.

[QUICK]: outline of features ::

- visual-~~nature~~ ^{ness} (3.1)
- people-centric (3.2)
- beautiful → unusual lang. (3.3)
- narrative? sourced? an interview? (3.4)
- ~~appeals~~ appeals to emotions? (3.5)
- amt of research content (3.6)
- readability (5, P2)
- good discourse coherence; transitions (5, P2)
- ~~interesting by others~~ prev. interestingness features
- bag of words (BOW)

mentioned some of this.

each is treated w/ a lot of care, and paper is pretty detailed.
As inspiration, can be expanded to apply to you. (visual → ^{political?} food-oriented)

• show MRC, to show kinds of data showed this pg

• go thru ~~beautiful~~ ^{unusual} language.

- lang. models.

<topic that deserves several lectures>

- use θ_i notation.

- ~~p~~ ref. to smoothing

- perplexity: requires explaining cross-entropy

→ explain entropy. (use 'surprised' idea - see Brandon's blog post) (-log₂)

- use Lagrange mult to show:

θ_i fixed
considered
"true"
or
"reference"

~~$-\sum p(x) \log q(x)$ min. when $q(x) = p(x)$.~~

~~$-\sum p(x) \log \theta$~~

$-\sum_i \theta_i^{\text{fixed}} \log \theta_i$ min for $\theta_i = \theta_i^{\text{fixed}}$.

<as opposed to Jensen's inequality for KL divergence>

perplexity: interpret as branching factor or # of choices.