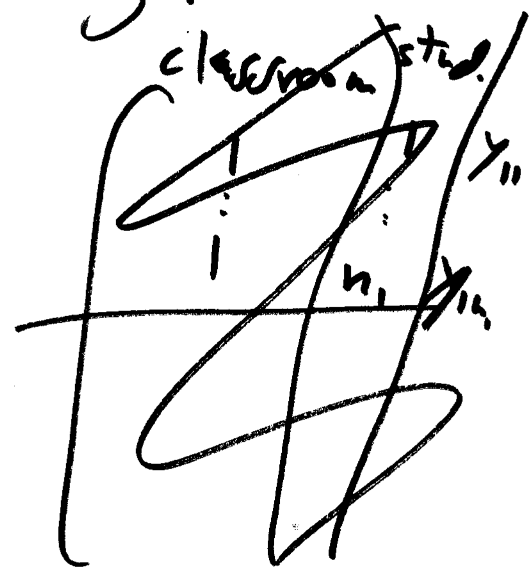
hierarchical data structure:

STAT 206
18 Feb 21

DD → H

(1) (education) y_{t} = math score in 6th grade (for student t)

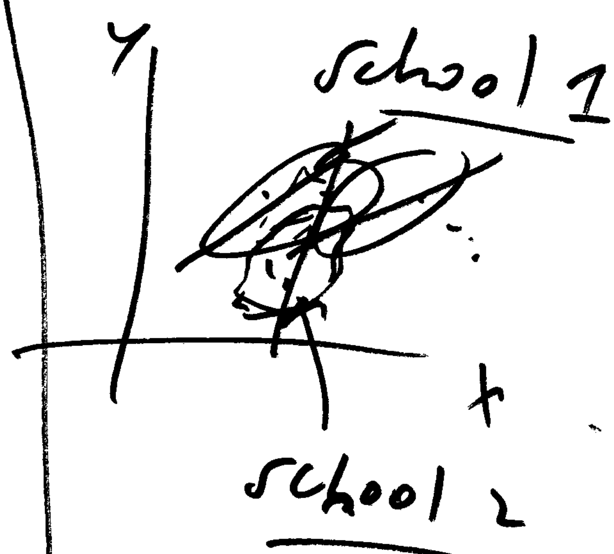
x = math — 5th grade



$$y_{ij} = (\beta_0)_i + (\beta_1)_i x_{ij} + e_{ij} \quad (*)$$

i (classroom), j ← student (nested in classroom)

(i) cl.	(j) st.	y	x
1	1	y_{11}	x_{11}
⋮	⋮	⋮	
1	n_1	y_{1n_1}	x_{1,n_1}
2	1	.	.
⋮	⋮		
2	n_2		
⋮	⋮	⋮	⋮
N	1		
⋮	⋮		
N	n_N		



$$\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}_i = \boxed{\begin{pmatrix}\beta_0^*\\ \beta_1^*\end{pmatrix}} + \begin{pmatrix}\epsilon_{1i}\\ \epsilon_{2i}\end{pmatrix}$$

y

Biden
other
Biden
Trump
⋮

$n = 1{,}265$

1 row for each sampled person

(raw data)

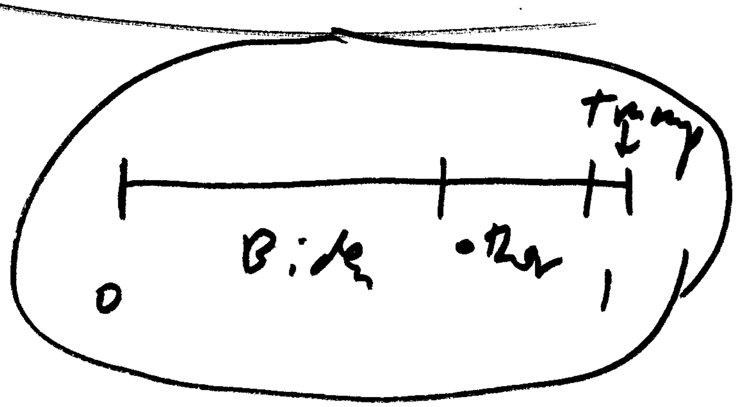
y ← qualitative (categorical)
nominal

sort

Biden
⋮
Biden — $n_1 = 659$

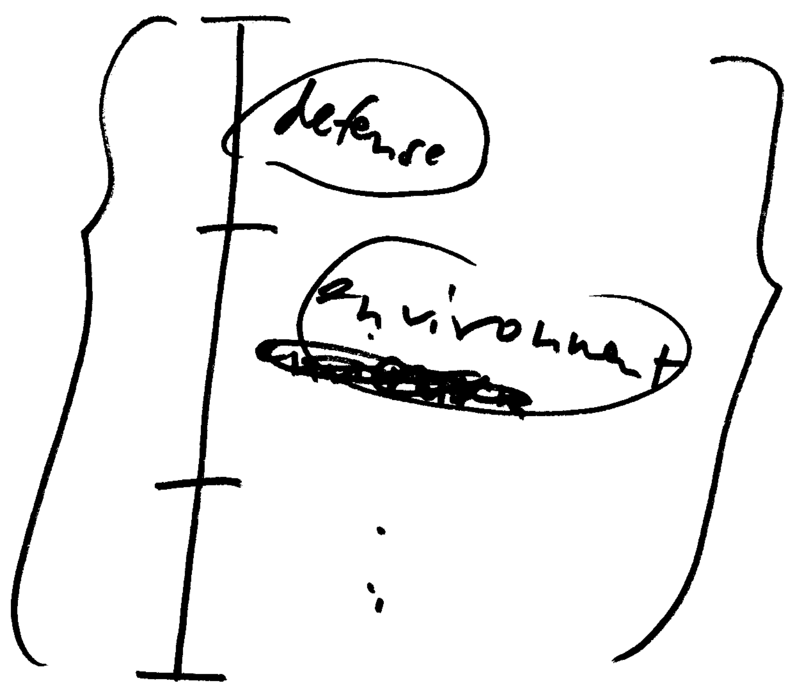
other
⋮
other — $n_3 = 52$

Trump
⋮
Trump — $n_2 = 554$



$ k T

2022 budget



THT2 | joint sampling dist. of $\underset{\sim}{N} = (N_1, \dots, N_k)$

$$p(\underset{\sim}{N} = \underset{\sim}{n} \mid \underset{\sim}{\theta}\, \mathcal{B}) = \boxed{C} \prod_{j=1}^{k} \theta_j^{n_j}, \qquad \left(\sum_{j=1}^{k} \theta_j = 1 \right)$$

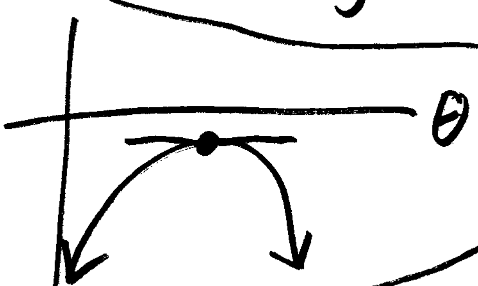
$C = \frac{n!}{\cdots}$ (constrained)

likelihood | $\ell_c(\underset{\sim}{\theta} \mid \underset{\sim}{N}\, \mathcal{B}) = \underline{c} \cdot \prod_{j=1}^{k} \theta_j^{n_j}$

$(c > 0)$ $\quad (c = 1)$

$$\ell\ell_c(\underset{\sim}{\theta} \mid \underset{\sim}{N}\, \mathcal{B}) = \sum_{j=1}^{k} n_j \log \theta_j$$
$$= n_1 \log \theta_1 + n_2 \log \theta_2 + \dots + n_k \log \theta_k$$

MLE | $\frac{\partial}{\partial \theta_j} \ell\ell_c(-) = \frac{n_j}{\theta_j}$ for $j = 1, \dots, k$



$$\left\{ \frac{n_1}{\theta_1} = 0, \ \frac{n_2}{\theta_2} = 0, \ \dots, \ \frac{n_k}{\theta_k} = 0 \right\}$$

we need to solve a constrained optimization problem: Lagrange multiplier (!)

(or) $(k=3)$ set $\theta_3 = 1 - \theta_1 - \theta_2$

$$\ell_u(\theta_1, \theta_2 \mid \underset{\sim}{N} \cap B) = \theta_1^{n_1} \cdot \theta_2^{n_2} \cdot (1-\theta_1-\theta_2)^{n_3}$$

(unconstrained)

$$\ell\ell_u(\theta_1, \theta_2 \mid \underset{\sim}{N} \cap B) =$$

$$n_1 \log \theta_1 + n_2 \log \theta_2 + n_3 \log(1-\theta_1-\theta_2)$$

$$\frac{\partial}{\partial \theta_1} \ell\ell($$

$$N \alpha \beta = (\cancel{175,000,000}) \cdot \left(\frac{500,000}{\cancel{175,000,000}}\right) \cdot (0.99)$$

$$= 455,000$$