this time: latent variable bootstrap models

next time: linear & logistic regression

read: Degroot & Schervish (DS) sec. 12.6

AMS 206

13 Mar 18

①

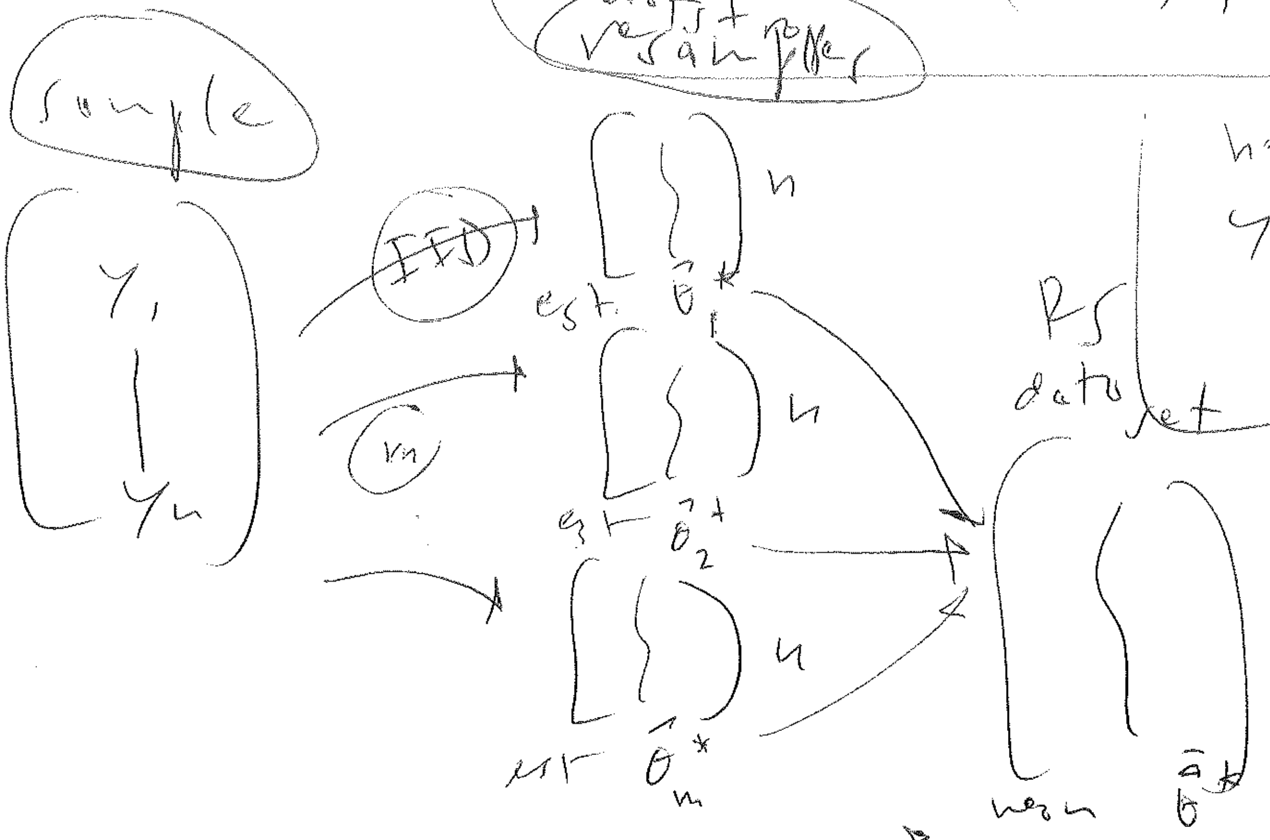
$$(\theta) \sim p(\theta)$$

lik → $$(Y_i \mid \theta B) \overset{IID}{\sim} p(Y_i \mid \theta B)$$

$\in \mathbb{R}$ $\quad (i = 1, \ldots, n)$

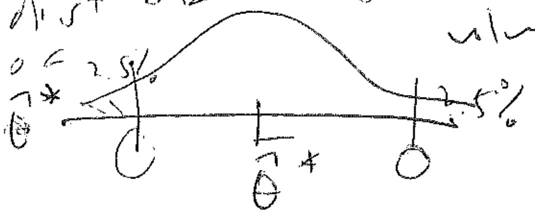
bootstrap resampler

how do you know p?

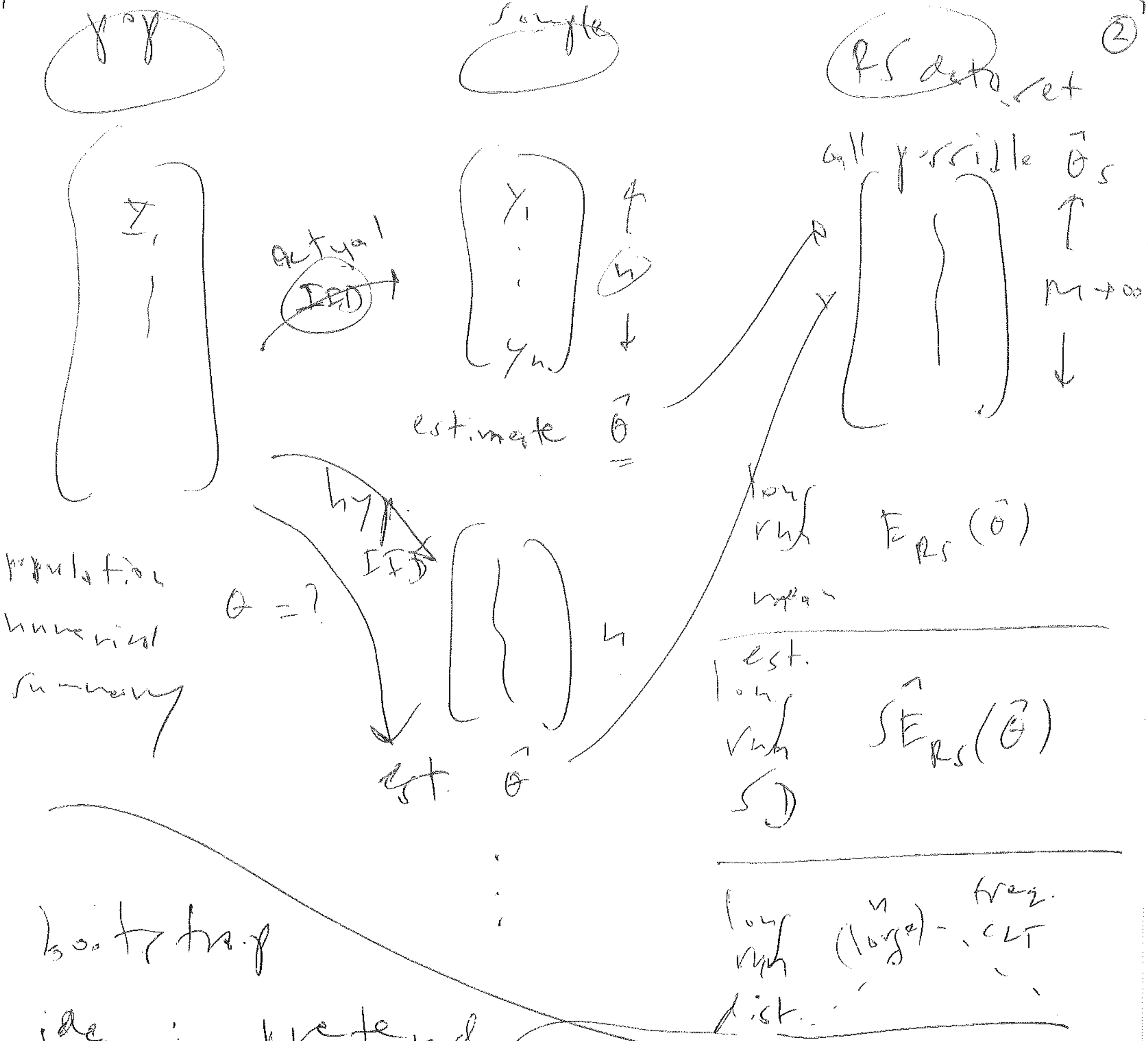


bootstrap estimate of $\theta = \frac{1}{m} \sum_{i=1}^{n} \hat{\theta}_i^* = \bar{\hat{\theta}}^*$

dist SE SD of the $\hat{\theta}^*$ values

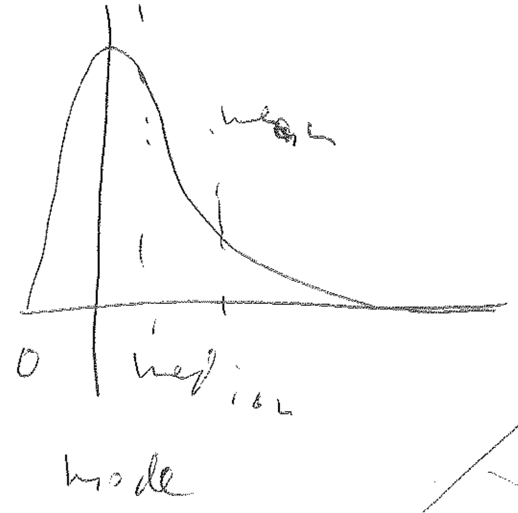


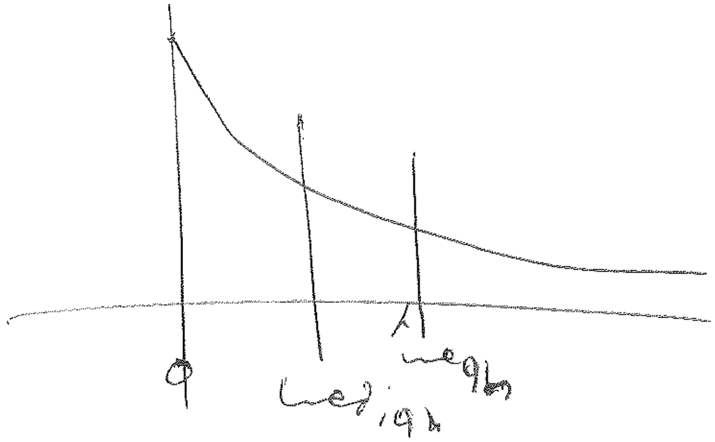
treat this like a post. dist for θ



bootstrap idea: pretend population = sample

bootstrap is a frequentist <u>nonparametric</u> method (no assumptions made about $p(y_i|\theta)$)





de Finetti's thm. for binary outcomes (1931)

binary

$Y_1, Y_2, \ldots Y_n, \ldots$ exch.

$$\left\{ \begin{array}{l} \theta \sim p(\theta) \\ (Y_i \mid \theta\, \mathcal{B}) \overset{IID}{\sim} \text{Bernoulli}(\theta) \\ \quad (i = 1, \ldots, n) \end{array} \right.$$

——— for continuous outcomes (1937)

$Y_1, Y_2, \ldots, Y_n, \ldots \in \mathbb{R}$

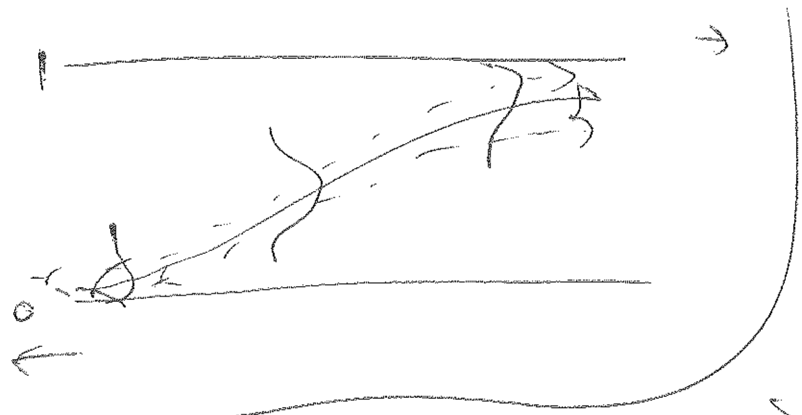
exch. →

$$\left\{ \begin{array}{l} F \sim p(F) \leftarrow \\ (Y_i \mid F\, \mathcal{B}) \overset{IID}{\sim} F \\ \quad (i = 1, \ldots, n) \end{array} \right\}$$

empirical CDF of → Ⓕ

(dist'n on functions)

Bayesian nonparametric model



Ferguson (1970)
Dirichlet Process
(DP) prior

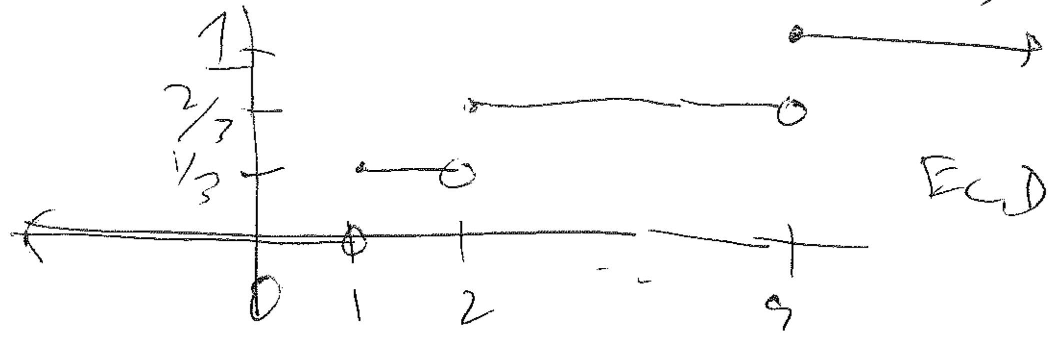
$$\left\{ \begin{array}{l} F \sim DP(\overset{\text{prior sample size}}{\overset{\downarrow}{\alpha}}, F_0) \\ (Y_i \mid F\, \mathcal{B}) \overset{IID}{\sim} F \\ (i = 1, \dots, n) \end{array} \right\} \to$$

$F_0 \leftarrow$ prior estimate of F

$$(\boxed{F} \mid y\, \mathcal{B}(DP)) \sim DP\left(\alpha + n, \frac{\alpha F_0 + n \hat{F}_n}{\alpha + n}\right)$$

$\hat{F}_n$ = empirical CDF of $(y_1, \dots, y_n)$

$$\begin{pmatrix} 1 \\ 2 \\ 9 \end{pmatrix}$$



ECDF

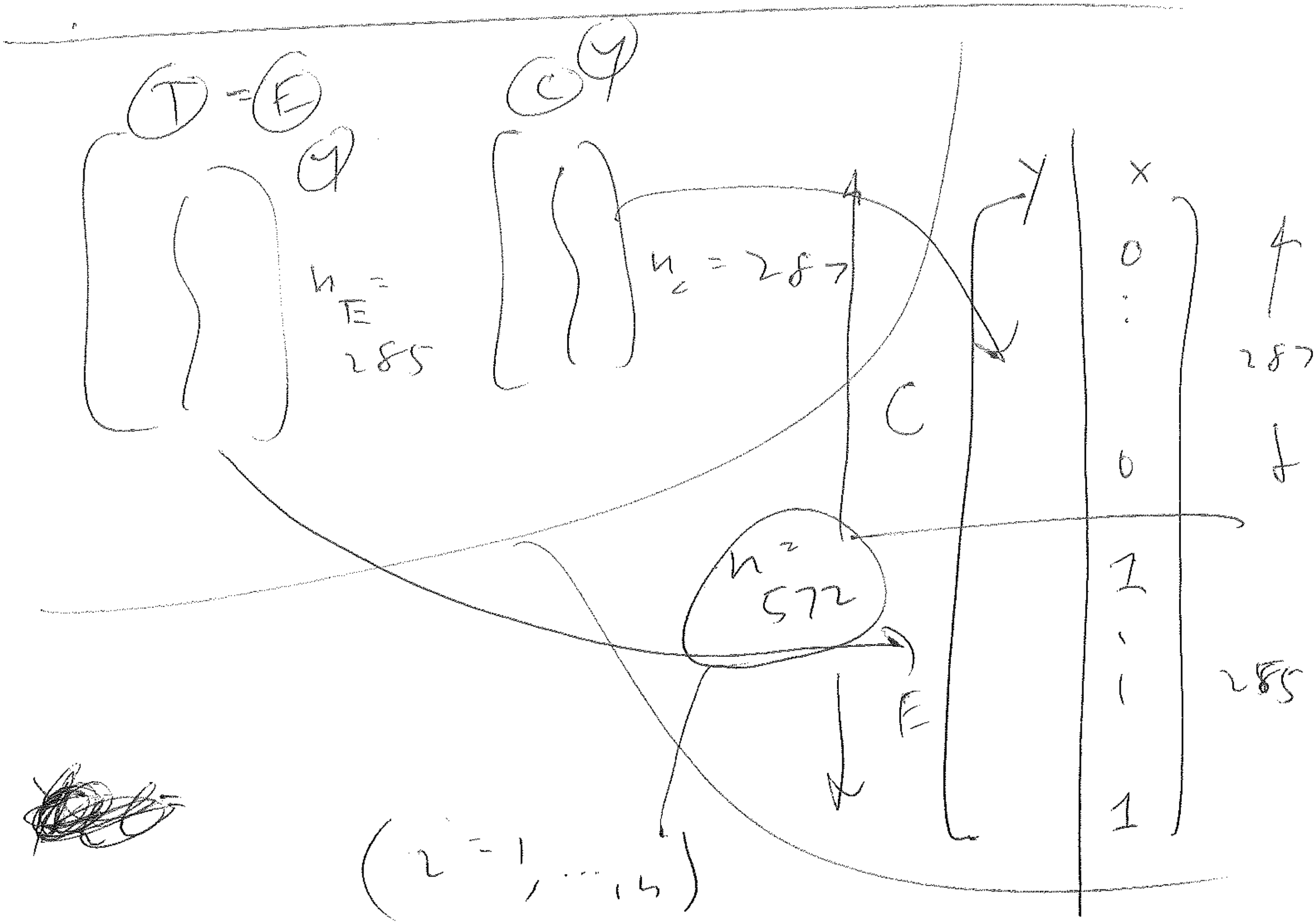
Terenin & Draper (2017)

bootstrap $\equiv DP(0, \cdot)$

$\to (F \mid y\, \mathcal{B}) \sim DP(n, \hat{F}_n)$

⑤

Poisson model $\lambda_C = 1500.2$



$(i = 1, \ldots, n)$

$$(Y_i \mid \lambda_i) \overset{I}{\sim} \text{Poisson}(\lambda_i)$$

$$\log \lambda_i = \boxed{\gamma_0} + \boxed{\gamma_1} x_i$$

$x_i = 0 \rightarrow$

$x_i = 1 \rightarrow \log \lambda_i = \gamma_0 + \gamma_1$ λ_C

$\lambda_E = \lambda_i = e^{\gamma_0 + \gamma_1} = \boxed{e^{\gamma_0}} \cdot e^{\gamma_1}$

$\log \lambda_i = \gamma_0$

$\lambda_C = \lambda_i = e^{\gamma_0}$

$\frac{\lambda_E}{\lambda_C} = e^{\gamma_1}$