Lec07

Monday, January 23, 2023 10:37 AM

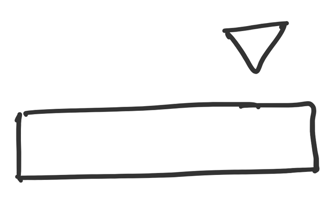
- Assingment due today

Recap

- Hypothesis testing : Random sample method
  (two sample t-test)

  Random design method
  (permutation method)

Paired t-test :



sample i → $y_{1,i}$, $y_{2,i}$

	Tip 1 ,	Tip 2	dent
1	y_{11}	y_{21}	$y_{21} - y_{11}$
2	y_{21}	y_{22}	⋮
3			
⋮	⋮		⋮
n			

$$d_i = y_{2,i} - y_{1,i} \sim N(\mu, \sigma)$$

Null hypothesis : No change between tip 1,2
$\mu = 0$

Aternative : $\mu \neq 0$, we expect some variation

$$t = \frac{\bar{d} - \mu}{s/\sqrt{n}} \qquad s = \frac{\sum (d_i - \bar{d})^2}{n-1}$$

- Randomization for inference : " we looked data we might have if one of our hypothesis is true"

- Bayesian hypothesis testing : our prior knowledge to specify likelihoods of our model and then compute evidences for our model

$$M_0 = N(0, \sigma) \ ; \ M_1 = N(\mu, \sigma)$$

$P(M_0)$, $P(M_1)$
(assume both are equal $P(M_0) = P(M_1) = 1/2$)

$$\underbrace{p(X \mid M)}_{\text{Model evidence}} = P(M)\, P(M \mid X)$$

given data X, likelihood of M being the model to describe it

$M \sim N(\mu, \sigma)$ $P(\mu)$, $P(\sigma)$
$P(\theta = [\mu, \sigma])$

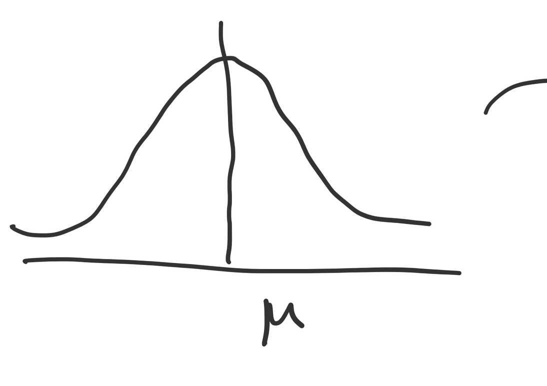
Bayesian model selection

$$\left[\text{Bayes factor} = \frac{P(X \mid M_0)}{P(X \mid M_1)} \right]$$

- Boot strap inference: sampling with replacement

  [10, 12, 14, 20, 20] → mean*

  Sample : [10, 10, 12, 14, 14] → mean^i



→ hypothesis testing

μ

S_M , S_{bM}

↓ bootstrap sample   ↓ bootstrap sample → t-test to do inference.

- wiki (Bootstrap hypothesis testing,
  Bayesian model selection, Bayes factor,
  Bayes information criteria)

- ANOVA (Analysis of Variance)

  → way of performing a series of paired t-tests

  → One-way ANOVA
  two-way ANOVA

- i → $\overset{\times}{e_0}$ e_1 $\overset{\times}{e_2}$ ... e_t

  perform a paired t-test between i (e_0, e_1)

  levels in two ways : Fixed levels
  Random effect levels

- gender (vs) salary
  (independt variable) (dependent variable)

- Null hypothesis : no difference across levels
  $\mu_0, \mu_1, \cdots, \mu_M$

  H_0 : $\mu_0 = \mu_1 = \cdots = \mu_M$

  H_1 : atleast one of μ_i is different irrespective of which level it is

- ANOVA ; $y_{ij} = \alpha + \beta_j + e_{ij}$

  i^{th} subject and undergoing j^{th} treatment

  $e_{ij} \sim N(0, \sigma)$ some variance across the different treatment (S_e^2)

  $\underbrace{\beta_j \neq 0}_{\text{(not purely random sampling)}}$ β_j → How much of a variation the level j contributes to the mean response α

  - sensitivity testing