

# Inclass 19: Posterior Predictive Distribution

[SCS4049] Machine Learning and Data Science

---

Seongsik Park (s.park@dgu.edu)

Department of Artificial Intelligence, Dongguk University

How to treat  $p(x' | x_1, x_2, \dots, x_N)$ ?

- Maximum likelihood estimate

$$\hat{\theta}_{\text{MLE}} = \arg \max p(x_1, x_2, \dots, x_N | \theta) \quad (1)$$

$$p(x' | x_1, x_2, \dots, x_N) \sim p(x' | \hat{\theta}_{\text{MLE}}) \quad (2)$$

- Maximum a posteriori estimate

$$\hat{\theta}_{\text{MAP}} = \arg \max p(\theta | x_1, x_2, \dots, x_N) \quad (3)$$

$$\text{where } p(\theta | x_1, x_2, \dots, x_N) \propto p(x_1, x_2, \dots, x_N | \theta)p(\theta) \quad (4)$$

$$p(x' | x_1, x_2, \dots, x_N) \sim p(x' | \hat{\theta}_{\text{MAP}}) \quad (5)$$

How to treat  $p(x' | x_1, x_2, \dots, x_N)$ ?

- Posterior predictive distribution

$$p(x' | x_1, x_2, \dots, x_N) = \int_{\theta} p(x' | \theta) p(\theta | x_1, x_2, \dots, x_N) d\theta \quad (6)$$