Posterior Predictive Distribution

Oliver Dürr

May 8, 2024

Posterior Predictive (Formal Derivation)

Definition 1: Conditional Distribution

$$p(Y_1, Y_2, Y_3 | X_1, X_2, X_3) = \frac{p(Y_1, Y_2, Y_3, X_1, X_2, X_3)}{p(X_1, X_2, X_3)}$$

Definition 2: Marginal Distribution

$$p(Y_1, Y_2) = \int p(Y_1, Y_2, Y_3)\, dY_3$$

Derivation of Posterior Predictive Distribution $P(Y|D)$

$$P(Y|D) = \frac{P(Y, D)}{P(D)} \quad \text{(by definition)}$$

$$\Rightarrow P(Y|D) = \int \frac{P(Y, D, \theta)}{P(D)}\, d\theta \quad \text{(marginalization over } \theta\text{)}$$

$$= \int P(Y|D, \theta) \frac{P(D|\theta)P(\theta)}{P(D)}\, d\theta \quad \text{(chain rule: } P(Y, D, \theta) = P(Y|D, \theta)P(D|\theta)P(\theta)\text{)}$$

$$= \int P(Y|D, \theta) P(\theta|D)\, d\theta \quad \text{(Bayes' rule for } P(\theta|D)\text{)}$$

(Assuming all data relevant to Y is captured by θ, $P(Y|D, \theta) = P(Y|\theta)$)

$$P(Y|D) = \int P(Y|\theta) P(\theta|D)\, d\theta \quad \text{(posterior predictive distribution)}$$